SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

DELEGATURA PROPIEDAD INDUSTRIAL

División de Nuevas Creaciones

PCT

SOLICITUD

PATENTE DE INVENCION

ENTRADA E I FASE NACIONAL

21. **EXPEDIENTE No** 07-333779

54 **TITULO** COMPOSICIONES Y METODOS PARA EL DIAGNOSTICO Y EL TRATAMIENTO DE TUMORES

51 **CLASIFICACION INTERNACIONAL** A61K 39/395

71 **SOLICITANTE** GENENTECH, INC.

DOMICILIO SAN FRANCISCO, CALIFORNIA, ESTADOS UNIDOS DE AMERICA

74 **APODERADO** JUAN PABLO CADENA SARMIENTO

22 **SANTAFE DE BOGOTA, D C**

FU 05 04 2007

(FORMA P 10)

Oficina de Comunicaciones

PETITORIO

AS: 103.343

Industria y Comercio
SUPERINTENDENCIA

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
Rad: 07-033378- -00000-0000
Fec: 2007-04-02 17:36:24 Dep: 2060 NUECREACIONES
Tra: 11 PATENTEYII
Cla: 379 FASENACIONALI
Act 411 PRESENTACION Folios: 1543

FORMULARIO ÚNICO DE SOLICITUD DE PATENTE
PCT
ENTRADA EN FASE NACIONAL

1 SOLICITUD DE:

☒ Patente de Invención. ☐ Patente de Modelo de Utilidad.

☒ Capítulo I. ☐ Capítulo II.

2 SOLICITANTE (71)	Nombre: **GENENTECH, INC.** Dirección: 1 DNA Way, South San Francisco, CA 94080-4990, US Nacionalidad o Domicilio: Estados Unidos Lugar de Constitución: Estados Unidos Teléfono: Fax: E-mail	**IDENTIFICACION** C.C. ☐ NIT ☐ C.E. ☐ Otro ☐ Cual_____ Número _____
3 REPRESENTANTE O APODERADO	Nombre: JUAN PABLO CADENA SARMIENTO Dirección: Carrera 7 # 71 – 21 Torre B – Piso 4 Teléfono: 744-2200 Fax: 744-2700 E-mail:	**IDENTIFICACION** C.C. ☒ NIT ☐ C.E. ☐ Otro ☐ Cual_____ Número 80.410.337 TP 57492
4 INVENTOR (ES) (72)	**Nombre:** Belinda CAIRNS, Ruihuan CHEN, Anan CHUNTHARAPAI, Gretchen FRANTZ, Kenneth HILLAN, Hartmut KOEPPEN, Heidi PHILLIPS, Paul POLAKIS, Chee Jeraka REED, Chie SAKANAKA, Victoria SMITH, Susan SPENCER, P. Mickey WILLIAMS, Thomas WU, Zemin ZHANG **Dirección:** 129 Paloma Avenue, Burlingame, California 94010, US; 1110 Polynesia Drive, #301, Foster City, CA 94404, US; 826 Ella Drive, Colma, CA 94015, US; 135 San Benito Way, San Francisco, CA 94127, US; 84 Seward Street, San Francisco, CA 94114, US; 1746 Milvia Street, Berkeley, CA 94709, US; 1511 Walnut Drive, Palo Alto, CA 94303, US; 1449 Cortez Avenue, Burlingame, CA 94010, US; 2396 Post Street, San Francisco, CA 94115, US; 1481 8th Avenue, San Francisco, CA 94122, US; 19 Dwight Road, Burlingame, CA 94010, US; 2395 Paradise Drive, Tiburon, CA 94920, US; 509 Alto Avenue, Half Moon Bay, CA 94019, US; 41 Nevada Street, San Francisco, CA 94110, US; 876 Taurus Drive, Foster City, CA 94404, US (Respectivamente) **Nacionalidad o Domicilio:** Estados Unidos de América (Respectivamente)	

5 PCT (86)
Solicitud Internacional No. PCT/US2005/031798 Fecha 07-09-2005
Publicación Internacional No. WO 2006/029183 A3 Fecha 16-03-2006

6 Título (54): COMPOSICIONES Y MÉTODOS PARA EL DIAGNÓSTICO Y EL TRATAMIENTO DE TUMORES

7 Clasificación Internacional (51)

8 Prioridad SI ☒ NO ☐	(33) País de Origen	(32) Fecha	(31) Número de Solicitud
	US	08-09-2004	10/936,626

9 Comprobante de pago No. 07-13,494 **Fecha** 02-04-2007

Instrucciones para el diligenciamiento del presente formulario, ver reverso de esta página.

2020-F06 (01-11-15)

10 **ANEXOS**

- [x] Comprobante de pago de la tasa de presentación de la solicitud. $482.000
- [x] Documento que acredita la existencia y representación legal cuando el solicitante sea persona jurídica.
- [x] Poderes, si fuere el caso. Fotocopia Autenticada
- [ ] Documento de cesión del inventor al solicitante o a su causante.
- [ ] Declaraciones
- [x] Copia de la solicitud internacional. (Resumen, descripción, reivindicaciones, dibujos)
- [x] Traducción de la solicitud internacional al castellano. (Resumen, descripción, reivindicaciones, dibujos)
- [ ] Copia del examen preliminar efectuado por la IPEA.
- [ ] Traducción del examen preliminar efectuado por la IPEA.
- [ ] De ser el caso, copia del contrato de acceso.
- [ ] De ser el caso, documento que acredite la licencia o autorización de uso de conocimientos tradicionales de las comunidades indígenas.
- [ ] De ser el caso, certificado de depósito del material biológico.
- [ ] Arte final 12x12.
- [ ] De ser el caso, información sobre otras solicitudes de patente o títulos obtenidos en el extranjero por el mismo titular o su causante, relacionadas parcial o totalmente con la invención de esta solicitud.
- [x] Otros: Listado de Secuencias
  Tarjeta de Propietarios
  Documentos PCT
  Reivindicaciones Modificadas

11 **FIGURA CARACTERÍSTICA:**

12 Solicito la concesión de la patente,

NOMBRE: JUAN PABLO CADENA SARMIENTO

FIRMA:

C.C. 80.410.337 de Usaquén TP 57492 del C.S.J.

Instrucciones para la presentación de los anexos, ver reverso de esta página.

Fs 102.243

3

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
Rad: 07-033378- -00000-0000
Fec:2007-04-02 17:36:24 Dep: 2020 NUECREACIONES
Tra: 11 PATENTEYR
Ene:370 FASENACIONALI
Anx:11 DIGIFIRMA CION.

NIT : 800176089-2

RECIBO OFICIAL DE CAJA : 07 - 13,494
FECHA : FEBRERO 14 DE 2007

***** C O N S I G N A C I O N *****

DEPOSITANTE	TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAG	Vr. PAGO
BRIGARD & CASTRO LTDA	CONSIGNACION	BANCO POPULAR	050-00110-6	52423538	14/02/2007	17,370,000.

***** C O N C E P T O *****

CANT.	RENTISTICO	CONCEPTO	TOTAL CONCEPTO
1	50005-01-01 SOLICITUDES	1 TRAMITES DE SOL. DE PATENTE DE IN	482,000.00
		TOTAL : $	482,000.

SON: CUATROCIENTOS OCHENTA Y DOS MIL PESOS

RESPONSABLE : _____

RECIBO DE CAJA APLICADO AL EXPEDIENTE No._____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
DIVISION DE NUEVAS CREACIONES

PATENTE [X] PCT/US2005/031798 MODELO DE UTILIDAD [ ] DISEÑO INDUSTRIAL [ ]

SOLICITANTE/S. RAZON SOCIAL O APELLIDOS Y NOMBRES GENENTECH, INC.

DOMICILIO SAN FRANCISCO, CALIFORNIA, ESTADOS UNIDOS DE AMERICA
APODERADO JUAN PABLO CADENA SARMIENTO
TITULO COMPOSICIONES Y METODOS PARA EL DIAGNOSTICO Y EL TRATAMIENTO DE TUMORES

CLASIFICACION
EXPEDIENTE No. FECHA DE SOLICITUD ROLLO MICROFILM
CERTIFICADO No. FECHA DE CONCESION VIGENCIA DESDE HASTA
PRORROGA CAMBIO DE NOMBRE
TRASPASO A
INVENTOR (ES) BERLINDA CAIRNS, RUIHUAN CHEN, ANAN CHUNTHARAPAI, GRETCHEN FRANTZ, KENNETH
OTROS HILLAN, KARTMUT KOEPPEN, HEIDI PHILLIPS, PAUL POLAKIS Y SIETE INVENTORES MAS

OFICINA DE COMUNICACIONES

103.343

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

PCT

(43) International Publication Date
16 March 2006 (16.03.2006)

(10) International Publication Number
WO 2006/029183 A3

(51) International Patent Classification:
C07K 16/30 (2006.01) *A61K 39/395* (2006.01)
A61K 47/48 (2006.01) *A61P 35/00* (2006.01)

(21) International Application Number:
PCT/US2005/031798

(22) International Filing Date:
7 September 2005 (07.09.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/936,626 8 September 2004 (08.09.2004) US

(71) Applicant (*for all designated States except US*): GENENTECH, INC. [US/US]; MS 49, 1 DNA Way, South San Francisco, California 94080-4990 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): CAIRNS, Belinda [GB/US]; 1229 Paloma Avenue, Burlingame, California 94010 (US). CHEN, Rufiusan [CN/US]; 1110 Polynesia Drive, #301, Foster City, CA 94404 (US). CHUNTHARAPAI, Anan [US/US]; 826 Ellis Drive, Colma, CA 94015 (US). FRANTZ, Gretchen [US/US]; 135 San Benito Way, San Francisco, CA 94127 (US). HILLAN, Kenneth [GB/US]; 64 Seward Street, San Francisco, CA 94114 (US). KOEPPEN, Hartmut [DE/US]; 1746 Milvia Street, Berkeley, CA 94709 (US). PHILLIPS, Heidi [US/US]; 1511 Walnut Drive, Palo Alto, CA 94303 (US). POLAKIS, Paul [US/US]; 1449 Cortez Avenue, Burlingame, CA 94010 (US). REED, Chase Janeba [US/US]; 2396 Post Street, San Francisco, CA 94115 (US). SAKANAKA, Chie [JP/US]; 1481 8th Avenue, San Francisco, CA 94122 (US). SMITH, Victoria [AU/US]; 19 Dwight Road, Burlingame, CA 94010 (US). SPENCER, Susan [US/US]; 2395 Paradise Drive, Tiburon, CA 94920 (US). WILLIAMS, P. Mickey [US/US]; 509 Alto Avenue, Half Moon Bay, CA 94019 (US). WU, Thomas [US/US]; 41 Nevada Street, San Francisco, CA 94110 (US). ZHANG, Zemin [US/US]; 876 Taurus Drive, Foster City, CA 94404 (US).

(74) Agents: CONLEY, Deirdre et al.; MS49 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
16 November 2006

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF TUMOR

(57) Abstract: The present invention is directed to compositions of matter useful for the diagnosis and treatment of tumor in mammals and to methods of using those compositions of matter for the same.

WO 2006/029183 A3

103·343

6

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 March 2006 (16.03.2006)

PCT

(10) International Publication Number
WO 2006/029183 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/US2005/031798

(22) International Filing Date:
7 September 2005 (07.09.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/936,626 8 September 2004 (08.09.2004) US

(71) Applicant (*for all designated States except US*): GENENTECH, INC. [US/US]; MS 49, 1 DNA Way, South San Francisco, California 94080-4990 (US).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): CAIRNS, Belinda [GB/US]; 1229 Paloma Avenue, Burlingame, California 94010 (US). CHEN, Ruihuan [CN/US]; 1110 Polynesia Drive, #301, Foster City, CA 94404 (US). CHUNTHARAPAI, Anan [US/US]; 826 Ellis Drive, Colma, CA 94015 (US). FRANTZ, Gretchen [US/US]; 135 San Benito Way, San Francisco, CA 94127 (US). HILLAN, Kenneth [GB/US]; 64 Seward Street, San Francisco, CA 94114 (US). KOEPPEN, Hartmut [DE/US]; 1746 Milvia Street, Berkeley, CA 94709 (US). PHILLIPS, Heidi [US/US]; 1511 Walnut Drive, Palo Alto, CA 94303 (US). POLAKIS, Paul [US/US]; 1449 Cortez Avenue, Burlingame, CA 94010 (US). REED, Chas Jancka [US/US]; 2396 Post Street, San Francisco, CA 94115 (US). SAKANAKA, Chie [JP/US]; 1481 8th Avenue, San Francisco, CA 94122 (US). SMITH, Victoria [AU/US]; 19 Dwight Road, Burlingame, CA 94010 (US). SPENCER, Susan [US/US]; 2395 Paradise Drive, Tiburon, CA 94920 (US). WILLIAMS, P. Mickey [US/US]; 509 Alto Avenue, Half Moon Bay, CA 94019 (US). WU, Thomas [US/US]; 41 Nevada Street, San Francisco, CA 94110 (US). ZHANG, Zemin [US/US]; 876 Taurus Drive, Foster City, CA 94404 (US).

(74) Agents: CONLEY, Deirdre et al.; MS49 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF TUMOR

(57) Abstract: The present invention is directed to compositions of matter useful for the diagnosis and treatment of tumor in mammals and to methods of using those compositions of matter for the same.

WO 2006/029183 A2

COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF TUMOR

RELATED APPLICATIONS

This application is a continuation-in-part of, and claims priority under 35 U.S.C. §120 to, U.S. Application Serial No. 10/177,488 filed 6/19/2002, which claims priority under 35 U.S.C. §119(e) to, U.S. Provisional Application Serial Nos. 60/299,500 filed 6/20/2001, 60/301,880 filed 6/29/2001, and 60/323,268 filed 9/18/2001, the disclosures of which are herein incorporated by reference.

FIELD OF THE INVENTION

The present invention is directed to compositions of matter useful for the diagnosis and treatment of tumor in mammals and to methods of using those compositions of matter for the same.

BACKGROUND OF THE INVENTION

Malignant tumors (cancers) are the second leading cause of death in the United States, after heart disease (Boring et al., *CA Cancel J. Clin.* 43:7 (1993)). Cancer is characterized by the increase in the number of abnormal, or neoplastic, cells derived from a normal tissue which proliferate to form a tumor mass, the invasion of adjacent tissues by these neoplastic tumor cells, and the generation of malignant cells which eventually spread via the blood or lymphatic system to regional lymph nodes and to distant sites via a process called metastasis. In a cancerous state, a cell proliferates under conditions in which normal cells would not grow. Cancer manifests itself in a wide variety of forms, characterized by different degrees of invasiveness and aggressiveness.

In attempts to discover effective cellular targets for cancer diagnosis and therapy, researchers have sought to identify transmembrane or otherwise membrane-associated polypeptides that are specifically expressed on the surface of one or more particular type(s) of cancer cell as compared to on one or more normal non-cancerous cell(s). Often, such membrane-associated polypeptides are more abundantly expressed on the surface of the cancer cells as compared to on the surface of the non-cancerous cells. The identification of such tumor-associated cell surface antigen polypeptides has given rise to the ability to specifically target cancer cells for destruction via antibody-based therapies. In this regard, it is noted that antibody-based therapy has proved very effective in the treatment of certain cancers. For example, HERCEPTIN® and RITUXAN® (both from Genentech Inc., South San Francisco, California) are antibodies that have been used successfully to treat breast cancer and non-Hodgkin's lymphoma, respectively. More specifically, HERCEPTIN® is a recombinant DNA-derived humanized monoclonal antibody that selectively binds to the extracellular domain of the human epidermal growth factor receptor 2 (HER2) proto-oncogene. HER2 protein overexpression is observed in 25-30% of primary breast cancers. RITUXAN® is a genetically engineered chimeric murine/human monoclonal antibody directed against the CD20 antigen found on the surface of normal and malignant B lymphocytes. Both these antibodies are recombinantly produced in CHO cells.

In other attempts to discover effective cellular targets for cancer diagnosis and therapy, researchers have sought to identify (1) non-membrane-associated polypeptides that are specifically produced by one or more particular type(s) of cancer cell(s) as compared to by one or more particular type(s) of non-cancerous normal cell(s), (2) polypeptides that are produced by cancer cells at an expression level that is significantly higher than that of one or more normal non-cancerous cell(s), or (3) polypeptides whose expression is specifically limited to only a single (or very limited number of different) tissue type(s) in both the cancerous and

(e.g., normal prostate and prostate tumor tissue). Such polypeptides may remain intracellularly located or may be secreted by the cancer cell. Moreover, such polypeptides may be expressed not by the cancer cell itself, but rather by cells which produce and/or secrete polypeptides having a potentiating or growth-enhancing effect on cancer cells. Such secreted polypeptides are often proteins that provide cancer cells with a growth advantage over normal cells and include such things as, for example, angiogenic factors, cellular adhesion factors, growth factors, and the like. Identification of antagonists of such non-membrane associated polypeptides would be expected to serve as effective therapeutic agents for the treatment of such cancers. Furthermore, identification of the expression pattern of such polypeptides would be useful for the diagnosis and treatment of particular cancers in mammals.

The ability to modulate gene expression in a mammal is still difficult. Traditionally, it has been done using tools such viral vectors that will express a polypeptide that the host is lacking. The ability to introduce a viral vector into a host that will reduce gene expression has not been perfected. Repression of gene expression has traditionally been done in mammals in a "knockout" fashion, where the test mammal, usually a mouse, has the gene ablated or "knocked out" through the technique of homolgous recombination. This technique in mice is slow, laborious and difficult, and impractical in humans. Therefore, Applicants turned to a vector system which will express an interfering RNA (si RNA) to reduce gene expression.

siRNAs have proven useful as a tool in studies of modulating gene expression where traditional antagonists such as small molecules or antibodies have failed. (Shi Y., *Trends in Genetics* 19(1):9-12 (2003)). In vitro synthesized, double stranded RNAs that are 21 to 23 nucleotides in length can act as interfering RNAs (iRNAs) and can specifically inhibit gene expression (Fire A., *Trends in Genetics* 391; 806-810 (1999)). These iRNAs act by mediating degradation of their target RNAs. Since they are under 30 nuclotides in length, however they do not trigger a cell antiviral defense mechanism. Such mechanisms include interferon production, and a general shutdown of host cell protein synthesis. Practically, siRNAs can by synthesized and then cloned into DNA vectors. Such vectors can be transfected and made to express the siRNA at high levels and/or in a tissue specific manner. The high level of siRNA expression is used to "knockdown" or significantly reduce the amount of protein produced in a cell, and thus it is useful in experiments where overexpression of a protein is believed to be linked to a disorders such as cancer. Despite advances in mammalian cancer therapy, there is a great need for therapeutic agents capable of effectively inhibiting neoplastic cell growth through reduction of gene expression. Accordingly, it is an objective of the present invention to identify a system that will modulate gene expression.

Improving the delivery of drugs and other agentsto target cells, tissues and tumors to achieve maximal efficacy and minimal toxicity has been the focus of considerable research for many years. Though many attempts have been made to develop effective methods for importing biologically active molecules into cells, both *vivo* and *in vitro*, none has proved to be entirely satisfactory. Optimizing the association of the drug with its intracellular target, while minimizing intercellular redistribution of the drug, e.g. to neighboring cells, is often difficult or inefficient.

Most agents currently administered to a patient parenterally are not targeted, resulting in systemic delivery of the agent to cells and tissues of the body where it is unnecessary, and often undesirable. This may

result in adverse drug side effects, and often limits the dose of a drug (e.g., chemotherapeutic (anti-cancer), cytotoxic, enzyme inhibitor agents and antiviral or antimicrobial drugs) that can be administered. By comparison, although oral administration of drugs is considered to be a convenient and economical mode of administration, it shares the same concerns of non-specific toxicity to unaffected cells once the drug has been absorbed into the systemic circulation. Further complications involve problems with oral bioavailability and residence of drug in the gut leading to additional exposure of gut to the drug and hence risk of gut toxicities. Accordingly, a major goal has been to develop methods for specifically targeting agents to cells and tissues. The benefits of such treatment include avoiding the general physiological effects of inappropriate delivery of such agents to other cells and tissues, such as uninfected cells. Intracellular targeting may be achieved by methods, compounds and formulations which allow accumulation or retention of biologically active agents, i.e. active metabolites, inside cells.

Monoclonal antibody therapy has been established for the targeted treatment of patients with cancer, immunological and angiogenic disorders.

There are a variety of useful toxins available. For example, the auristatin peptides, auristain E (AE) and monomethylauristatin (MMAE), synthetic analogs of dolastatin, were conjugated to: (i) chimeric monoclonal antibodies cBR96 (specific to Lewis Y on carcinomas); (ii) cAC10 which is specific to CD30 on hematological malignancies (Klussman, et al (2004), Bioconjugate Chemistry 15(4):765-773; Doronina et al (2003) Nature Biotechnology 21(7):778-784; "Monomethylvaline Compounds Capable of Conjugation to Ligands"; Francisco et al (2003) Blood 102(4):1458-1465; US 2004/0018194; (iii) anti-CD20 antibodies such as rituxan (WO 04/032828) for the treatment of CD20-expressing cancers and immune disorders; (iv) anti-EphB2 antibodies 2H9 and anti-IL-8 for treatment of colorectal cancer (Mao, et al (2004) Cancer Research 64(3):781-788); (v) E-selectin antibody (Bhaskar et al (2003) Cancer Res. 63:6387-6394); and (vi) other anti-CD30 antibodies (WO 03/043583). Monomethylauristatin (MMAE) has also been conjugated to 2H9, an antibody against EphB2R which is a type 1 TM tyrosine kinase receptor with close homology between mouse and human, and is over-expressed in colorectal cancer cells (Mao et al (2004) Cancer Res. 64:781-788).

Monomethylauristatin MMAF, a variant of auristatin E (MMAE) with a phenylalanine at the C-terminus (US 5767237; US 6124431), has been reported to be less potent than MMAE, but more potent when conjugated to monoclonal antibodies (Senter et al, Proceedings fo the American Association for Cancer Research, Volume 45, Abstract Number 623, presented March 28, 2004). Auristatin F phenylene diamin(AFP); a phenylalanine variant of MMAE was linked to an anti-CD70 mAb, 1F6, through the C-terminus of 1F6 via a phenylene diamine spacer (Law et al, Proceedings of the American Association for Cancer Research, Volume 45, Abstract Number 625, presented March 28, 2004)

Despite the above identified advances in mammalian cancer therapy, there is a great need for additional diagnostic and therapeutic agents capable of detecting the presence of tumor in a mammal and for effectively inhibiting neoplastic cell growth, respectively. Accordingly, it is an objective of the present invention to identify: (1) cell membrane-associated polypeptides that are more abundantly expressed on one or more type(s) of cancer cell(s) as compared to on normal cells or on other different cancer cells, (2) non-membrane-associated

polypeptides that are specifically produced by one or more particular type(s) of cancer cell(s) (or by other cells that produce polypeptides having a potentiating effect on the growth of cancer cells) as compared to by one or more particular type(s) of non-cancerous normal cell(s), (3) non-membrane-associated polypeptides that are produced by cancer cells at an expression level that is significantly higher than that of one or more normal non-cancerous cell(s), or (4) polypeptides whose expression is specifically limited to only a single (or very limited number of different) tissue type(s) in both a cancerous and non-cancerous state (e.g., normal prostate and prostate tumor tissue), and to use those polypeptides, and their encoding nucleic acids, to produce compositions of matter useful in the therapeutic treatment and diagnostic detection of cancer in mammals. It is also an objective of the present invention to identify cell membrane-associated, secreted or intracellular polypeptides whose expression is limited to a single or very limited number of tissues, and to use those polypeptides, and their encoding nucleic acids, to produce compositions of matter useful in the therapeutic treatment and diagnostic detection of cancer in mammals.

SUMMARY OF THE INVENTION

A. Embodiments

In the present specification, Applicants describe for the first time the identification of various cellular polypeptides (and their encoding nucleic acids or fragments thereof) which are expressed to a greater degree on the surface of or by one or more types of cancer cell(s) as compared to on the surface of or by one or more types of normal non-cancer cells. Alternatively, such polypeptides are expressed by cells which produce and/or secrete polypeptides having a potentiating or growth-enhancing effect on cancer cells. Again alternatively, such polypeptides may not be overexpressed by tumor cells as compared to normal cells of the same tissue type, but rather may be specifically expressed by both tumor cells and normal cells of only a single or very limited number of tissue types (preferably tissues which are not essential for life, e.g., prostate, etc.). All of the above polypeptides are herein referred to as Tumor-associated Antigenic Target polypeptides ("TAT" polypeptides) and are expected to serve as effective targets for cancer therapy and diagnosis in mammals.

Accordingly, in one embodiment of the present invention, the invention provides an isolated nucleic acid molecule having a nucleotide sequence that encodes a tumor-associated antigenic target polypeptide or fragment thereof (a "TAT" polypeptide).

In certain aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% nucleic acid sequence identity, to (a) a DNA molecule encoding a full-length TAT polypeptide having an amino acid sequence as disclosed herein, a TAT polypeptide amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane TAT polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length TAT polypeptide amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In other aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least

about 80% nucleic acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% nucleic acid sequence identity, to (a) a DNA molecule comprising the coding sequence of a full-length TAT polypeptide cDNA as disclosed herein, the coding sequence of a TAT polypeptide lacking the signal peptide as disclosed herein, the coding sequence of an extracellular domain of a transmembrane TAT polypeptide, with or without the signal peptide, as disclosed herein or the coding sequence of any other specifically defined fragment of the full-length TAT polypeptide amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In further aspects, the invention concerns an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% nucleic acid sequence identity, to (a) a DNA molecule that encodes the same mature polypeptide encoded by the full-length coding region of any of the human protein cDNAs deposited with the ATCC as disclosed herein, or (b) the complement of the DNA molecule of (a).

Another aspect of the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a TAT polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain(s) of such polypeptide(s) are disclosed herein. Therefore, soluble extracellular domains of the herein described TAT polypeptides are contemplated.

In other aspects, the present invention is directed to isolated nucleic acid molecules which hybridize to (a) a nucleotide sequence encoding a TAT polypeptide having a full-length amino acid sequence as disclosed herein, a TAT polypeptide amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane TAT polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length TAT polypeptide amino acid sequence as disclosed herein, or (b) the complement of the nucleotide sequence of (a). In this regard, an embodiment of the present invention is directed to fragments of a full-length TAT polypeptide coding sequence, or the complement thereof, as disclosed herein, that may find use as, for example, hybridization probes useful as, for example, diagnostic probes, antisense oligonucleotide probes, or for encoding fragments of a full-length TAT polypeptide that may optionally encode a polypeptide comprising a binding site for an anti-TAT polypeptide antibody, a TAT binding oligopeptide or other small organic molecule that binds to a TAT polypeptide. Such nucleic acid fragments are usually at least about 5 nucleotides in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, or 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length. It is noted

that novel fragments of a TAT polypeptide-encoding nucleotide sequence may be determined in a routine manner by aligning the TAT polypeptide-encoding nucleotide sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which TAT polypeptide-encoding nucleotide sequence fragment(s) are novel. All of such novel fragments of TAT polypeptide-encoding nucleotide sequences are contemplated herein. Also contemplated are the TAT polypeptide fragments encoded by these nucleotide molecule fragments, preferably those TAT polypeptide fragments that comprise a binding site for an anti-TAT antibody, a TAT binding oligopeptide or other small organic molecule that binds to a TAT polypeptide.

In another embodiment, the invention provides isolated TAT polypeptides encoded by any of the isolated nucleic acid sequences hereinabove identified.

In a certain aspect, the invention concerns an isolated TAT polypeptide, comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% amino acid sequence identity, to a TAT polypeptide having a full-length amino acid sequence as disclosed herein, a TAT polypeptide amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane TAT polypeptide protein, with or without the signal peptide, as disclosed herein, an amino acid sequence encoded by any of the nucleic acid sequences disclosed herein or any other specifically defined fragment of a full-length TAT polypeptide amino acid sequence as disclosed herein.

In a further aspect, the invention concerns an isolated TAT polypeptide comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% amino acid sequence identity, to an amino acid sequence encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein.

In a specific aspect, the invention provides an isolated TAT polypeptide without the N-terminal signal sequence and/or without the initiating methionine and is encoded by a nucleotide sequence that encodes such an amino acid sequence as hereinbefore described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the TAT polypeptide and recovering the TAT polypeptide from the cell culture.

Another aspect of the invention provides an isolated TAT polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the TAT polypeptide and recovering the TAT polypeptide from the cell culture.

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described polypeptides. Host cells comprising any such vector are also provided. By way of example, the host cells may be CHO cells, *E. coli* cells, or yeast cells. A process for producing any of the

herein described polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired polypeptide and recovering the desired polypeptide from the cell culture.

In other embodiments, the invention provides isolated chimeric polypeptides comprising any of the herein described TAT polypeptides fused to a heterologous (non-TAT) polypeptide. Example of such chimeric molecules comprise any of the herein described TAT polypeptides fused to a heterologous polypeptide such as, for example, an epitope tag sequence or a Fc region of an immunoglobulin.

In another embodiment, the invention provides an antibody which binds, preferably specifically, to any of the above or below described polypeptides. Optionally, the antibody is a monoclonal antibody, antibody fragment, chimeric antibody, humanized antibody, single-chain antibody or antibody that competitively inhibits the binding of an anti-TAT polypeptide antibody to its respective antigenic epitope. Antibodies of the present invention may optionally be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The antibodies of the present invention may optionally be produced in CHO cells or bacterial cells and preferably induce death of a cell to which they bind. For diagnostic purposes, the antibodies of the present invention may be detectably labeled, attached to a solid support, or the like.

In another embodiment of the present invention comprising antibody conjugated to a cytotoxin, the toxin may be selected from any of the toxins known in the art. In one embodiment, the toxin is linked to the antibody by a peptide or peptidomimetic linker. In one embodiment, the toxin is linked to the antibody via a linker comprising valine-citrulline (-vc-). In one embodiment the toxin is MMAE. In one embodiment the toxin is MMAE. In one embodiment, the toxin is an auristatin peptide.

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described antibodies. Host cell comprising any such vector are also provided. By way of example, the host cells may be CHO cells,*E. coli* cells, or yeast cells. A process for producing any of the herein described antibodies is further provided and comprises culturing host cells under conditions suitable for expression of the desired antibody and recovering the desired antibody from the cell culture.

In another embodiment, the invention provides oligopeptides ("TAT binding oligopeptides") which bind, preferably specifically, to any of the above or below described TAT polypeptides. Optionally, the TAT binding oligopeptides of the present invention may be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The TAT binding oligopeptides of the present invention may optionally be produced in CHO cells or bacterial cells and preferably induce death of a cell to which they bind. For diagnostic purposes, the TAT binding oligopeptides of the present invention may be detectably labeled, attached to a solid support, or the like.

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described TAT binding oligopeptides. Host cell comprising any such vector are also provided. By way of example, the host cells may be CHO cells,*E. coli* cells, or yeast cells. A process for

producing any of the herein described TAT binding oligopeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired oligopeptide and recovering the desired oligopeptide from the cell culture.

In another embodiment, the invention provides small organic molecules ("TAT binding organic molecules") which bind, preferably specifically, to any of the above or below described TAT polypeptides. Optionally, the TAT binding organic molecules of the present invention may be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The TAT binding organic molecules of the present invention preferably induce death of a cell to which they bind. For diagnostic purposes, the TAT binding organic molecules of the present invention may be detectably labeled, attached to a solid support, or the like.

In a still further embodiment, the invention concerns a composition of matter comprising a TAT polypeptide as described herein, a chimeric TAT polypeptide as described herein, an anti-TAT antibody as described herein, a TAT binding oligopeptide as described herein, or a TAT binding organic molecule as described herein, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

In a still further embodiment, the invention concerns a composition of matter comprising a TAT polypeptide as described herein, a chimeric TAT polypeptide as described herein, an anti-TAT antibody as described herein, a TAT binding oligopeptide as described herein, a TAT binding interfering RNA (siRNA) or a TAT binding organic molecule as described herein, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier. Specifically the TAT polypeptide is TAT188, also referred to herein as H16 or as TAT188(H16). The anti-TAT188 (also referred to as anti-H16 or anti-TAT188(H16)) antibodies of one embodiment of the invention include without limitation 3B5, 12B9, and 12G12, as described herein. The siRNA of the embodiment includes siTAT188 (also referred to as siH16 or siTAT188(H16 or TAT188 siRNA or H16 siRNA or TAT188(H16) siRNA)) as described herein.

In yet another embodiment, the invention concerns an article of manufacture comprising a container and a composition of matter contained within the container, wherein the composition of matter may comprise a TAT polypeptide as described herein, a chimeric TAT polypeptide as described herein, an anti-TAT antibody as described herein, a TAT binding oligopeptide as described herein, or a TAT binding organic molecule as described herein. The article may further optionally comprise a label affixed to the container, or a package insert included with the container, that refers to the use of the composition of matter for the therapeutic treatment or diagnostic detection of a tumor.

Another embodiment of the present invention is directed to the use of a TAT polypeptide as described herein, a chimeric TAT polypeptide as described herein, an anti-TAT polypeptide antibody as described herein, a TAT binding oligopeptide as described herein, or a TAT binding organic molecule as described herein, for the preparation of a medicament useful in the treatment of a condition which is responsive to the TAT polypeptide, chimeric TAT polypeptide, anti-TAT polypeptide antibody, TAT binding oligopeptide, or TAT binding organic molecule.

B. Additional Embodiments

Another embodiment of the present invention is directed to a method for inhibiting the growth of a cell that expresses a TAT polypeptide, wherein the method comprises contacting the cell with an antibody, an oligopeptide or a small organic molecule that binds to the TAT polypeptide, and wherein the binding of the antibody, oligopeptide or organic molecule to the TAT polypeptide causes inhibition of the growth of the cell expressing the TAT polypeptide. In preferred embodiments, the cell is a cancer cell and binding of the antibody, oligopeptide or organic molecule to the TAT polypeptide causes death of the cell expressing the TAT polypeptide. Optionally, the antibody is a monoclonal antibody, antibody fragment, chimeric antibody, humanized antibody, or single-chain antibody. Antibodies, TAT binding oligopeptides and TAT binding organic molecules employed in the methods of the present invention may optionally be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The antibodies and TAT binding oligopeptides employed in the methods of the present invention may optionally be produced in CHO cells or bacterial cells.

Yet another embodiment of the present invention is directed to a method of therapeutically treating a mammal having a cancerous tumor comprising cells that express a TAT polypeptide, wherein the method comprises administering to the mammal a therapeutically effective amount of an antibody, an oligopeptide or a small organic molecule that binds to the TAT polypeptide, thereby resulting in the effective therapeutic treatment of the tumor. Optionally, the antibody is a monoclonal antibody, antibody fragment, chimeric antibody, humanized antibody, or single-chain antibody. Antibodies, TAT binding oligopeptides and TAT binding organic molecules employed in the methods of the present invention may optionally be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The antibodies and oligopeptides employed in the methods of the present invention may optionally be produced in CHO cells or bacterial cells.

Yet another embodiment of the present invention is directed to a method of determining the presence of a TAT polypeptide in a sample suspected of containing the TAT polypeptide, wherein the method comprises exposing the sample to an antibody, oligopeptide or small organic molecule that binds to the TAT polypeptide and determining binding of the antibody, oligopeptide or organic molecule to the TAT polypeptide in the sample, wherein the presence of such binding is indicative of the presence of the TAT polypeptide in the sample. Optionally, the sample may contain cells (which may be cancer cells) suspected of expressing the TAT polypeptide. The antibody, TAT binding oligopeptide or TAT binding organic molecule employed in the method may optionally be detectably labeled, attached to a solid support, or the like.

A further embodiment of the present invention is directed to a method of diagnosing the presence of a tumor in a mammal, wherein the method comprises detecting the level of expression of a gene encoding a TAT polypeptide (a) in a test sample of tissue cells obtained from said mammal, and (b) in a control sample of known normal non-cancerous cells of the same tissue origin or type, wherein a higher level of expression of

the TAT polypeptide in the test sample, as compared to the control sample, is indicative of the presence of tumor in the mammal from which the test sample was obtained.

Another embodiment of the present invention is directed to a method of diagnosing the presence of a tumor in a mammal, wherein the method comprises (a) contacting a test sample comprising tissue cells obtained from the mammal with an antibody, oligopeptide or small organic molecule that binds to a TAT polypeptide and (b) detecting the formation of a complex between the antibody, oligopeptide or small organic molecule and the TAT polypeptide in the test sample, wherein the formation of a complex is indicative of the presence of a tumor in the mammal. Optionally, the antibody, TAT binding oligopeptide or TAT binding organic molecule employed is detectably labeled, attached to a solid support, or the like, and/or the test sample of tissue cells is obtained from an individual suspected of having a cancerous tumor.

Yet another embodiment of the present invention is directed to a method for treating or preventing a cell proliferative disorder associated with altered, preferably increased, expression or activity of a TAT polypeptide, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a TAT polypeptide. Preferably, the cell proliferative disorder is cancer and the antagonist of the TAT polypeptide is an anti-TAT polypeptide antibody, TAT binding oligopeptide, TAT binding organic molecule or antisense oligonucleotide. Effective treatment or prevention of the cell proliferative disorder may be a result of direct killing or growth inhibition of cells that express a TAT polypeptide or by antagonizing the cell growth potentiating activity of a TAT polypeptide.

Yet another embodiment of the present invention is directed to a method of binding an antibody, oligopeptide or small organic molecule to a cell that expresses a TAT polypeptide, wherein the method comprises contacting a cell that expresses a TAT polypeptide with said antibody, oligopeptide or small organic molecule under conditions which are suitable for binding of the antibody, oligopeptide or small organic molecule to said TAT polypeptide and allowing binding therebetween.

Other embodiments of the present invention are directed to the use of (a) a TAT polypeptide, (b) a nucleic acid encoding a TAT polypeptide or a vector or host cell comprising that nucleic acid, (c) an anti-TAT polypeptide antibody, (d) a TAT-binding oligopeptide, or (e) a TAT-binding small organic molecule in the preparation of a medicament useful for (i) the therapeutic treatment or diagnostic detection of a cancer or tumor, or (ii) the therapeutic treatment or prevention of a cell proliferative disorder.

Another embodiment of the present invention is directed to a method for inhibiting the growth of a cancer cell, wherein the growth of said cancer cell is at least in part dependent upon the growth potentiating effect(s) of a TAT polypeptide (wherein the TAT polypeptide may be expressed either by the cancer cell itself or a cell that produces polypeptide(s) that have a growth potentiating effect on cancer cells), wherein the method comprises contacting the TAT polypeptide with an antibody, an oligopeptide or a small organic molecule that binds to the TAT polypeptide, thereby antagonizing the growth-potentiating activity of the TAT polypeptide and, in turn, inhibiting the growth of the cancer cell. Preferably the growth of the cancer cell is completely inhibited. Even more preferably, binding of the antibody, oligopeptide or small organic molecule to the TAT polypeptide induces the death of the cancer cell. Optionally, the antibody is a monoclonal

antibody, antibody fragment, chimeric antibody, humanized antibody, or single-chain antibody. Antibodies, TAT binding oligopeptides and TAT binding organic molecules employed in the methods of the present invention may optionally be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The antibodies and TAT binding oligopeptides employed in the methods of the present invention may optionally be produced in CHO cells or bacterial cells.

Yet another embodiment of the present invention is directed to a method of therapeutically treating a tumor in a mammal, wherein the growth of said tumor is at least in part dependent upon the growth potentiating effect(s) of a TAT polypeptide, wherein the method comprises administering to the mammal a therapeutically effective amount of an antibody, an oligopeptide or a small organic molecule that binds to the TAT polypeptide, thereby antagonizing the growth potentiating activity of said TAT polypeptide and resulting in the effective therapeutic treatment of the tumor. Optionally, the antibody is a monoclonal antibody, antibody fragment, chimeric antibody, humanized antibody, or single-chain antibody. Antibodies, TAT binding oligopeptides and TAT binding organic molecules employed in the methods of the present invention may optionally be conjugated to a growth inhibitory agent or cytotoxic agent such as a toxin, including, for example, a maytansinoid or calicheamicin, an antibiotic, a radioactive isotope, a nucleolytic enzyme, or the like. The antibodies and oligopeptides employed in the methods of the present invention may optionally be produced in CHO cells or bacterial cells.

C. Further Additional Embodiments

In yet further embodiments, the invention is directed to the following set of potential claims for this application:

1. Isolated nucleic acid having a nucleotide sequence that has at least 80% nucleic acid sequence identity to:

(a) a DNA molecule encoding the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) a DNA molecule encoding the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) a DNA molecule encoding an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) a DNA molecule encoding an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78);

(f) the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(g) the complement of (a), (b), (c), (d), (e) or (f).

2. Isolated nucleic acid having:

(a) a nucleotide sequence that encodes the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) a nucleotide sequence that encodes the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) a nucleotide sequence that encodes an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) a nucleotide sequence that encodes an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78);

(f) the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(g) the complement of (a), (b), (c), (d), (e) or (f).

3. Isolated nucleic acid that hybridizes to:

(a) a nucleic acid that encodes the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) a nucleic acid that encodes the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) a nucleic acid that encodes an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) a nucleic acid that encodes an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78);

(f) the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(g) the complement of (a), (b), (c), (d), (e) or (f).

4. The nucleic acid of Claim 3, wherein the hybridization occurs under stringent conditions.

5. The nucleic acid of Claim 3 which is at least about 5 nucleotides in length.

6. An expression vector comprising the nucleic acid of Claim 1, 2 or 3.

7. The expression vector of Claim 6, wherein said nucleic acid is operably linked to control sequences recognized by a host cell transformed with the vector.

8. A host cell comprising the expression vector of Claim 7.

9. The host cell of Claim 8 which is a CHO cell, an *E. coli* cell or a yeast cell.

10. A process for producing a polypeptide comprising culturing the host cell of Claim 8 under conditions suitable for expression of said polypeptide and recovering said polypeptide from the cell culture.

11. An isolated polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

12. An isolated polypeptide having:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

13. A chimeric polypeptide comprising the polypeptide of Claim 11 or 12 fused to a heterologous polypeptide.

14. The chimeric polypeptide of Claim 13, wherein said heterologous polypeptide is an epitope tag sequence or an Fc region of an immunoglobulin.

15. An isolated antibody that binds to a polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

16. An isolated antibody that binds to a polypeptide having:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

17. The antibody of Claim 15 or 16 which is a monoclonal antibody.

18. The antibody of Claim 15 or 16 which is an antibody fragment.

19. The antibody of Claim 15 or 16 which is a chimeric or a humanized antibody.

20. The antibody of Claim 15 or 16 which is conjugated to a growth inhibitory agent.

21. The antibody of Claim 15 or 16 which is conjugated to a cytotoxic agent.

22. The antibody of Claim 21, wherein the cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

23. The antibody of Claim 21, wherein the cytotoxic agent is a toxin.

24. The antibody of Claim 23, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

25. The antibody of Claim 23, wherein the toxin is a maytansinoid.

26. The antibody of Claim 15 or 16 which is produced in bacteria.

27. The antibody of Claim 15 or 16 which is produced in CHO cells.

28. The antibody of Claim 15 or 16 which induces death of a cell to which it binds.

29. The antibody of Claim 15 or 16 which is detectably labeled.

30. An isolated nucleic acid having a nucleotide sequence that encodes the antibody of Claim 15 or 16.

31. An expression vector comprising the nucleic acid of Claim 30 operably linked to control sequences recognized by a host cell transformed with the vector.

32. A host cell comprising the expression vector of Claim 31.

33. The host cell of Claim 32 which is a CHO cell, an *E. coli* cell or a yeast cell.

34. A process for producing an antibody comprising culturing the host cell of Claim 32 under conditions suitable for expression of said antibody and recovering said antibody from the cell culture.

35. An isolated oligopeptide that binds to a polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

36. An isolated oligopeptide that binds to a polypeptide having:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

37. The oligopeptide of Claim 35 or 36 which is conjugated to a growth inhibitory agent.

38. The oligopeptide of Claim 35 or 36 which is conjugated to a cytotoxic agent.

39. The oligopeptide of Claim 38, wherein the cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

40. The oligopeptide of Claim 38, wherein the cytotoxic agent is a toxin.

41. The oligopeptide of Claim 40, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

42. The oligopeptide of Claim 40, wherein the toxin is a maytansinoid.

43. The oligopeptide of Claim 35 or 36 which induces death of a cell to which it binds.

44. The oligopeptide of Claim 35 or 36 which is detectably labeled.

45. A TAT binding organic molecule that binds to a polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

46. The organic molecule of Claim 45 that binds to a polypeptide having:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

47. The organic molecule of Claim 45 or 46 which is conjugated to a growth inhibitory agent.

48. The organic molecule of Claim 45 or 46 which is conjugated to a cytotoxic agent.

49. The organic molecule of Claim 48, wherein the cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

50. The organic molecule of Claim 48, wherein the cytotoxic agent is a toxin.

51. The organic molecule of Claim 50, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

52. The organic molecule of Claim 50, wherein the toxin is a maytansinoid.

53. The organic molecule of Claim 45 or 46 which induces death of a cell to which it binds.

54. The organic molecule of Claim 45 or 46 which is detectably labeled.

55. A composition of matter comprising:

(a) the polypeptide of Claim 11;

(b) the polypeptide of Claim 12;

(c) the chimeric polypeptide of Claim 13;

(d) the antibody of Claim 15;

(e) the antibody of Claim 16;

(f) the oligopeptide of Claim 35;

(g) the oligopeptide of Claim 36;

(h) the TAT binding organic molecule of Claim 45; or

(i) the TAT binding organic molecule of Claim 46; in combination with a carrier.

56. The composition of matter of Claim 55, wherein said carrier is a pharmaceutically acceptable carrier.

57. An article of manufacture comprising:

(a) a container; and

(b) the composition of matter of Claim 55 contained within said container.

58. The article of manufacture of Claim 57 further comprising a label affixed to said container, or a package insert included with said container, referring to the use of said composition of matter for the therapeutic treatment of or the diagnostic detection of a cancer.

59. A method of inhibiting the growth of a cell that expresses a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising contacting said cell with an antibody, oligopeptide or organic molecule that binds to said protein, the binding of said antibody, oligopeptide or organic molecule to said protein thereby causing an inhibition of growth of said cell.

60. The method of Claim 59, wherein said antibody is a monoclonal antibody.

61. The method of Claim 59, wherein said antibody is an antibody fragment.

62. The method of Claim 59, wherein said antibody is a chimeric or a humanized antibody.

63. The method of Claim 59, wherein said antibody, oligopeptide or organic molecule is conjugated to a growth inhibitory agent.

64. The method of Claim 59, wherein said antibody, oligopeptide or organic molecule is conjugated to a cytotoxic agent.

65. The method of Claim 64, wherein said cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

66. The method of Claim 64, wherein the cytotoxic agent is a toxin.

67. The method of Claim 66, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

68. The method of Claim 66, wherein the toxin is a maytansinoid.

69. The method of Claim 59, wherein said antibody is produced in bacteria.

70. The method of Claim 59, wherein said antibody is produced in CHO cells.

71. The method of Claim 59, wherein said cell is a cancer cell.

72. The method of Claim 71, wherein said cancer cell is further exposed to radiation treatment or a chemotherapeutic agent.

73. The method of Claim 71, wherein said cancer cell is selected from the group consisting of a breast cancer cell, a colorectal cancer cell, a lung cancer cell, an ovarian cancer cell, a central nervous system cancer cell, a liver cancer cell, a bladder cancer cell, a pancreatic cancer cell, a cervical cancer cell, a melanoma cell and a leukemia cell.

74. The method of Claim 71, wherein said protein is more abundantly expressed by said cancer cell as compared to a normal cell of the same tissue origin.

75. The method of Claim 59 which causes the death of said cell.

76. The method of Claim 59, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

77. A method of therapeutically treating a mammal having a cancerous tumor comprising cells that express a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising administering to said mammal a therapeutically effective amount of an antibody, oligopeptide or organic molecule that binds to said protein, thereby effectively treating said mammal.

78. The method of Claim 77, wherein said antibody is a monoclonal antibody.

79. The method of Claim 77, wherein said antibody is an antibody fragment.

80. The method of Claim 77, wherein said antibody is a chimeric or a humanized antibody.

81. The method of Claim 77, wherein said antibody, oligopeptide or organic molecule is conjugated to a growth inhibitory agent.

82. The method of Claim 77, wherein said antibody, oligopeptide or organic molecule is conjugated to a cytotoxic agent.

83. The method of Claim 82, wherein said cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

84. The method of Claim 82, wherein the cytotoxic agent is a toxin.

85. The method of Claim 84, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

86. The method of Claim 84, wherein the toxin is a maytansinoid.

87. The method of Claim 77, wherein said antibody is produced in bacteria.

88. The method of Claim 77, wherein said antibody is produced in CHO cells.

89. The method of Claim 77, wherein said tumor is further exposed to radiation treatment or a chemotherapeutic agent.

90. The method of Claim 77, wherein said tumor is a breast tumor, a colorectal tumor, a lung tumor, an ovarian tumor, a central nervous system tumor, a liver tumor, a bladder tumor, a pancreatic tumor, or a cervical tumor.

91. The method of Claim 77, wherein said protein is more abundantly expressed by the cancerous cells of said tumor as compared to a normal cell of the same tissue origin.

92. The method of Claim 77, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of

Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

93. A method of determining the presence of a protein in a sample suspected of containing said protein, wherein said protein has at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising exposing said sample to an antibody, oligopeptide or organic molecule that binds to said protein and determining binding of said antibody, oligopeptide or organic molecule to said protein in said sample, wherein binding of the antibody, oligopeptide or organic molecule to said protein is indicative of the presence of said protein in said sample.

94. The method of Claim 93, wherein said sample comprises a cell suspected of expressing said protein.

95. The method of Claim 94, wherein said cell is a cancer cell.

96. The method of Claim 93, wherein said antibody, oligopeptide or organic molecule is detectably labeled.

97. The method of Claim 93, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence

shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

98. A method of diagnosing the presence of a tumor in a mammal, said method comprising determining the level of expression of a gene encoding a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), in a test sample of tissue cells obtained from said mammal and in a control sample of known normal cells of the same tissue origin, wherein a higher level of expression of said protein in the test sample, as compared to the control sample, is indicative of the presence of tumor in the mammal from which the test sample was obtained.

99. The method of Claim 98, wherein the step of determining the level of expression of a gene encoding said protein comprises employing an oligonucleotide in an *in situ* hybridization or RT-PCR analysis.

100. The method of Claim 98, wherein the step determining the level of expression of a gene encoding said protein comprises employing an antibody in an immunohistochemistry or Western blot analysis.

101. The method of Claim 98, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

102. A method of diagnosing the presence of a tumor in a mammal, said method comprising contacting a test sample of tissue cells obtained from said mammal with an antibody, oligopeptide or organic molecule that binds to a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), and detecting the formation of a complex between said antibody, oligopeptide or organic molecule and said protein in the test sample, wherein the formation of a complex is indicative of the presence of a tumor in said mammal.

103. The method of Claim 102, wherein said antibody, oligopeptide or organic molecule is detectably labeled.

104. The method of Claim 102, wherein said test sample of tissue cells is obtained from an individual suspected of having a cancerous tumor.

105. The method of Claim 102, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

106. A method for treating or preventing a cell proliferative disorder associated with increased expression or activity of a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID

NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising administering to a subject in need of such treatment an effective amount of an antagonist of said protein, thereby effectively treating or preventing said cell proliferative disorder.

107. The method of Claim 106, wherein said cell proliferative disorder is cancer.

108. The method of Claim 106, wherein said antagonist is an anti-TAT polypeptide antibody, TAT binding oligopeptide, TAT binding organic molecule or antisense oligonucleotide.

109. A method of binding an antibody, oligopeptide or organic molecule to a cell that expresses a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising contacting said cell with an antibody, oligopeptide or organic molecule that binds to said protein and allowing the binding of the antibody, oligopeptide or organic molecule to said protein to occur, thereby binding said antibody, oligopeptide or organic molecule to said cell.

110. The method of Claim 109, wherein said antibody is a monoclonal antibody.

111. The method of Claim 109, wherein said antibody is an antibody fragment.

112. The method of Claim 109, wherein said antibody is a chimeric or a humanized antibody.

113. The method of Claim 109, wherein said antibody, oligopeptide or organic molecule is conjugated to a growth inhibitory agent.

114. The method of Claim 109, wherein said antibody, oligopeptide or organic molecule is conjugated to a cytotoxic agent.

115. The method of Claim 114, wherein said cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

116. The method of Claim 114, wherein the cytotoxic agent is a toxin.

117. The method of Claim 116, wherein the toxin is selected from the group consisting of

maytansinoid and calicheamicin.

118. The method of Claim 116, wherein the toxin is a maytansinoid.

119. The method of Claim 109, wherein said antibody is produced in bacteria.

120. The method of Claim 109, wherein said antibody is produced in CHO cells.

121. The method of Claim 109, wherein said cell is a cancer cell.

122. The method of Claim 121, wherein said cancer cell is further exposed to radiation treatment or a chemotherapeutic agent.

123. The method of Claim 121, wherein said cancer cell is selected from the group consisting of a breast cancer cell, a colorectal cancer cell, a lung cancer cell, an ovarian cancer cell, a central nervous system cancer cell, a liver cancer cell, a bladder cancer cell, a pancreatic cancer cell, a cervical cancer cell, a melanoma cell and a leukemia cell.

124. The method of Claim 123, wherein said protein is more abundantly expressed by said cancer cell as compared to a normal cell of the same tissue origin..

125. The method of Claim 109 which causes the death of said cell.

126. Use of a nucleic acid as claimed in any of Claims 1 to 5 or 30 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

127. Use of a nucleic acid as claimed in any of Claims 1 to 5 or 30 in the preparation of a medicament for treating a tumor.

128. Use of a nucleic acid as claimed in any of Claims 1 to 5 or 30 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

129. Use of an expression vector as claimed in any of Claims 6, 7 or 31 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

130. Use of an expression vector as claimed in any of Claims 6, 7 or 31 in the preparation of medicament for treating a tumor.

131. Use of an expression vector as claimed in any of Claims 6, 7 or 31 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

132. Use of a host cell as claimed in any of Claims 8, 9, 32, or 33 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

133. Use of a host cell as claimed in any of Claims 8, 9, 32 or 33 in the preparation of a medicament for treating a tumor.

134. Use of a host cell as claimed in any of Claims 8, 9, 32 or 33 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

135. Use of a polypeptide as claimed in any of Claims 11 to 14 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

136. Use of a polypeptide as claimed in any of Claims 11 to 14 in the preparation of a medicament for treating a tumor.

137. Use of a polypeptide as claimed in any of Claims 11 to 14 in the preparation of a

medicament for treatment or prevention of a cell proliferative disorder.

138. Use of an antibody as claimed in any of Claims 15 to 29 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

139. Use of an antibody as claimed in any of Claims 15 to 29 in the preparation of a medicament for treating a tumor.

140. Use of an antibody as claimed in any of Claims 15 to 29 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

141. Use of an oligopeptide as claimed in any of Claims 35 to 44 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

142. Use of an oligopeptide as claimed in any of Claims 35 to 44 in the preparation of a medicament for treating a tumor.

143. Use of an oligopeptide as claimed in any of Claims 35 to 44 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

144. Use of a TAT binding organic molecule as claimed in any of Claims 45 to 54 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

145. Use of a TAT binding organic molecule as claimed in any of Claims 45 to 54 in the preparation of a medicament for treating a tumor.

146. Use of a TAT binding organic molecule as claimed in any of Claims 45 to 54 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

147. Use of a composition of matter as claimed in any of Claims 55 or 56 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

148. Use of a composition of matter as claimed in any of Claims 55 or 56 in the preparation of a medicament for treating a tumor.

149. Use of a composition of matter as claimed in any of Claims 55 or 56 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

150. Use of an article of manufacture as claimed in any of Claims 57 or 58 in the preparation of a medicament for the therapeutic treatment or diagnostic detection of a cancer.

151. Use of an article of manufacture as claimed in any of Claims 57 or 58 in the preparation of a medicament for treating a tumor.

152. Use of an article of manufacture as claimed in any of Claims 57 or 58 in the preparation of a medicament for treatment or prevention of a cell proliferative disorder.

153. A method for inhibiting the growth of a cell, wherein the growth of said cell is at least in part dependent upon a growth potentiating effect of a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising contacting said protein with an antibody, oligopeptide or organic molecule that binds to said protein, there by inhibiting the growth of said cell.

154. The method of Claim 153, wherein said cell is a cancer cell.

155. The method of Claim 153, wherein said protein is expressed by said cell.

156. The method of Claim 153, wherein the binding of said antibody, oligopeptide or organic molecule to said protein antagonizes a cell growth-potentiating activity of said protein.

157. The method of Claim 153, wherein the binding of said antibody, oligopeptide or organic molecule to said protein induces the death of said cell.

158. The method of Claim 153, wherein said antibody is a monoclonal antibody.

159. The method of Claim 153, wherein said antibody is an antibody fragment.

160. The method of Claim 153, wherein said antibody is a chimeric or a humanized antibody.

161. The method of Claim 153, wherein said antibody, oligopeptide or organic molecule is conjugated to a growth inhibitory agent.

162. The method of Claim 153, wherein said antibody, oligopeptide or organic molecule is conjugated to a cytotoxic agent.

163. The method of Claim 162, wherein said cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

164. The method of Claim 162, wherein the cytotoxic agent is a toxin.

165. The method of Claim 164, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

166. The method of Claim 164, wherein the toxin is a maytansinoid.

167. The method of Claim 153, wherein said antibody is produced in bacteria.

168. The method of Claim 153, wherein said antibody is produced in CHO cells.

169. The method of Claim 153, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

170. A method of therapeutically treating a tumor in a mammal, wherein the growth of said tumor is at least in part dependant upon a growth potentiating effect of a protein having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78), said method comprising contacting said protein with an antibody, oligopeptide or organic molecule that binds to said protein, thereby effectively treating said tumor.

171. The method of Claim 170, wherein said protein is expressed by cells of said tumor.

172. The method of Claim 170, wherein the binding of said antibody, oligopeptide or organic molecule to said protein antagonizes a cell growth-potentiating activity of said protein.

173. The method of Claim 170, wherein said antibody is a monoclonal antibody.

174. The method of Claim 170, wherein said antibody is an antibody fragment.

175. The method of Claim 170, wherein said antibody is a chimeric or a humanized antibody.

176. The method of Claim 170, wherein said antibody, oligopeptide or organic molecule is conjugated to a growth inhibitory agent.

177. The method of Claim 170, wherein said antibody, oligopeptide or organic molecule is conjugated to a cytotoxic agent.

178. The method of Claim 177, wherein said cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

179. The method of Claim 177, wherein the cytotoxic agent is a toxin.

180. The method of Claim 179, wherein the toxin is selected from the group consisting of maytansinoid and calicheamicin.

181. The method of Claim 179, wherein the toxin is a maytansinoid.

182. The method of Claim 170, wherein said antibody is produced in bacteria.

183. The method of Claim 170, wherein said antibody is produced in CHO cells.

184. The method of Claim 170, wherein said protein has:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

185. An isolated antibody that binds to a polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

186. An isolated antibody that binds to a polypeptide having:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

187. The antibody of Claim 185 which is a monoclonal antibody.

188. The antibody of Claim 185 which is an antibody fragment.

189. The antibody of Claim 185 which is a chimeric or a humanized antibody.

190. The antibody of Claim 185 which is conjugated to a growth inhibitory agent.

191. The antibody of Claim 185 which is conjugated to a cytotoxic agent.

192. The antibody of Claim 191, wherein the cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

193. The antibody of Claim 191, wherein the cytotoxic agent is a toxin.

194. The antibody of Claim 193, wherein the toxin is selected from the group consisting of auristatin, maytansinoid and calicheamicin.

195. The antibody of Claim 193, wherein the toxin is a maytansinoid.

196. The antibody of Claim 185 which is produced in bacteria.

197. The antibody of Claim 185 which is produced in CHO cells.

198. The antibody of Claim 185 which induces death of a cell to which it binds.

199. The antibody of Claim 185 which is detectably labeled.

200. The antibody of Claim 193, wherein the toxin is selected from the group consisting of MMAE (mono-methyl auristatin E), MMAF, (auristatin E valeryl benzylhydrazone), and AFP (Auristatin F phenylene diamine).

201. The antibody of claim 193, wherein the toxin is covalently attached to the antibody by a linker.

202. The antibody of claim 201, wherein the linker is selected from the group consisting of maleimidocaproyl (MC), valine-citrulline (val-cit, vc), citrulline (2-amino-5-ureido pentanoic acid), PAB (p-aminobenzylcarbamoyl), Me (N-methyl-valine citrulline), MC(PEG)6-OH (maleimidocaproyl- polyethylene glycol), SPP (N-Succinimidyl 4-(2-pyridylthio) pentanoate), SMCC (N-Succinimidyl 4-(N-maleimidomethyl) cyclohexane-1 carboxylate), and MC-vc-PAB.

203. The antibody of claim 200, wherein the toxin is MMAE.

204. The antibody of claim 200, wherein the toxin is MMAF.

205. The antibody of claim 202, wherein the linker is MC-vc-PAB and the toxin is MMAE or MMAF.

206. The antibody of any one of claims 185-205, wherein the antibody binds the TAT188 polypeptide.

207. The antibody of claim 206, wherein the antibody inhibits proliferation or promotes cell death of a cell expressing TAT188.

208. The antibody of claim 207, wherein the cell is a cancer cell.

209. The antibody of claim 208, where the cancer cell is selected from the group of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

210. An isolated antibody that competes with binding to the epitopes of TAT188 polypeptide bound by an antibody produced by the hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195).

211. An isolated antibody having the biological activity of an antibody produced by the hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195), wherein the biological activity is inhibition of cell proliferation or promotion of cell death in a cell expressing TAT188.

212. An isolated antibody comprising in a corresponding complementary determining region (CDR) an amino acid sequence having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%,

95%, 96%, 97%, 98%, 99% or 100% of the amino acid sequence of at least 1, 2, 3, 4, 5, or 6 of the CDR(s) of the antibody produced by a hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195).

213. The antibody of any one of claims 185-205, wherein the antibody comprises in a corresponding complementary determining region (CDR) an amino acid sequence having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid sequence of at least 1, 2, 3, 4, 5, or 6 of the CDR(s) of the antibody produced by a hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195).

214. The antibody of any one of claims 185-205, wherein the antibody exhibits the biological activity of an antibody produced by the hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195), wherein the biological activity is inhibition of cell proliferation or promotion of cell death in a cell expressing TAT188.

215. The antibody of claim 214, wherein the cell is a cancer cell.

216. The antibody of claim 215, wherein the cancer cell is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

217. A method of inhibiting growth of a cell expressing TAT188, the method comprising contacting the cell with an antibody of any one of claims 185-205.

218. The method of claim 217, wherein the cell is a cancer cell.

219. The method of claim 218, wherein the cancer cell is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

220. The method of claim 219, wherein the cancer cell is a mammalian cell.

221. The method of claim 220, wherein the mammalian cell is a human cell.

222. A method of inhibiting growth of a cell expressing TAT188, the method comprising contacting the cell with an antibody of any one of claims 185-205, wherein the antibody comprises in a corresponding complementary determining region (CDR) an amino acid sequence having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid sequence of at least 1, 2, 3, 4, 5, or 6 of the CDR(s) of the antibody produced by a hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195).

223. A method of detecting the level of TAT188 polypeptide expressed in a test cell relative to a control cell, the method comprising:

(a) contacting the test cell and the control cell with an isolated anti-TAT188 antibody of claim 210;

(b) detecting binding of the antibody; and

(c) determining the relative binding of the antibody to the test and control cell.

224. The method of claim 223, wherein the test cell and control cell are lysed .

225. The method of claim 223, wherein the test cell is in a tissue.

226. The method of claim 225, wherein the tissue is a tumor tissue.

227. The method of claim 226, wherein the tumor tissue is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

228. A method of detecting the level of TAT188 polypeptide or a polypeptide having at least 80% sequence identity to the amino acid sequence shown in Figure 115 (SEQ ID NO:115) in a test cell relative to a control cell,, the method comprising:

(a) contacting the test cell and the control cell with an isolated antibody of any one of claims 185-189, 196-199, and 210-212;

(b) detecting binding of the antibody; and

(c) determining the relative binding of the antibody to the test and control cell.

229. The method of claim 228, wherein the level of TAT188 polypeptide in the test cell is greater than that in the control cell.

230. The method of claim 229, wherein the method diagnoses cancer in a tissue containing or having contained the test cell.

231. A TAT binding interfering RNA (siRNA) which binds to a nucleic acid having at least 80% sequence identity to:
(a) a nucleotide sequence shown in Figure 37 (SEQ ID NO:37; and
(b) the complement of (a), wherein the siRNA reduces expression of TAT188.

232. An expression vector comprising the siRNA of Claim 231.

233. The expression vector of Claim 232, wherein said siRNA is operably linked to control sequences recognized by a host cell transfected with the vector.

234. A host cell comprising the expression vector of Claim 233.

235. A composition of matter comprising:
(a) the antibody of Claim 185, or
(b) the siRNA of claim 231,
in combination with a carrier.

236. The composition of matter of Claim 235, wherein said carrier is a pharmaceutically acceptable carrier.

237. An article of manufacture:
(a) a container; and
(b) the composition of matter of Claim 235 contained within said container.

238. The article of manufacture of Claim 237, further comprising a label affixed to said container, or a package insert included with said container, referring to the use of said composition of matter for the therapeutic treatment of or the diagnostic detection of a cancer.

239. A method of inhibiting the growth of a cancer cell that expresses a polypeptide having at least 80% amino acid sequence identity to the amino acid sequence shown in Figure 115 (SEQ ID NO:115), said method comprising contacting said cancer cell with a siRNA that binds to a nucleic acid encoding the amino acid in said cancer cell, thereby inhibiting the growth of said cancer cell.

240. The method of claim 239, wherein the nucleic acid has the sequence shown in Figure 37 (SEQ ID NO:37).

241. The method of Claim 228, wherein the detecting the level of expression of the polypeptide

comprises employing an antibody in an immunohistochemistry analysis.

242. A method for treating or preventing a cell proliferative disorder associated with increased expression or activity of a polypeptide having at least 80% amino acid sequence identity to the amino acid sequence shown in Figure 115 (SEQ ID NO:115), said method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a TAT188 polypeptide.

243. The method of claim 242, wherein said antagonist is an isolated anti-TAT188 polypeptide antibody of any one of claims 1-21 and 23-28.

244. The method of claim 243, wherein the cell proliferative disorder is cancer.

245. The method of claim 244, wherein the cancer is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

Yet further embodiments of the present invention will be evident to the skilled artisan upon a reading of the present specification.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a nucleotide sequence (SEQ ID NO:1) of a TAT161 cDNA, wherein SEQ ID NO:1 is a clone designated herein as "DNA77507".

Figures 2A-B show a nucleotide sequence (SEQ ID NO:2) of a TAT101 cDNA, wherein SEQ ID NO:2 is a clone designated herein as "DNA80894".

Figure 3 shows a nucleotide sequence (SEQ ID NO:3) of a TAT157 cDNA, wherein SEQ ID NO:3 is a clone designated herein as "DNA82343".

Figure 4 shows a nucleotide sequence (SEQ ID NO:4) of a TAT160 cDNA, wherein SEQ ID NO:4 is a clone designated herein as "DNA87994".

Figure 5 shows a nucleotide sequence (SEQ ID NO:5) of a TAT158 cDNA, wherein SEQ ID NO:5 is a clone designated herein as "DNA88131".

Figure 6 shows a nucleotide sequence (SEQ ID NO:6) of a TAT110 cDNA, wherein SEQ ID NO:6 is a clone designated herein as "DNA95930".

Figure 7 shows a nucleotide sequence (SEQ ID NO:7) of a TAT210 cDNA, wherein SEQ ID NO:7 is a clone designated herein as "DNA95930-1".

Figure 8 shows a nucleotide sequence (SEQ ID NO:8) of a TAT159 cDNA, wherein SEQ ID NO:8 is a clone designated herein as "DNA96917".

Figure 9 shows a nucleotide sequence (SEQ ID NO:9) of a TAT112 cDNA, wherein SEQ ID NO:9 is a clone designated herein as "DNA96930".

Figure 10 shows a nucleotide sequence (SEQ ID NO:10) of a TAT147 cDNA, wherein SEQ ID NO:10 is a clone designated herein as "DNA96936".

Figure 11 shows a nucleotide sequence (SEQ ID NO:11) of a TAT145 cDNA, wherein SEQ ID NO:11 is a clone designated herein as "DNA98565".

Figure 12 shows a nucleotide sequence (SEQ ID NO:12) of a TAT152 cDNA, wherein SEQ ID NO:12 is a clone designated herein as "DNA246435".

Figure 13 shows a nucleotide sequence (SEQ ID NO:13) of a TAT162 cDNA, wherein SEQ ID NO:13 is a clone designated herein as "DNA98591".

Figure 14 shows a nucleotide sequence (SEQ ID NO:14) of a TAT114 cDNA, wherein SEQ ID NO:14 is a clone designated herein as "DNA108809".

Figure 15 shows a nucleotide sequence (SEQ ID NO:15) of a TAT119 cDNA, wherein SEQ ID NO:15 is a clone designated herein as "DNA119488".

Figure 16 shows a nucleotide sequence (SEQ ID NO:16) of a TAT103 cDNA, wherein SEQ ID NO:16 is a clone designated herein as "DNA143493".

Figures 17A-B show a nucleotide sequence (SEQ ID NO:17) of a TAT130 cDNA, wherein SEQ ID NO:17 is a clone designated herein as "DNA167234".

Figure 18 shows a nucleotide sequence (SEQ ID NO:18) of a TAT166 cDNA, wherein SEQ ID NO:18 is a clone designated herein as "DNA235621".

Figure 19 shows a nucleotide sequence (SEQ ID NO:19) of a TAT132 cDNA, wherein SEQ ID NO:19 is a clone designated herein as "DNA176766".

Figure 20 shows a nucleotide sequence (SEQ ID NO:20) of a TAT150 cDNA, wherein SEQ ID NO:20 is a clone designated herein as "DNA236463".

Figure 21 shows a nucleotide sequence (SEQ ID NO:21) of a TAT129 cDNA, wherein SEQ ID NO:21 is a clone designated herein as "DNA181162".

Figure 22 shows a nucleotide sequence (SEQ ID NO:22) of a TAT111 cDNA, wherein SEQ ID NO:22 is a clone designated herein as "DNA188221".

Figure 23 shows a nucleotide sequence (SEQ ID NO:23) of a TAT146 cDNA, wherein SEQ ID NO:23 is a clone designated herein as "DNA233876".

Figure 24 shows a nucleotide sequence (SEQ ID NO:24) of a TAT148 cDNA, wherein SEQ ID NO:24 is a clone designated herein as "DNA193891".

Figure 25 shows a nucleotide sequence (SEQ ID NO:25) of a TAT187 cDNA, wherein SEQ ID NO:25 is a clone designated herein as "DNA248170".

Figure 26 shows a nucleotide sequence (SEQ ID NO:26) of a TAT118 cDNA, wherein SEQ ID NO:26 is a clone designated herein as "DNA194628".

Figure 27 shows a nucleotide sequence (SEQ ID NO:27) of a TAT167 cDNA, wherein SEQ ID NO:27 is a clone designated herein as "DNA246415".

Figure 28 shows a nucleotide sequence (SEQ ID NO:28) of a TAT123 cDNA, wherein SEQ ID

NO:28 is a clone designated herein as "DNA210499".

Figure 29 shows a nucleotide sequence (SEQ ID NO:29) of a TAT211 cDNA, wherein SEQ ID NO:29 is a clone designated herein as "DNA219894".

Figure 30 shows a nucleotide sequence (SEQ ID NO:30) of a TAT113 cDNA, wherein SEQ ID NO:30 is a clone designated herein as "DNA215609".

Figure 31 shows a nucleotide sequence (SEQ ID NO:31) of a TAT128 cDNA, wherein SEQ ID NO:31 is a clone designated herein as "DNA220432".

Figures 32A-B show a nucleotide sequence (SEQ ID NO:32) of a TAT164 cDNA, wherein SEQ ID NO:32 is a clone designated herein as "DNA226094".

Figure 33 shows a nucleotide sequence (SEQ ID NO:33) of a TAT122 cDNA, wherein SEQ ID NO:33 is a clone designated herein as "DNA226165".

Figure 34 shows a nucleotide sequence (SEQ ID NO:34) of a TAT117 cDNA, wherein SEQ ID NO:34 is a clone designated herein as "DNA226237".

Figure 35 shows a nucleotide sequence (SEQ ID NO:35) of a TAT168 cDNA, wherein SEQ ID NO:35 is a clone designated herein as "DNA246450".

Figure 36 shows a nucleotide sequence (SEQ ID NO:36) of a TAT144 cDNA, wherein SEQ ID NO:36 is a clone designated herein as "DNA226456".

Figure 37 shows a nucleotide sequence (SEQ ID NO:37) of a TAT188 cDNA, wherein SEQ ID NO:37 is a clone designated herein as "DNA237637".

Figure 38 shows a nucleotide sequence (SEQ ID NO:38) of a TAT126 cDNA, wherein SEQ ID NO:38 is a clone designated herein as "DNA226539".

Figure 39 shows a nucleotide sequence (SEQ ID NO:39) of a TAT151 cDNA, wherein SEQ ID NO:39 is a clone designated herein as "DNA236511".

Figure 40 shows a nucleotide sequence (SEQ ID NO:40) of a TAT115 cDNA, wherein SEQ ID NO:40 is a clone designated herein as "DNA226771".

Figure 41 shows a nucleotide sequence (SEQ ID NO:41) of a TAT163 cDNA, wherein SEQ ID NO:41 is a clone designated herein as "DNA227087".

Figure 42 shows a nucleotide sequence (SEQ ID NO:42) of a TAT227 cDNA, wherein SEQ ID NO:42 is a clone designated herein as "DNA266307".

Figure 43 shows a nucleotide sequence (SEQ ID NO:43) of a TAT228 cDNA, wherein SEQ ID NO:43 is a clone designated herein as "DNA266311".

Figure 44 shows a nucleotide sequence (SEQ ID NO:44) of a TAT229 cDNA, wherein SEQ ID NO:44 is a clone designated herein as "DNA266312".

Figure 45 shows a nucleotide sequence (SEQ ID NO:45) of a TAT230 cDNA, wherein SEQ ID NO:45 is a clone designated herein as "DNA266313".

Figure 46 shows a nucleotide sequence (SEQ ID NO:46) of a TAT121 cDNA, wherein SEQ ID NO:46 is a clone designated herein as "DNA227224".

Figure 47 shows a nucleotide sequence (SEQ ID NO:47) of a TAT183 cDNA, wherein SEQ ID NO:47 is a clone designated herein as "DNA247486".

Figure 48 shows a nucleotide sequence (SEQ ID NO:48) of a TAT165 cDNA, wherein SEQ ID NO:48 is a clone designated herein as "DNA227578".

Figure 49 shows a nucleotide sequence (SEQ ID NO:49) of a TAT131 cDNA, wherein SEQ ID NO:49 is a clone designated herein as "DNA227800".

Figure 50 shows a nucleotide sequence (SEQ ID NO:50) of a TAT140 cDNA, wherein SEQ ID NO:50 is a clone designated herein as "DNA227904".

Figure 51 shows a nucleotide sequence (SEQ ID NO:51) of a TAT127 cDNA, wherein SEQ ID NO:51 is a clone designated herein as "DNA228199".

Figure 52 shows a nucleotide sequence (SEQ ID NO:52) of a TAT116 cDNA, wherein SEQ ID NO:52 is a clone designated herein as "DNA228201".

Figure 53 shows a nucleotide sequence (SEQ ID NO:53) of a TAT189 cDNA, wherein SEQ ID NO:53 is a clone designated herein as "DNA247488".

Figure 54 shows a nucleotide sequence (SEQ ID NO:54) of a TAT190 cDNA, wherein SEQ ID NO:54 is a clone designated herein as "DNA236538".

Figure 55 shows a nucleotide sequence (SEQ ID NO:55) of a TAT191 cDNA, wherein SEQ ID NO:55 is a clone designated herein as "DNA247489".

Figure 56 shows a nucleotide sequence (SEQ ID NO:56) of a TAT133 cDNA, wherein SEQ ID NO:56 is a clone designated herein as "DNA228211".

Figure 57 shows a nucleotide sequence (SEQ ID NO:57) of a TAT186 cDNA, wherein SEQ ID NO:57 is a clone designated herein as "DNA233937".

Figure 58 shows a nucleotide sequence (SEQ ID NO:58) of a TAT120 cDNA, wherein SEQ ID NO:58 is a clone designated herein as "DNA228993".

Figure 59 shows a nucleotide sequence (SEQ ID NO:59) of a TAT124 cDNA, wherein SEQ ID NO:59 is a clone designated herein as "DNA228994".

Figure 60 shows a nucleotide sequence (SEQ ID NO:60) of a TAT105 cDNA, wherein SEQ ID NO:60 is a clone designated herein as "DNA229410".

Figures 61A-B show a nucleotide sequence (SEQ ID NO:61) of a TAT107 cDNA, wherein SEQ ID NO:61 is a clone designated herein as "DNA229411".

Figures 62A-B show a nucleotide sequence (SEQ ID NO:62) of a TAT108 cDNA, wherein SEQ ID NO:62 is a clone designated herein as "DNA229413".

Figures 63A-B show a nucleotide sequence (SEQ ID NO:63) of a TAT139 cDNA, wherein SEQ ID NO:63 is a clone designated herein as "DNA229700".

Figure 64 shows a nucleotide sequence (SEQ ID NO:64) of a TAT143 cDNA, wherein SEQ ID NO:64 is a clone designated herein as "DNA231312".

Figure 65 shows a nucleotide sequence (SEQ ID NO:65) of a TAT100 cDNA, wherein SEQ ID

NO:65 is a clone designated herein as "DNA231542".

Figure 66 shows a nucleotide sequence (SEQ ID NO:66) of a TAT284 cDNA, wherein SEQ ID NO:66 is a clone designated herein as "DNA231542-1".

Figure 67 shows a nucleotide sequence (SEQ ID NO:67) of a TAT285 cDNA, wherein SEQ ID NO:67 is a clone designated herein as "DNA231542-2".

Figure 68 shows a nucleotide sequence (SEQ ID NO:68) of a TAT285-1 cDNA, wherein SEQ ID NO:68 is a clone designated herein as "DNA297393".

Figure 69 shows a nucleotide sequence (SEQ ID NO:69) of a TAT125 cDNA, wherein SEQ ID NO:69 is a clone designated herein as "DNA232754".

Figure 70 shows a nucleotide sequence (SEQ ID NO:70) of a TAT149 cDNA, wherein SEQ ID NO:70 is a clone designated herein as "DNA234833".

Figure 71 shows a nucleotide sequence (SEQ ID NO:71) of a TAT231 cDNA, wherein SEQ ID NO:71 is a clone designated herein as "DNA268022".

Figure 72 shows a nucleotide sequence (SEQ ID NO:72) of a TAT153 cDNA, wherein SEQ ID NO:72 is a clone designated herein as "DNA236246".

Figure 73 shows a nucleotide sequence (SEQ ID NO:73) of a TAT104 cDNA, wherein SEQ ID NO:73 is a clone designated herein as "DNA236343".

Figure 74 shows a nucleotide sequence (SEQ ID NO:74) of a TAT141 cDNA, wherein SEQ ID NO:74 is a clone designated herein as "DNA236493".

Figure 75 shows a nucleotide sequence (SEQ ID NO:75) of a TAT102 cDNA, wherein SEQ ID NO:75 is a clone designated herein as "DNA236534".

Figure 76 shows a nucleotide sequence (SEQ ID NO:76) of a TAT109 cDNA, wherein SEQ ID NO:76 is a clone designated herein as "DNA246430".

Figure 77 shows a nucleotide sequence (SEQ ID NO:77) of a TAT142 cDNA, wherein SEQ ID NO:77 is a clone designated herein as "DNA247480".

Figures 78A-B show a nucleotide sequence (SEQ ID NO:78) of a TAT106 cDNA, wherein SEQ ID NO:78 is a clone designated herein as "DNA264454".

Figure 79 shows the amino acid sequence (SEQ ID NO:79) derived from the coding sequence of SEQ ID NO:1 shown in Figure 1.

Figure 80 shows the amino acid sequence (SEQ ID NO:80) derived from the coding sequence of SEQ ID NO:2 shown in Figure 2.

Figure 81 shows the amino acid sequence (SEQ ID NO:81) derived from the coding sequence of SEQ ID NO:3 shown in Figure 3.

Figure 82 shows the amino acid sequence (SEQ ID NO:82) derived from the coding sequence of SEQ ID NO:4 shown in Figure 4.

Figure 83 shows the amino acid sequence (SEQ ID NO:83) derived from the coding sequence of SEQ ID NO:5 shown in Figure 5.

Figure 84 shows the amino acid sequence (SEQ ID NO:84) derived from the coding sequence of SEQ ID NO:6 shown in Figure 6.

Figure 85 shows the amino acid sequence (SEQ ID NO:85) derived from the coding sequence of SEQ ID NO:7 shown in Figure 7.

Figure 86 shows the amino acid sequence (SEQ ID NO:86) derived from the coding sequence of SEQ ID NO:8 shown in Figure 8.

Figure 87 shows the amino acid sequence (SEQ ID NO:87) derived from the coding sequence of SEQ ID NO:9 shown in Figure 9.

Figure 88 shows the amino acid sequence (SEQ ID NO:88) derived from the coding sequence of SEQ ID NO:10 shown in Figure 10.

Figure 89 shows the amino acid sequence (SEQ ID NO:89) derived from the coding sequence of SEQ ID NO:11 shown in Figure 11.

Figure 90 shows the amino acid sequence (SEQ ID NO:90) derived from the coding sequence of SEQ ID NO:12 shown in Figure 12.

Figure 91 shows the amino acid sequence (SEQ ID NO:91) derived from the coding sequence of SEQ ID NO:13 shown in Figure 13.

Figure 92 shows the amino acid sequence (SEQ ID NO:92) derived from the coding sequence of SEQ ID NO:14 shown in Figure 14.

Figure 93 shows the amino acid sequence (SEQ ID NO:93) derived from the coding sequence of SEQ ID NO:15 shown in Figure 15.

Figure 94 shows the amino acid sequence (SEQ ID NO:94) derived from the coding sequence of SEQ ID NO:16 shown in Figure 16.

Figure 95 shows the amino acid sequence (SEQ ID NO:95) derived from the coding sequence of SEQ ID NO:17 shown in Figures 17A-B.

Figure 96 shows the amino acid sequence (SEQ ID NO:96) derived from the coding sequence of SEQ ID NO:18 shown in Figure 18.

Figure 97 shows the amino acid sequence (SEQ ID NO:97) derived from the coding sequence of SEQ ID NO:19 shown in Figure 19.

Figure 98 shows the amino acid sequence (SEQ ID NO:98) derived from the coding sequence of SEQ ID NO:20 shown in Figure 20.

Figure 99 shows the amino acid sequence (SEQ ID NO:99) derived from the coding sequence of SEQ ID NO:21 shown in Figure 21.

Figure 100 shows the amino acid sequence (SEQ ID NO:100) derived from the coding sequence of SEQ ID NO:22 shown in Figure 22.

Figure 101 shows the amino acid sequence (SEQ ID NO:101) derived from the coding sequence of SEQ ID NO:23 shown in Figure 23.

Figure 102 shows the amino acid sequence (SEQ ID NO:102) derived from the coding sequence of

SEQ ID NO:24 shown in Figure 24.

Figure 103 shows the amino acid sequence (SEQ ID NO:103) derived from the coding sequence of SEQ ID NO:25 shown in Figure 25.

Figure 104 shows the amino acid sequence (SEQ ID NO:104) derived from the coding sequence of SEQ ID NO:26 shown in Figure 26.

Figure 105 shows the amino acid sequence (SEQ ID NO:105) derived from the coding sequence of SEQ ID NO:27 shown in Figure 27.

Figure 106 shows the amino acid sequence (SEQ ID NO:106) derived from the coding sequence of SEQ ID NO:28 shown in Figure 28.

Figure 107 shows the amino acid sequence (SEQ ID NO:107) derived from the coding sequence of SEQ ID NO:29 shown in Figure 29.

Figure 108 shows the amino acid sequence (SEQ ID NO:108) derived from the coding sequence of SEQ ID NO:30 shown in Figure 30.

Figure 109 shows the amino acid sequence (SEQ ID NO:109) derived from the coding sequence of SEQ ID NO:31 shown in Figure 31.

Figures 110A-B shows the amino acid sequence (SEQ ID NO:110) derived from the coding sequence of SEQ ID NO:32 shown in Figures 32A-B.

Figure 111 shows the amino acid sequence (SEQ ID NO:111) derived from the coding sequence of SEQ ID NO:33 shown in Figure 33.

Figure 112 shows the amino acid sequence (SEQ ID NO:112) derived from the coding sequence of SEQ ID NO:34 shown in Figure 34.

Figure 113 shows the amino acid sequence (SEQ ID NO:113) derived from the coding sequence of SEQ ID NO:35 shown in Figure 35.

Figure 114 shows the amino acid sequence (SEQ ID NO:114) derived from the coding sequence of SEQ ID NO:36 shown in Figure 36.

Figure 115 shows the amino acid sequence (SEQ ID NO:115) derived from the coding sequence of SEQ ID NO:37 shown in Figure 37.

Figure 116 shows the amino acid sequence (SEQ ID NO:116) derived from the coding sequence of SEQ ID NO:38 shown in Figure 38.

Figure 117 shows the amino acid sequence (SEQ ID NO:117) derived from the coding sequence of SEQ ID NO:39 shown in Figure 39.

Figure 118 shows the amino acid sequence (SEQ ID NO:118) derived from the coding sequence of SEQ ID NO:40 shown in Figure 40.

Figure 119 shows the amino acid sequence (SEQ ID NO:119) derived from the coding sequence of SEQ ID NO:41 shown in Figure 41.

Figure 120 shows the amino acid sequence (SEQ ID NO:120) derived from the coding sequence of SEQ ID NO:42 shown in Figure 42.

Figure 121 shows the amino acid sequence (SEQ ID NO:121) derived from the coding sequence of SEQ ID NO:43 shown in Figure 43.

Figure 122 shows the amino acid sequence (SEQ ID NO:122) derived from the coding sequence of SEQ ID NO:44 shown in Figure 44.

Figure 123 shows the amino acid sequence (SEQ ID NO:123) derived from the coding sequence of SEQ ID NO:45 shown in Figure 45.

Figure 124 shows the amino acid sequence (SEQ ID NO:124) derived from the coding sequence of SEQ ID NO:46 shown in Figure 46.

Figure 125 shows the amino acid sequence (SEQ ID NO:125) derived from the coding sequence of SEQ ID NO:47 shown in Figure 47.

Figure 126 shows the amino acid sequence (SEQ ID NO:126) derived from the coding sequence of SEQ ID NO:48 shown in Figure 48.

Figure 127 shows the amino acid sequence (SEQ ID NO:127) derived from the coding sequence of SEQ ID NO:49 shown in Figure 49.

Figure 128 shows the amino acid sequence (SEQ ID NO:128) derived from the coding sequence of SEQ ID NO:50 shown in Figure 50.

Figure 129 shows the amino acid sequence (SEQ ID NO:129) derived from the coding sequence of SEQ ID NO:51 shown in Figure 51.

Figure 130 shows the amino acid sequence (SEQ ID NO:130) derived from the coding sequence of SEQ ID NO:52 shown in Figure 52.

Figure 131 shows the amino acid sequence (SEQ ID NO:131) derived from the coding sequence of SEQ ID NO:53 shown in Figure 53.

Figure 132 shows the amino acid sequence (SEQ ID NO:132) derived from the coding sequence of SEQ ID NO:54 shown in Figure 54.

Figure 133 shows the amino acid sequence (SEQ ID NO:133) derived from the coding sequence of SEQ ID NO:55 shown in Figure 55.

Figure 134 shows the amino acid sequence (SEQ ID NO:134) derived from the coding sequence of SEQ ID NO:56 shown in Figure 56.

Figure 135 shows the amino acid sequence (SEQ ID NO:135) derived from the coding sequence of SEQ ID NO:57 shown in Figure 57.

Figure 136 shows the amino acid sequence (SEQ ID NO:136) derived from the coding sequence of SEQ ID NO:58 shown in Figure 58.

Figure 137 shows the amino acid sequence (SEQ ID NO:137) derived from the coding sequence of SEQ ID NO:59 shown in Figure 59.

Figure 138 shows the amino acid sequence (SEQ ID NO:138) derived from the coding sequence of SEQ ID NO:60 shown in Figure 60.

Figure 139 shows the amino acid sequence (SEQ ID NO:139) derived from the coding sequence of

SEQ ID NO:61 shown in Figures 61A-B.

Figure 140 shows the amino acid sequence (SEQ ID NO:140) derived from the coding sequence of SEQ ID NO:62 shown in Figures 62A-B.

Figure 141 shows the amino acid sequence (SEQ ID NO:141) derived from the coding sequence of SEQ ID NO:63 shown in Figures 63A-B.

Figure 142 shows the amino acid sequence (SEQ ID NO:142) derived from the coding sequence of SEQ ID NO:64 shown in Figure 64.

Figure 143 shows the amino acid sequence (SEQ ID NO:143) derived from the coding sequence of SEQ ID NO:66 shown in Figure 66.

Figure 144 shows the amino acid sequence (SEQ ID NO:144) derived from the coding sequence of SEQ ID NO:67 shown in Figure 67.

Figure 145 shows the amino acid sequence (SEQ ID NO:145) derived from the coding sequence of SEQ ID NO:68 shown in Figure 68.

Figure 146 shows the amino acid sequence (SEQ ID NO:146) derived from the coding sequence of SEQ ID NO:69 shown in Figure 69.

Figure 147 shows the amino acid sequence (SEQ ID NO:147) derived from the coding sequence of SEQ ID NO:70 shown in Figure 70.

Figure 148 shows the amino acid sequence (SEQ ID NO:148) derived from the coding sequence of SEQ ID NO:71 shown in Figure 71.

Figure 149 shows the amino acid sequence (SEQ ID NO:149) derived from the coding sequence of SEQ ID NO:73 shown in Figure 73.

Figure 150 shows the amino acid sequence (SEQ ID NO:150) derived from the coding sequence of SEQ ID NO:74 shown in Figure 74.

Figure 151 shows the amino acid sequence (SEQ ID NO:151) derived from the coding sequence of SEQ ID NO:75 shown in Figure 75.

Figure 152 shows the amino acid sequence (SEQ ID NO:152) derived from the coding sequence of SEQ ID NO:76 shown in Figure 76.

Figure 153 shows the amino acid sequence (SEQ ID NO:153) derived from the coding sequence of SEQ ID NO:77 shown in Figure 77.

Figure 154 shows the amino acid sequence (SEQ ID NO:154) derived from the coding sequence of SEQ ID NO:78 shown in Figures 78A-B.

Figure 155 is a diagram depicting the three dimensional structure of the E16 polypeptide (TAT188) as a subunit of the sodium ion independent large neutral amino acid transporter.

Figure 156 shows the graphical results of FACS plots demonstrating binding of anti-TAT188 (anti-E16) antibodies to PC3 cell surface (green plot) and reduction in binding with reduction in TAT188 expression (red plot).

Figure 157 shows internalization of anti-TAT188 antibody after binding to the surface of TAT188-

expressing PC3 cells.

Figure 158 is a bar graph showing changes in TAT188 amino acid transport activity in the presence of anti-TAT188 antibody with time.

Figure 159 provides plots of cell viability in the presence of increasing amounts of MC-vc-PAB-MMAE toxin-conjugated anti-TAT188 antibody in PC3 and Colo205 cells.

Figure 160 is a plot of cell viability in the presence of increasing amounts of MC-vc-PAB-MMAE or MC-vc-PAB-MMAF toxin-conjugated anti-TAT188 antibody in Colo205 cells.

Figure 161shows the results anti-human E16 antibody binding cells endogenously expressing E16 polypeptide, where the cells are monkey COS7 cells, human MCF10A breast cancer cells, and mouse NIH-3T3 cells. Figure 161A shows the FACS results of binding of anti-E16 antibodies 3B5, 12B9, and 12G12 to monkey COS. Figure 161B shows binding of anti-E16 3B5 to human breast tumor cell line MCF10A and no binding to mouse NIH-3T3 cells.

Figure 162 is a plot of mean tumor volume changes over time in xenograft mice administered naked (not toxin-conjugated) anti-E16 antibodies.

Figure 163 is a graph of mean tumor volume changes over time in xenograft mice administerd toxin-conjugated anti-E16 antibodies, where the toxin was either MC-vc-PAB-MMAE or MC-vc-PAB-MMAF.

Figure 164 shows that E16-GFP fusion polypeptide expression is reduced in cells transfected with siRNA that targets the E16 gene. The top panel are photographs of the results of bioluminescence assays for the presence of GFP. The lower panel are photographs of phase contrast visualization showing the GFP bioluminescence reduction was not caused by reduction in cell number.

Figure 165 is a Western blot showing that E16-GFP fusion polypeptide expression is inhibited in PC3 cells transiently transfected with E16 siRNA. β-tubulin is a gel loading control.

Figure 166 shows that transient transfection of PC3 cells with E16 siRNA is associated with a reduction in amino acid transport function, consistent with reduced expression of E16.

Figure 167 is a plot of cell proliferation over time with and without transient transfection with E16 siRNA. Cell proliferation is reduced in PC3 cells transiently transfected with E16.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

I. Definitions

The terms "TAT polypeptide" and "TAT" as used herein and when immediately followed by a numerical designation, refer to various polypeptides, wherein the complete designation (i.e.,TAT/number) refers to specific polypeptide sequences as described herein. The terms "TAT/number polypeptide" and "TAT/number" wherein the term "number" is provided as an actual numerical designation as used herein encompass native sequence polypeptides, polypeptide variants and fragments of native sequence polypeptides and polypeptide variants (which are further defined herein). The TAT polypeptides described herein may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods. The term "TAT polypeptide" refers to each individual TAT/number

polypeptide disclosed herein. All disclosures in this specification which refer to the "TAT polypeptide" refer to each of the polypeptides individually as well as jointly. For example, descriptions of the preparation of, purification of, derivation of, formation of antibodies to or against, formation of TAT binding oligopeptides to or against, formation of TAT binding organic molecules to or against, administration of, compositions containing, treatment of a disease with, etc., pertain to each polypeptide of the invention individually. The term "TAT polypeptide" also includes variants of the TAT/number polypeptides disclosed herein.

A "native sequence TAT polypeptide" comprises a polypeptide having the same amino acid sequence as the corresponding TAT polypeptide derived from nature. Such native sequence TAT polypeptides can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence TAT polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of the specific TAT polypeptide (*e.g.*, an extracellular domain sequence), naturally-occurring variant forms (*e.g.*, alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In certain embodiments of the invention, the native sequence TAT polypeptides disclosed herein are mature or full-length native sequence polypeptides comprising the full-length amino acids sequences shown in the accompanying figures. Start and stop codons (if indicated) are shown in bold font and underlined in the figures. Nucleic acid residues indicated as "N" in the accompanying figures are any nucleic acid residue. However, while the TAT polypeptides disclosed in the accompanying figures are shown to begin with methionine residues designated herein as amino acid position 1 in the figures, it is conceivable and possible that other methionine residues located either upstream or downstream from the amino acid position 1 in the figures may be employed as the starting amino acid residue for the TAT polypeptides.

The TAT polypeptide "extracellular domain" or "ECD" refers to a form of the TAT polypeptide which is essentially free of the transmembrane and cytoplasmic domains. Ordinarily, a TAT polypeptide ECD will have less than 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than 0.5% of such domains. It will be understood that any transmembrane domains identified for the TAT polypeptides of the present invention are identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified herein. Optionally, therefore, an extracellular domain of a TAT polypeptide may contain from about 5 or fewer amino acids on either side of the transmembrane domain/extracellular domain boundary as identified in the Examples or specification and such polypeptides, with or without the associated signal peptide, and nucleic acid encoding them, are contemplated by the present invention.

The approximate location of the "signal peptides" of the various TAT polypeptides disclosed herein may be shown in the present specification and/or the accompanying figures. It is noted, however, that the C-terminal boundary of a signal peptide may vary, but most likely by no more than about 5 amino acids on either side of the signal peptide C-terminal boundary as initially identified herein, wherein the C-terminal boundary of the signal peptide may be identified pursuant to criteria routinely employed in the art for identifying that type of amino acid sequence element (e.g., Nielsen et al., Prot. Eng. 10:1-6 (1997) and von Heinje et al.,

Nucl. Acids. Res. 14:4683-4690 (1986)). Moreover, it is also recognized that, in some cases, cleavage of a signal sequence from a secreted polypeptide is not entirely uniform, resulting in more than one secreted species. These mature polypeptides, where the signal peptide is cleaved within no more than about 5 amino acids on either side of the C-terminal boundary of the signal peptide as identified herein, and the polynucleotides encoding them, are contemplated by the present invention.

"TAT polypeptide variant" means a TAT polypeptide, preferably an active TAT polypeptide, as defined herein having at least about 80% amino acid sequence identity with a full-length native sequence TAT polypeptide sequence as disclosed herein, a TAT polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a TAT polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length TAT polypeptide sequence as disclosed herein (such as those encoded by a nucleic acid that represents only a portion of the complete coding sequence for a full-length TAT polypeptide). Such TAT polypeptide variants include, for instance, TAT polypeptides wherein one or more amino acid residues are added, or deleted, at the N– or C-terminus of the full-length native amino acid sequence. Ordinarily, a TAT polypeptide variant will have at least about 80% amino acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% amino acid sequence identity, to a full-length native sequence TAT polypeptide sequence as disclosed herein, a TAT polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a TAT polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length TAT polypeptide sequence as disclosed herein. Ordinarily, TAT variant polypeptides are at least about 10 amino acids in length, alternatively at least about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600 amino acids in length, or more. Optionally, TAT variant polypeptides will have no more than one conservative amino acid substitution as compared to the native TAT polypeptide sequence, alternatively no more than 2, 3, 4, 5, 6, 7, 8, 9, or 10 conservative amino acid substitution as compared to the native TAT polypeptide sequence.

"Percent (%) amino acid sequence identity" with respect to the TAT polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific TAT polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is

provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

In situations where ALIGN-2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

100 times the fraction X/Y

where X is the number of amino acid residues scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A. As examples of % amino acid sequence identity calculations using this method, Tables 2 and 3 demonstrate how to calculate the % amino acid sequence identity of the amino acid sequence designated "Comparison Protein" to the amino acid sequence designated "TAT", wherein "TAT" represents the amino acid sequence of a hypothetical TAT polypeptide of interest, "Comparison Protein" represents the amino acid sequence of a polypeptide against which the "TAT" polypeptide of interest is being compared, and "X, "Y" and "Z" each represent different hypothetical amino acid residues. Unless specifically stated otherwise, all % amino acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program.

"TAT variant polynucleotide" or "TAT variant nucleic acid sequence" means a nucleic acid molecule which encodes a TAT polypeptide, preferably an active TAT polypeptide, as defined herein and which has at least about 80% nucleic acid sequence identity with a nucleotide acid sequence encoding a full-length native sequence TAT polypeptide sequence as disclosed herein, a full-length native sequence TAT polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a TAT polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length TAT polypeptide sequence as disclosed herein (such as those encoded by a nucleic acid that represents only a portion of the complete coding sequence for a full-length TAT polypeptide). Ordinarily, a TAT variant polynucleotide will have at least about 80% nucleic acid sequence identity, alternatively at least about 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% nucleic acid sequence identity with a nucleic acid sequence encoding a full-length native sequence TAT

polypeptide sequence as disclosed herein, a full-length native sequence TAT polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a TAT polypeptide, with or without the signal sequence, as disclosed herein or any other fragment of a full-length TAT polypeptide sequence as disclosed herein. Variants do not encompass the native nucleotide sequence.

Ordinarily, TAT variant polynucleotides are at least about 5 nucleotides in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, or 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length.

"Percent (%) nucleic acid sequence identity" with respect to TAT-encoding nucleic acid sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the TAT nucleic acid sequence of interest, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. For purposes herein, however, % nucleic acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

In situations where ALIGN-2 is employed for nucleic acid sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

100 times the fraction W/Z

where W is the number of nucleotides scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will

be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C. As examples of % nucleic acid sequence identity calculations, Tables 4 and 5, demonstrate how to calculate the % nucleic acid sequence identity of the nucleic acid sequence designated "Comparison DNA" to the nucleic acid sequence designated "TAT-DNA", wherein "TAT-DNA" represents a hypothetical TAT-encoding nucleic acid sequence of interest, "Comparison DNA" represents the nucleotide sequence of a nucleic acid molecule against which the "TAT-DNA" nucleic acid molecule of interest is being compared, and "N", "L" and "V" each represent different hypothetical nucleotides. Unless specifically stated otherwise, all % nucleic acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program.

In other embodiments, TAT variant polynucleotides are nucleic acid molecules that encode a TAT polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding a full-length TAT polypeptide as disclosed herein. TAT variant polypeptides may be those that are encoded by a TAT variant polynucleotide.

The term "full-length coding region" when used in reference to a nucleic acid encoding a TAT polypeptide refers to the sequence of nucleotides which encode the full-length TAT polypeptide of the invention (which is often shown between start and stop codons, inclusive thereof, in the accompanying figures). The term "full-length coding region" when used in reference to an ATCC deposited nucleic acid refers to the TAT polypeptide-encoding portion of the cDNA that is inserted into the vector deposited with the ATCC (which is often shown between start and stop codons, inclusive thereof, in the accompanying figures).

"Isolated," when used to describe the various TAT polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the TAT polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

An "isolated" TAT polypeptide-encoding nucleic acid or other polypeptide-encoding nucleic acid is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the polypeptide-encoding nucleic acid. An isolated polypeptide-encoding nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated polypeptide-encoding nucleic acid molecules therefore are distinguished from the specific polypeptide-encoding nucleic acid molecule as it exists in natural cells. However, an isolated polypeptide-

encoding nucleic acid molecule includes polypeptide-encoding nucleic acid molecules contained in cells that ordinarily express the polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

"Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

"Stringent conditions" or "high stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) overnight hybridization in a solution that employs 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with a 10 minute wash at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) followed by a 10 minute high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

"Moderately stringent conditions" may be identified as described by Sambrook et al., Molecular

Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent that those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising a TAT polypeptide or anti-TAT antibody fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

"Active" or "activity" for the purposes herein refers to form(s) of a TAT polypeptide which retain a biological and/or an immunological activity of native or naturally-occurring TAT, wherein "biological" activity refers to a biological function (either inhibitory or stimulatory) caused by a native or naturally-occurring TAT other than the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring TAT and an "immmological" activity refers to the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring TAT.

The term "antagonist" is used in the broadest sense, and includes any molecule that partially or fully blocks, inhibits, or neutralizes a biological activity of a native TAT polypeptide disclosed herein. In a similar manner, the term "agonist" is used in the broadest sense and includes any molecule that mimics a biological activity of a native TAT polypeptide disclosed herein. Suitable agonist or antagonist molecules specifically include agonist or antagonist antibodies or antibody fragments, fragments or amino acid sequence variants of native TAT polypeptides, peptides, antisense oligonucleotides, small organic molecules, etc. Methods for identifying agonists or antagonists of a TAT polypeptide may comprise contacting a TAT polypeptide with a candidate agonist or antagonist molecule and measuring a detectable change in one or more biological activities normally associated with the TAT polypeptide.

"Treating" or "treatment" or "alleviation" refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathologic condition or disorder. Those in need of treatment include those already with the disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented. A subject or mammal is successfully "treated" for a TAT polypeptide-expressing cancer if, after receiving a therapeutic amount of an anti-TAT antibody, TAT binding oligopeptide or TAT binding organic molecule according to the methods of the present invention, the patient shows observable and/or measurable reduction in or absence of one or more

of the following: reduction in the number of cancer cells or absence of the cancer cells; reduction in the tumor size; inhibition (i.e., slow to some extent and preferably stop) of cancer cell infiltration into peripheral organs including the spread of cancer into soft tissue and bone; inhibition (i.e., slow to some extent and preferably stop) of tumor metastasis; inhibition, to some extent, of tumor growth; and/or relief to some extent, one or more of the symptoms associated with the specific cancer; reduced morbidity and mortality, and improvement in quality of life issues. To the extent the anti-TAT antibody or TAT binding oligopeptide may prevent growth and/or kill existing cancer cells, it may be cytostatic and/or cytotoxic. Reduction of these signs or symptoms may also be felt by the patient.

The above parameters for assessing successful treatment and improvement in the disease are readily measurable by routine procedures familiar to a physician. For cancer therapy, efficacy can be measured, for example, by assessing the time to disease progression (TTP) and/or determining the response rate (RR). Metastasis can be determined by staging tests and by bone scan and tests for calcium level and other enzymes to determine spread to the bone. CT scans can also be done to look for spread to the pelvis and lymph nodes in the area. Chest X-rays and measurement of liver enzyme levels by known methods are used to look for metastasis to the lungs and liver, respectively. Other routine methods for monitoring the disease include transrectal ultrasonography (TRUS) and transrectal needle biopsy (TRNB).

For bladder cancer, which is a more localized cancer, methods to determine progress of disease include urinary cytologic evaluation by cystoscopy, monitoring for presence of blood in the urine, visualization of the urothelial tract by sonography or an intravenous pyelogram, computed tomography (CT) and magnetic resonance imaging (MRI). The presence of distant metastases can be assessed by CT of the abdomen, chest x-rays, or radionuclide imaging of the skeleton.

"Chronic" administration refers to administration of the agent(s) in a continuous mode as opposed to an acute mode, so as to maintain the initial therapeutic effect (activity) for an extended period of time. "Intermittent" administration is treatment that is not consecutively done without interruption, but rather is cyclic in nature.

"Mammal" for purposes of the treatment of, alleviating the symptoms of or diagnosis of a cancer refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, etc. Preferably, the mammal is human.

Administration "in combination with" one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

"Carriers" as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are nontoxic to the cell or mammal being exposed thereto at the dosages and concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such

as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN®, polyethylene glycol (PEG), and PLURONICS®.

By "solid phase" or "solid support" is meant a non-aqueous matrix to which an antibody, TAT binding oligopeptide or TAT binding organic molecule of the present invention can adhere or attach. Examples of solid phases encompassed herein include those formed partially or entirely of glass (e.g., controlled pore glass), polysaccharides (e.g., agarose), polyacrylamides, polystyrene, polyvinyl alcohol and silicones. In certain embodiments, depending on the context, the solid phase can comprise the well of an assay plate; in others it is a purification column (e.g., an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Patent No. 4,275,149.

A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug (such as a TAT polypeptide, an antibody thereto or a TAT binding oligopeptide) to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes.

A "small" molecule or "small" organic molecule is defined herein to have a molecular weight below about 500 Daltons.

An "effective amount" of a polypeptide, antibody, TAT binding oligopeptide, TAT binding organic molecule or an agonist or antagonist thereof as disclosed herein is an amount sufficient to carry out a specifically stated purpose. An "effective amount" may be determined empirically and in a routine manner, in relation to the stated purpose.

The term "therapeutically effective amount" refers to an amount of an antibody, polypeptide, TAT binding oligopeptide, TAT binding organic molecule or other drug effective to "treat" a disease or disorder in a subject or mammal. In the case of cancer, the therapeutically effective amount of the drug may reduce the number of cancer cells; reduce the tumor size; inhibit (i.e., slow to some extent and preferably stop) cancer cell infiltration into peripheral organs; inhibit (i.e., slow to some extent and preferably stop) tumor metastasis; inhibit, to some extent, tumor growth; and/or relieve to some extent one or more of the symptoms associated with the cancer. See the definition herein of "treating". To the extent the drug may prevent growth and/or kill existing cancer cells, it may be cytostatic and/or cytotoxic.

A "growth inhibitory amount" of an anti-TAT antibody, TAT polypeptide, TAT binding oligopeptide, TAT siRNA (such as a TAT188 siRNA) or TAT binding organic molecule is an amount capable of inhibiting the growth of a cell, especially tumor, e.g., cancer cell, either *in vitro* or *in vivo*. A "growth inhibitory amount" of an anti-TAT antibody, TAT polypeptide, TAT binding oligopeptide, TAT siRNA (such as a TAT188 siRNA) or TAT binding organic molecule for purposes of inhibiting neoplastic cell growth may be determined empirically and in a routine manner.

A "cytotoxic amount" of an anti-TAT antibody, TAT polypeptide, TAT binding oligopeptid, TAT

siRNA (such as a TAT188 siRNA) e or TAT binding organic molecule is an amount capable of causing the destruction of a cell, especially tumor, e.g., cancer cell, either *in vitro* or *in vivo*. A "cytotoxic amount" of an anti-TAT antibody, TAT polypeptide, TAT binding oligopeptide, TAT siRNA (such as a TAT188 siRNA) or TAT binding organic molecule for purposes of inhibiting neoplastic cell growth may be determined empirically and in a routine manner.

The term "antibody" is used in the broadest sense and specifically covers, for example, single anti-TAT monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies), anti-TAT antibody compositions with polyepitopic specificity, polyclonal antibodies, single chain anti-TAT antibodies, and fragments of anti-TAT antibodies (see below) as long as they exhibit the desired biological or immunological activity. The term "immunoglobulin" (Ig) is used interchangeable with antibody herein.

An "isolated antibody" is one which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or nonreducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody in situ within recombinant cells since at least one component of the antibody's natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

The basic 4-chain antibody unit is a heterotetrameric glycoprotein composed of two identical light (L) chains and two identical heavy (H) chains (an IgM antibody consists of 5 of the basic heterotetramer unit along with an additional polypeptide called J chain, and therefore contain 10 antigen binding sites, while secreted IgA antibodies can polymerize to form polyvalent assemblages comprising 2-5 of the basic 4-chain units along with J chain). In the case of IgGs, the 4-chain unit is generally about 150,000 daltons. Each L chain is linked to a H chain by one covalent disulfide bond, while the two H chains are linked to each other by one or more disulfide bonds depending on the H chain isotype. Each H and L chain also has regularly spaced intrachain disulfide bridges. Each H chain has at the N-terminus, a variable domain (V_H) followed by three constant domains (C_H) for each of the α and γ chains and four C_H domains for μ and ϵ isotypes. Each L chain has at the N-terminus, a variable domain (V_L) followed by a constant domain (C_L) at its other end. The V_L is aligned with the V_H and the C_L is aligned with the first constant domain of the heavy chain (C_H1). Particular amino acid residues are believed to form an interface between the light chain and heavy chain variable domains. The pairing of a V_H and V_L together forms a single antigen-binding site. For the structure and properties of the different classes of antibodies, see, e.g., Basic and Clinical Immunology, 8th edition, Daniel P. Stites, Abba I. Terr and Tristram G. Parslow (eds.), Appleton & Lange, Norwalk, CT, 1994, page 71 and Chapter 6.

The L chain from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant domains. Depending on the amino acid sequence of the constant domain of their heavy chains (C_H), immunoglobulins can be assigned to different classes or isotypes. There are five classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM, having heavy chains designated α, δ, ε, γ, and μ, respectively. The γ and α classes are further divided into subclasses on the basis of relatively minor differences in C_H sequence and function, e.g., humans express the following subclasses: IgG1, IgG2, IgG3, IgG4, IgA1, and IgA2.

The term "variable" refers to the fact that certain segments of the variable domains differ extensively in sequence among antibodies. The V domain mediates antigen binding and define specificity of a particular antibody for its particular antigen. However, the variability is not evenly distributed across the 110-amino acid span of the variable domains. Instead, the V regions consist of relatively invariant stretches called framework regions (FRs) of 15-30 amino acids separated by shorter regions of extreme variability called "hypervariable regions" that are each 9-12 amino acids long. The variable domains of native heavy and light chains each comprise four FRs, largely adopting a β-sheet configuration, connected by three hypervariable regions, which form loops connecting, and in some cases forming part of, the β-sheet structure. The hypervariable regions in each chain are held together in close proximity by the FRs and, with the hypervariable regions from the other chain, contribute to the formation of the antigen-binding site of antibodies (see Kabat et al., Sequences of Proteins of Immunological Interest, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD. (1991)). The constant domains are not involved directly in binding an antibody to an antigen, but exhibit various effector functions, such as participation of the antibody in antibody dependent cellular cytotoxicity (ADCC).

The term "hypervariable region" when used herein refers to the amino acid residues of an antibody which are responsible for antigen-binding. The hypervariable region (HVR) generally comprises amino acid residues from a "complementarity determining region" or "CDR" (e.g. around about residues 24-34 (L1), 50-56 (L2) and 89-97 (L3) in the V_L, and around about 1-35 (H1), 50-65 (H2) and 95-102 (H3) in the V_H; Kabat et al., Sequences of Proteins of Immunological Interest, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD. (1991)) and/or those residues from a "hypervariable loop" (*e.g.* residues 26-32 (L1), 50-52 (L2) and 91-96 (L3) in the V_L, and 26-32 (H1), 53-55 (H2) and 96-101 (H3) in the V_H; Chothia and Lesk *J. Mol. Biol.* 196:901-917 (1987)). The HVRs are also defined as encompassing the sequences as disclosed in US Application Serial No. , filed February 19, 2004, incorporated herein by reference in its entirety, wherein extended HVRs include some of the amino acid sequence positions within the variable regions that are defined as being in contact with a bound ligand of the antibody based on an analysis of complex crystal structures. As used herein the terms "HVR" and "CDR" are used interchangeably to encompass the extend HVRs defined as follows.

The term "hypervariable region" when used herein refers to the regions of an antibody variable domain which are hypervariable in sequence and/or form structurally defined loops. Generally, antibodies comprise six hypervariable regions; three in the VH (H1, H2, H3), and three in the VL (L1, L2, L3). A number of hypervariable region delineations are in use and are encompassed herein. The Kabat Complementarity Determining Regions (CDRs) are based on sequence variability and are the most commonly used (Kabat *et al., Sequences of Proteins of Immunological Interest*, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD. (1991)). Chothia refers instead to the location of the structural loops (Chothia and Lesk *J. Mol. Biol.* 196:901-917 (1987)). The AbM hypervariable regions represent a compromise between the Kabat CDRs and Chothia structural loops, and are used by Oxford Molecular's AbM antibody modeling software. The "contact" hypervariable regions are based on an analysis of the available complex crystal structures. The residues from each of these hypervariable regions are noted below.

Table HVR

Loop	Kabat	AbM	Chothia	Contact
L1	L24-L34	L24-L34	L26-L32	L30-L36
L2	L50-L56	L50-L56	L50-L52	L46-L55
L3	L89-L97	L89-L97	L91-L96	L89-L96
H1 (Kabat Numbering)	H31-H35B	H26-H35B	H26-H32..34	H30-H35B
H1 (Chothia Numbering)	H31-H35	H26-H35	H26-H32	H30-H35
H2	H50-H65	H50-H58	H53-H55	H47-H58
H3	H95-H102	H95-H102	H96-H101	H93-H101

Hypervariable regions may comprise "extended hypervariable regions" as follows: 24-34 (L1), 50-56 (L2) and 89-97 (L3) in the VL and 26-35 (H1), 50-65 or 49-65 (H2) and 93-102, 94-102 or 95-102 (H3) in the VH. The variable domain residues are numbered according to Kabat *et al., supra* for each of these definitions.

"Framework" or "FR" residues are those variable domain residues other than the hypervariable region residues as herein defined.

The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally occurring mutations that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site. Furthermore, in contrast to polyclonal antibody preparations which include different antibodies directed against different determinants (epitopes), each monoclonal antibody is directed against a single determinant on the antigen. In addition to their specificity, the monoclonal antibodies are advantageous in that they may be synthesized uncontaminated by other antibodies. The modifier "monoclonal" is not to be construed as requiring production of the antibody by any particular method. For example, the monoclonal antibodies useful in the present invention may be prepared by the hybridoma methodology first described by Kohler et al., Nature, 256:495 (1975), or may be made using recombinant DNA methods in bacterial, eukaryotic animal or plant cells (see, e.g., U.S. Patent No. 4,816,567). The "monoclonal antibodies" may also be isolated from phage antibody libraries using the techniques described in Clackson et al., Nature, 352:624-628 (1991) and Marks et al., J. Mol. Biol., 222:581-597 (1991), for example.

The monoclonal antibodies herein include "chimeric" antibodies in which a portion of the heavy and/or light chain is identical with or homologous to corresponding sequences in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is identical with or homologous to corresponding sequences in antibodies derived from another species or belonging to another antibody class or subclass, as well as fragments of such antibodies, so long as they exhibit the desired biological activity (see U.S. Patent No. 4,816,567; and Morrison et al., Proc. Natl. Acad. Sci. USA, 81:6851-6855 (1984)). Chimeric antibodies of interest herein include "primatized" antibodies comprising variable domain antigen-binding sequences derived from a non-human primate (*e.g.* Old World Monkey, Ape etc), and human constant region sequences.

An "intact" antibody is one which comprises an antigen-binding site as well as a C_L and at least heavy chain constant domains, C_H1, C_H2 and C_H3. The constant domains may be native sequence constant domains (*e.g.* human native sequence constant domains) or amino acid sequence variant thereof. Preferably, the intact antibody has one or more effector functions.

"Antibody fragments" comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', $F(ab')_2$, and Fv fragments; diabodies; linear antibodies (see U.S. Patent No. 5,641,870, Example 2; Zapata et al., Protein Eng. 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. The Fab fragment consists of an entire L chain along with the variable region domain of the H chain (V_H), and the first constant domain of one heavy chain (C_H1). Each Fab fragment is monovalent with respect to antigen binding, i.e., it has a single antigen-binding site. Pepsin treatment of an antibody yields a single large $F(ab')_2$ fragment

which roughly corresponds to two disulfide linked Fab fragments having divalent antigen-binding activity and is still capable of cross-linking antigen. Fab' fragments differ from Fab fragments by having additional few residues at the carboxy terminus of the C_H1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. $F(ab')_2$ antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

The Fc fragment comprises the carboxy-terminal portions of both H chains held together by disulfides. The effector functions of antibodies are determined by sequences in the Fc region, which region is also the part recognized by Fc receptors (FcR) found on certain types of cells.

"Fv" is the minimum antibody fragment which contains a complete antigen-recognition and -binding site. This fragment consists of a dimer of one heavy- and one light-chain variable region domain in tight, non-covalent association. From the folding of these two domains emanate six hypervariable loops (3 loops each from the H and L chain) that contribute the amino acid residues for antigen binding and confer antigen binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

"Single-chain Fv" also abbreviated as "sFv" or "scFv" are antibody fragments that comprise the V_H and V_L antibody domains connected into a single polypeptide chain. Preferably, the sFv polypeptide further comprises a polypeptide linker between the V_H and V_L domains which enables the sFv to form the desired structure for antigen binding. For a review of sFv, see Pluckthun in The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenburg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994); Borrebaeck 1995, infra.

The term "diabodies" refers to small antibody fragments prepared by constructing sFv fragments (see preceding paragraph) with short linkers (about 5-10 residues) between the V_H and V_L domains such that inter-chain but not intra-chain pairing of the V domains is achieved, resulting in a bivalent fragment, i.e., fragment having two antigen-binding sites. Bispecific diabodies are heterodimers of two "crossover" sFv fragments in which the V_H and V_L domains of the two antibodies are present on different polypeptide chains. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and Hollinger et al., Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993).

"Humanized" forms of non-human (*e.g.*, rodent) antibodies are chimeric antibodies that contain minimal sequence derived from the non-human antibody. For the most part, humanized antibodies are human immunoglobulins (recipient antibody) in which residues from a hypervariable region of the recipient are replaced by residues from a hypervariable region of a non-human species (donor antibody) such as mouse, rat, rabbit or non-human primate having the desired antibody specificity, affinity, and capability. In some instances, framework region (FR) residues of the human immunoglobulin are replaced by corresponding non-human residues. Furthermore, humanized antibodies may comprise residues that are not found in the recipient antibody or in the donor antibody. These modifications are made to further refine antibody performance. In

general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the hypervariable loops correspond to those of a non-human immunoglobulin and all or substantially all of the FRs are those of a human immunoglobulin sequence. The humanized antibody optionally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin. For further details, see Jones et al., Nature 321:522-525 (1986); Riechmann et al., Nature 332:323-329 (1988); and Presta, Curr. Op. Struct. Biol. 2:593-596 (1992).

A "species-dependent antibody," e.g., a mammalian anti-human IgE antibody, is an antibody which has a stronger binding affinity for an antigen from a first mammalian species than it has for a homologue of that antigen from a second mammalian species. Normally, the species-dependent antibody "bind specifically" to a human antigen (i.e., has a binding affinity (Kd) value of no more than about 1×10^{-7} M, preferably no more than about 1×10^{-8} and most preferably no more than about 1×10^{-9} M) but has a binding affinity for a homologue of the antigen from a second non-human mammalian species which is at least about 50 fold, or at least about 500 fold, or at least about 1000 fold, weaker than its binding affinity for the human antigen. The species-dependent antibody can be of any of the various types of antibodies as defined above, but preferably is a humanized or human antibody.

A "TAT binding oligopeptide" is an oligopeptide that binds, preferably specifically, to a TAT polypeptide as described herein. TAT binding oligopeptides may be chemically synthesized using known oligopeptide synthesis methodology or may be prepared and purified using recombinant technology. TAT binding oligopeptides are usually at least about 5 amino acids in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 amino acids in length or more, wherein such oligopeptides that are capable of binding, preferably specifically, to a TAT polypeptide as described herein. TAT binding oligopeptides may be identified without undue experimentation using well known techniques. In this regard, it is noted that techniques for screening oligopeptide libraries for oligopeptides that are capable of specifically binding to a polypeptide target are well known in the art (see, e.g., U.S. Patent Nos. 5,556,762, 5,750,373, 4,708,871, 4,833,092, 5,223,409, 5,403,484, 5,571,689, 5,663,143; PCT Publication Nos. WO 84/03506 and WO84/03564; Geysen et al., Proc. Natl. Acad. Sci. U.S.A., 81:3998-4002 (1984); Geysen et al., Proc. Natl. Acad. Sci. U.S.A., 82:178-182 (1985); Geysen et al., in Synthetic Peptides as Antigens, 130-149 (1986); Geysen et al., J. Immunol. Meth., 102:259-274 (1987); Schoofs et al., J. Immunol., 140:611-616 (1988), Cwirla, S. E. et al. (1990) Proc. Natl. Acad. Sci. USA, 87:6378; Lowman, H.B. et al. (1991) Biochemistry, 30:10832; Clackson, T. et al. (1991) Nature, 352: 624; Marks, J. D. et al. (1991), J. Mol. Biol., 222:581; Kang, A.S. et al. (1991) Proc. Natl. Acad. Sci. USA, 88:8363, and Smith, G. P. (1991) Current Opin. Biotechnol., 2:668).

A "TAT binding organic molecule" is an organic molecule other than an oligopeptide or antibody as defined herein that binds, preferably specifically, to a TAT polypeptide as described herein. TAT binding

organic molecules may be identified and chemically synthesized using known methodology (see, e.g., PCT Publication Nos. WO00/00823 and WO00/39585). TAT binding organic molecules are usually less than about 2000 daltons in size, alternatively less than about 1500, 750, 500, 250 or 200 daltons in size, wherein such organic molecules that are capable of binding, preferably specifically, to a TAT polypeptide as described herein may be identified without undue experimentation using well known techniques. In this regard, it is noted that techniques for screening organic molecule libraries for molecules that are capable of binding to a polypeptide target are well known in the art (see, e.g., PCT Publication Nos. WO00/00823 and WO00/39585).

An "interfering RNA" or "small interfering RNA (siRNA)" is a double stranded RNA molecule less than 30 nucleotides in length that reduces expression of a target gene. A "TAT interfering RNA" or "TAT siRNA" binds, preferably specifically, to a TAT nucleic acid and reduces its expression. This means the expression of the TAT molecule is lower with the interfering RNA present as compared to expression of the TAT molecule in the control where the interfering RNA is not present. TAT interfering RNAs may be identified and synthesized using known methods (Shi Y.*Trends in Genetics* 19(1):9-12 (2003), WO/2003056012 and WO2003064621).

An antibody, oligopeptide, siRNA or other organic molecule "which binds" an antigen of interest, e.g. a tumor-associated polypeptide antigen target, is one that binds the antigen with sufficient affinity such that the antibody, oligopeptide or other organic molecule is useful as a diagnostic and/or therapeutic agent in targeting a cell or tissue expressing the antigen, and does not significantly cross-react with other proteins. In such embodiments, the extent of binding of the antibody, oligopeptide, siRNA or other organic molecule to a "non-target" protein will be less than about 10% of the binding of the antibody, oligopeptide, siRNA or other organic molecule to its particular target protein as determined by fluorescence activated cell sorting (FACS) analysis or radioimmunoprecipitation (RIA). With regard to the binding of an antibody, oligopeptide or other organic molecule to a target molecule, the term "specific binding" or "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide target means binding that is measurably different from a non-specific interaction. Specific binding can be measured, for example, by determining binding of a molecule compared to binding of a control molecule, which generally is a molecule of similar structure that does not have binding activity. For example, specific binding can be determined by competition with a control molecule that is similar to the target, for example, an excess of non-labeled target. In this case, specific binding is indicated if the binding of the labeled target to a probe is competitively inhibited by excess unlabeled target. The term "specific binding" or "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide target as used herein can be exhibited, for example, by a molecule having a Kd for the target of at least about 10^{-4} M, alternatively at least about 10^{-5} M, alternatively at least about 10^{-6} M, alternatively at least about 10^{-7} M, alternatively at least about 10^{-8} M, alternatively at least about 10^{-9} M, alternatively at least about 10^{-10} M, alternatively at least about 10^{-11} M, alternatively at least about 10^{-12} M, or greater. In one embodiment, the term "specific binding" refers to binding where a molecule

binds to a particular polypeptide or epitope on a particular polypeptide without substantially binding to any other polypeptide or polypeptide epitope.

An antibody, oligopeptide, siRNA or other organic molecule that "inhibits the growth of tumor cells expressing a TAT polypeptide" or a "growth inhibitory" antibody, oligopeptide or other organic molecule is one which results in measurable growth inhibition of cancer cells expressing or overexpressing the appropriate TAT polypeptide. The TAT polypeptide may be a transmembrane polypeptide expressed on the surface of a cancer cell or may be a polypeptide that is produced and secreted by a cancer cell. Preferred growth inhibitory anti-TAT antibodies, oligopeptides or organic molecules inhibit growth of TAT-expressing tumor cells by greater than 20%, preferably from about 20% to about 50%, and even more preferably, by greater than 50% (e.g., from about 50% to about 100%) as compared to the appropriate control, the control typically being tumor cells not treated with the antibody, oligopeptide or other organic molecule being tested. In one embodiment, growth inhibition can be measured at an antibody concentration of about 0.1 to 30 μg/ml or about 0.5 nM to 200 nM in cell culture, where the growth inhibition is determined 1-10 days after exposure of the tumor cells to the antibody. Growth inhibition of tumor cells *in vivo* can be determined in various ways such as is described in the Experimental Examples section below. The antibody is growth inhibitory in vivo if administration of the anti-TAT antibody at about 1 μg/kg to about 100 mg/kg body weight results in reduction in tumor size or tumor cell proliferation within about 5 days to 3 months from the first administration of the antibody, preferably within about 5 to 30 days.

An antibody, oligopeptide, siRNA or other organic molecule which "induces apoptosis" is one which induces programmed cell death as determined by binding of annexin V, fragmentation of DNA, cell shrinkage, dilation of endoplasmic reticulum, cell fragmentation, and/or formation of membrane vesicles (called apoptotic bodies). The cell is usually one which overexpresses a TAT polypeptide. Preferably the cell is a tumor cell, e.g., a prostate, breast, ovarian, stomach, endometrial, lung, kidney, colon, colorectal, bladder cell. Various methods are available for evaluating the cellular events associated with apoptosis. For example, phosphatidyl serine (PS) translocation can be measured by annexin binding; DNA fragmentation can be evaluated through DNA laddering; and nuclear/chromatin condensation along with DNA fragmentation can be evaluated by any increase in hypodiploid cells. Preferably, the antibody, oligopeptide or other organic molecule which induces apoptosis is one which results in about 2 to 50 fold, preferably about 5 to 50 fold, and most preferably about 10 to 50 fold, induction of annexin binding relative to untreated cell in an annexin binding assay.

Antibody "effector functions" refer to those biological activities attributable to the Fc region (a native sequence Fc region or amino acid sequence variant Fc region) of an antibody, and vary with the antibody isotype. Examples of antibody effector functions include: C1q binding and complement dependent cytotoxicity; Fc receptor binding; antibody-dependent cell-mediated cytotoxicity (ADCC); phagocytosis; down regulation of cell surface receptors (e.g., B cell receptor); and B cell activation.

"Antibody-dependent cell-mediated cytotoxicity" or "ADCC" refers to a form of cytotoxicity in which secreted Ig bound onto Fc receptors (FcRs) present on certain cytotoxic cells (e.g., Natural Killer (NK)

cells, neutrophils, and macrophages) enable these cytotoxic effector cells to bind specifically to an antigen-bearing target cell and subsequently kill the target cell with cytotoxins. The antibodies "arm" the cytotoxic cells and are absolutely required for such killing. The primary cells for mediating ADCC, NK cells, express FcγRIII only, whereas monocytes express FcγRI, FcγRII and FcγRIII. FcR expression on hematopoietic cells is summarized in Table 3 on page 464 of Ravetch and Kinet, Annu. Rev. Immunol. 9:457-92 (1991). To assess ADCC activity of a molecule of interest, an in vitro ADCC assay, such as that described in US Patent No. 5,500,362 or 5,821,337 may be performed. Useful effector cells for such assays include peripheral blood mononuclear cells (PBMC) and Natural Killer (NK) cells. Alternatively, or additionally, ADCC activity of the molecule of interest may be assessed in vivo, e.g., in a animal model such as that disclosed in Clynes et al. (USA) 95:652-656 (1998).

"Fc receptor" or "FcR" describes a receptor that binds to the Fc region of an antibody. The preferred FcR is a native sequence human FcR. Moreover, a preferred FcR is one which binds an IgG antibody (a gamma receptor) and includes receptors of the FcγRI, FcγRII and FcγRIII subclasses, including allelic variants and alternatively spliced forms of these receptors. FcγRII receptors include FcγRIIA (an "activating receptor") and FcγRIIB (an "inhibiting receptor"), which have similar amino acid sequences that differ primarily in the cytoplasmic domains thereof. Activating receptor FcγRIIA contains an immunoreceptor tyrosine-based activation motif (ITAM) in its cytoplasmic domain. Inhibiting receptor FcγRIIB contains an immunoreceptor tyrosine-based inhibition motif (ITIM) in its cytoplasmic domain. (see review M. in Daëron, Annu. Rev. Immunol. 15:203-234 (1997)). FcRs are reviewed in Ravetch and Kinet, Annu. Rev. Immunol. 9:457-492 (1991); Capel et al., Immunomethods 4:25-34 (1994); and de Haas et al., J. Lab. Clin. Med. 126:330-41 (1995). Other FcRs, including those to be identified in the future, are encompassed by the term "FcR" herein. The term also includes the neonatal receptor, FcRn, which is responsible for the transfer of maternal IgGs to the fetus (Guyer et al., J. Immunol. 117:587 (1976) and Kim et al., J. Immunol. 24:249 (1994)).

"Human effector cells" are leukocytes which express one or more FcRs and perform effector functions. Preferably, the cells express at least FcγRIII and perform ADCC effector function. Examples of human leukocytes which mediate ADCC include peripheral blood mononuclear cells (PBMC), natural killer (NK) cells, monocytes, cytotoxic T cells and neutrophils; with PBMCs and NK cells being preferred. The effector cells may be isolated from a native source, e.g., from blood.

"Complement dependent cytotoxicity" or "CDC" refers to the lysis of a target cell in the presence of complement. Activation of the classical complement pathway is initiated by the binding of the first component of the complement system (C1q) to antibodies (of the appropriate subclass) which are bound to their cognate antigen. To assess complement activation, a CDC assay, e.g., as described in Gazzano-Santoro et al., J. Immunol. Methods 202:163 (1996), may be performed.

The terms "cancer" and "cancerous" refer to or describe the physiological condition in mammals that is typically characterized by unregulated cell growth. Examples of cancer include, but are not limited to, carcinoma, lymphoma, blastoma, sarcoma, and leukemia or lymphoid malignancies. More particular

examples of such cancers include squamous cell cancer (e.g., epithelial squamous cell cancer), lung cancer including small-cell lung cancer, non-small cell lung cancer, adenocarcinoma of the lung and squamous carcinoma of the lung, cancer of the peritoneum, hepatocellular cancer, gastric or stomach cancer including gastrointestinal cancer, pancreatic cancer, glioblastoma, cervical cancer, ovarian cancer, liver cancer, bladder cancer, cancer of the urinary tract, hepatoma, breast cancer, colon cancer, rectal cancer, colorectal cancer, endometrial or uterine carcinoma, salivary gland carcinoma, kidney or renal cancer, prostate cancer, vulval cancer, thyroid cancer, hepatic carcinoma, anal carcinoma, penile carcinoma, melanoma, multiple myeloma and B-cell lymphoma, brain, as well as head and neck cancer, and associated metastases.

The terms "cell proliferative disorder" and "proliferative disorder" refer to disorders that are associated with some degree of abnormal cell proliferation. In one embodiment, the cell proliferative disorder is cancer.

"Tumor", as used herein, refers to all neoplastic cell growth and proliferation, whether malignant or benign, and all pre-cancerous and cancerous cells and tissues.

An antibody, oligopeptide or other organic molecule which "induces cell death" is one which causes a viable cell to become nonviable. The cell is one which expresses a TAT polypeptide, preferably a cell that overexpresses a TAT polypeptide as compared to a normal cell of the same tissue type. The TAT polypeptide may be a transmembrane polypeptide expressed on the surface of a cancer cell or may be a polypeptide that is produced and secreted by a cancer cell. Preferably, the cell is a cancer cell, e.g., a breast, ovarian, stomach, endometrial, salivary gland, lung, kidney, colon, colorectal, thyroid, pancreatic or bladder cell. Cell death *in vitro* may be determined in the absence of complement and immune effector cells to distinguish cell death induced by antibody-dependent cell-mediated cytotoxicity (ADCC) or complement dependent cytotoxicity (CDC). Thus, the assay for cell death may be performed using heat inactivated serum (i.e., in the absence of complement) and in the absence of immune effector cells. To determine whether the antibody, oligopeptide or other organic molecule is able to induce cell death, loss of membrane integrity as evaluated by uptake of propidium iodide (PI), trypan blue (see Moore et al. Cytotechnology 17:1-11 (1995)) or 7AAD can be assessed relative to untreated cells. Preferred cell death-inducing antibodies, oligopeptides or other organic molecules are those which induce PI uptake in the PI uptake assay in BT474 cells.

A "TAT-expressing cell" is a cell which expresses an endogenous or transfected TAT polypeptide either on the cell surface or in a secreted form. A "TAT-expressing cancer" is a cancer comprising cells that have a TAT polypeptide present on the cell surface or that produce and secrete a TAT polypeptide. A "TAT-expressing cancer" optionally produces sufficient levels of TAT polypeptide on the surface of cells thereof, such that an anti-TAT antibody, oligopeptide or other organic molecule can bind thereto and have a therapeutic effect with respect to the cancer. In another embodiment, a "TAT-expressing cancer" optionally produces and secretes sufficient levels of TAT polypeptide, such that an anti-TAT antibody, oligopeptide or other organic molecule antagonist can bind thereto and have a therapeutic effect with respect to the cancer. With regard to the latter, the antagonist may be an antisense oligonucleotide which reduces, inhibits or prevents production and secretion of the secreted TAT polypeptide by tumor cells. A cancer which

"overexpresser" a TAT polypeptide is one which has significantly higher levels of TAT polypeptide at the cell surface thereof, or produces and secretes, compared to a noncancerous cell of the same tissue type. Such overexpression may be caused by gene amplification or by increased transcription or translation. TAT polypeptide overexpression may be determined in a diagnostic or prognostic assay by evaluating increased levels of the TAT protein present on the surface of a cell, or secreted by the cell (e.g., via an immunohistochemistry assay using anti-TAT antibodies prepared against an isolated TAT polypeptide which may be prepared using recombinant DNA technology from an isolated nucleic acid encoding the TAT polypeptide; FACS analysis, etc.). Alternatively, or additionally, one may measure levels of TAT polypeptide-encoding nucleic acid or mRNA in the cell, e.g., via fluorescent *in situ* hybridization using a nucleic acid based probe corresponding to a TAT-encoding nucleic acid or the complement thereof; (FISH; see WO98/45479 published October, 1998), Southern blotting, Northern blotting, or polymerase chain reaction (PCR) techniques, such as real time quantitative PCR (RT-PCR). One may also study TAT polypeptide overexpression by measuring shed antigen in a biological fluid such as serum, e.g, using antibody-based assays (see also, e.g., U.S. Patent No. 4,933,294 issued June 12, 1990; WO91/05264 published April 18, 1991; U.S. Patent 5,401,638 issued March 28, 1995; and Sias et al., J. Immunol. Methods 132:73-80 (1990)). Aside from the above assays, various *in vivo* assays are available to the skilled practitioner. For example, one may expose cells within the body of the patient to an antibody which is optionally labeled with a detectable label, e.g., a radioactive isotope, and binding of the antibody to cells in the patient can be evaluated, e.g., by external scanning for radioactivity or by analyzing a biopsy taken from a patient previously exposed to the antibody.

As used herein, the term "immunoadhesin" designates antibody-like molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site of an antibody (i.e., is "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD or IgM.

The word "label" when used herein refers to a detectable compound or composition which is conjugated directly or indirectly to the antibody, oligopeptide or other organic molecule so as to generate a "labeled" antibody, oligopeptide or other organic molecule. The label may be detectable by itself (e.g. radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition which is detectable.

The term "cytotoxic agent" as used herein refers to a substance that inhibits or prevents the function of cells and/or causes destruction of cells. The term is intended to include radioactive isotopes (e.g., At^{211}, I^{131}, I^{125}, Y^{90}, Re^{186}, Re^{188}, Sm^{153}, Bi^{212}, P^{32} and radioactive isotopes of Lu), chemotherapeutic agents e.g.

methotrexate, adriamicin, vinca alkaloids (vincristine, vinblastine, etoposide), doxorubicin, melphalan, mitomycin C, chlorambucil, daunorubicin or other intercalating agents, enzymes and fragments thereof such as nucleolytic enzymes, antibiotics, and toxins such as small molecule toxins or enzymatically active toxins of bacterial, fungal, plant or animal origin, including fragments and/or variants thereof, and the various antitumor or anticancer agents disclosed below. Other cytotoxic agents are described below. A tumoricidal agent causes destruction of tumor cells. Cytotoxins may be covalently attached to an antibody to target the toxin to a particular cell of interest which expresses the antigen. Useful cytotoxins are their linker include, without limitation, the following:

LINKERS:

MC = maleimidocaproyl

Val Cit = valine-citrulline, dipeptide site in protease cleavable linker.

Citrulline = 2-amino-5-ureido pentanoic acid

PAB = p-aminobenzylcarbamoyl ("self immolative" portion of linker)

Me = N-methyl-valine citrulline where the linker peptide bond has been modified to prevent its cleavage by cathepsin B

MC(PEG)6-OH = maleimidocaproyl- polyethylene glycol, attached to antibody cysteines.

SPP = N-Succinimidyl 4-(2-pyridylthio) pentanoate

SMCC = N-Succinimidyl 4-(N-maleimidomethyl) cyclohexane-1 carboxylate

CYTOTOXIC DRUGS:

MMAE = mono-methyl auristatin E (MW 718)

MMAF = variant of auristatin E (MMAE) with a phenylalanine at the C-terminus of the drug (MW 731.5)

AEVB = auristatin E valeryl benzylhydrazone, acid labile linker through the C-terminus of AE (MW 732)

AFP =Auristatin F phenylene diamine; (the phenylalanine variant linked to the antibody through the C-terminus via a phenylene diamine spacer) (MW 732).

A "growth inhibitory agent" when used herein refers to a compound or composition which inhibits growth of a cell, especially a TAT-expressing cancer cell, either *in vitro* or *in vivo*. Thus, the growth inhibitory agent may be one which significantly reduces the percentage of TAT-expressing cells in S phase. Examples of growth inhibitory agents include agents that block cell cycle progression (at a place other than S phase), such as agents that induce G1 arrest and M-phase arrest. Classical M-phase blockers include the vincas (vincristine and vinblastine), taxanes, and topoisomerase II inhibitors such as doxorubicin, epirubicin, daunorubicin, etoposide, and bleomycin. Those agents that arrest G1 also spill over into S-phase arrest, for example, DNA alkylating agents such as tamoxifen, prednisone, dacarbazine, mechlorethamine, cisplatin, methotrexate, 5-fluorouracil, and ara-C. Further information can be found in The Molecular Basis of Cancer, Mendelsohn and Israel, eds., Chapter 1, entitled "Cell cycle regulation, oncogenes, and antineoplastic drugs" by Murakami et al. (WB Saunders: Philadelphia, 1995), especially p. 13. The taxanes (paclitaxel and docetaxel) are anticancer

drugs both derived from the yew tree. Docetaxel (TAXOTHERE®, Rhone-Poulenc Rorer), derived from the European yew, is a semisynthetic analogue of paclitaxel (TAXOL®, Bristol-Myers Squibb). Paclitaxel and docetaxel promote the assembly of microtubules from tubulin dimers and stabilize microtubules by preventing depolymerization, which results in the inhibition of mitosis in cells.

"Doxorubicin" is an anthracycline antibiotic. The full chemical name of doxorubicin is (8S-cis)-10-[(3-amino-2,3,6-trideoxy-α-L-lyxo-hexapyranosyl)oxy]-7,8,9,10-tetrahydro-6,8,11-trihydroxy-8-(hydroxyacetyl)-1-methoxy-5,12-naphthacenedione.

The term "cytokine" is a generic term for proteins released by one cell population which act on another cell as intercellular mediators. Examples of such cytokines are lymphokines, monokines, and traditional polypeptide hormones. Included among the cytokines are growth hormone such as human growth hormone, N-methionyl human growth hormone, and bovine growth hormone; parathyroid hormone; thyroxine; insulin; proinsulin; relaxin; prorelaxin; glycoprotein hormones such as follicle stimulating hormone (FSH), thyroid stimulating hormone (TSH), and luteinizing hormone (LH); hepatic growth factor; fibroblast growth factor; prolactin; placental lactogen; tumor necrosis factor-α and -β; mullerian-inhibiting substance; mouse gonadotropin-associated peptide; inhibin; activin; vascular endothelial growth factor; integrin; thrombopoietin (TPO); nerve growth factors such as NGF-β; platelet-growth factor; transforming growth factors (TGFs) such as TGF-α and TGF-β; insulin-like growth factor-I and -II; erythropoietin (EPO); osteoinductive factors; interferons such as interferon -α, -β, and -γ; colony stimulating factors (CSFs) such as macrophage-CSF (M-CSF); granulocyte-macrophage-CSF (GM-CSF); and granulocyte-CSF (G-CSF); interleukins (ILs) such as IL-1, IL-1a, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-11, IL-12; a tumor necrosis factor such as TNF-α or TNF-β; and other polypeptide factors including LIF and kit ligand (KL). As used herein, the term cytokine includes proteins from natural sources or from recombinant cell culture and biologically active equivalents of the native sequence cytokines.

The term "package insert" is used to refer to instructions customarily included in commercial packages of therapeutic products, that contain information about the indications, usage, dosage, administration, contraindications and/or warnings concerning the use of such therapeutic products.

Table 1

```
/*
*
* C-C increased from 12 to 15
* Z is average of EQ
* B is average of ND
* match with stop is _M; stop-stop = 0; J (joker) match = 0
*/
#define _M      -8      /* value of a match with a stop */

int     _day[26][26] = {
/*    A B C D E F G H I J K L M N O P Q R S T U V W X Y Z */
/* A */  { 2, 0,-2, 0, 0,-4, 1,-1,-1, 0,-1,-2,-1, 0,_M, 1, 0,-2, 1, 1, 0, 0,-6, 0,-3, 0},
/* B */  { 0, 3,-4, 3, 2,-5, 0, 1,-2, 0, 0,-3,-2, 2,_M,-1, 1, 0, 0, 0, 0,-2,-5, 0,-3, 1},
/* C */  {-2,-4,15,-5,-5,-4,-3,-3,-2, 0,-5,-6,-5,-4,_M,-3,-5,-4, 0,-2, 0,-2,-8, 0, 0,-5},
/* D */  { 0, 3,-5, 4, 3,-6, 1, 1,-2, 0, 0,-4,-3, 2,_M,-1, 2,-1, 0, 0, 0,-2,-7, 0,-4, 2},
/* E */  { 0, 2,-5, 3, 4,-5, 0, 1,-2, 0, 0,-3,-2, 1,_M,-1, 2,-1, 0, 0, 0,-2,-7, 0,-4, 3},
/* F */  {-4,-5,-4,-6,-5, 9,-5,-2, 1, 0,-5, 2, 0,-4,_M,-5,-5,-4,-3,-3, 0,-1, 0, 0, 7,-5},
/* G */  { 1, 0,-3, 1, 0,-5, 5,-2,-3, 0,-2,-4,-3, 0,_M,-1,-1,-3, 1, 0, 0,-1,-7, 0,-5, 0},
/* H */  {-1, 1,-3, 1, 1,-2,-2, 6,-2, 0, 0,-2,-2, 2,_M, 0, 3, 2,-1,-1, 0,-2,-3, 0, 0, 2},
/* I */  {-1,-2,-2,-2,-2, 1,-3,-2, 5, 0,-2, 2, 2,-2,_M,-2,-2,-2,-1, 0, 0, 4,-5, 0,-1,-2},
/* J */  { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,_M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */  {-1, 0,-5, 0, 0,-5,-2, 0,-2, 0, 5,-3, 0, 1,_M,-1, 1, 3, 0, 0, 0,-2,-3, 0,-4, 0},
/* L */  {-2,-3,-6,-4,-3, 2,-4,-2, 2, 0,-3, 6, 4,-3,_M,-3,-2,-3,-3,-1, 0, 2,-2, 0,-1,-2},
/* M */  {-1,-2,-5,-3,-2, 0,-3,-2, 2, 0, 0, 4, 6,-2,_M,-2,-1, 0,-2,-1, 0, 2,-4, 0,-2,-1},
/* N */  { 0, 2,-4, 2, 1,-4, 0, 2,-2, 0, 1,-3,-2, 2,_M,-1, 1, 0, 1, 0, 0,-2,-4, 0,-2, 1},
/* O */  {_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M, 0,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M},
/* P */  { 1,-1,-3,-1,-1,-5,-1, 0,-2, 0,-1,-3,-2,-1,_M, 6, 0, 0, 1, 0, 0,-1,-6, 0,-5, 0},
/* Q */  { 0, 1,-5, 2, 2,-5,-1, 3,-2, 0, 1,-2,-1, 1,_M, 0, 4, 1,-1,-1, 0,-2,-5, 0,-4, 3},
/* R */  {-2, 0,-4,-1,-1,-4,-3, 2,-2, 0, 3,-3, 0, 0,_M, 0, 1, 6, 0,-1, 0,-2, 2, 0,-4, 0},
/* S */  { 1, 0, 0, 0, 0,-3, 1,-1,-1, 0, 0,-3,-2, 1,_M, 1,-1, 0, 2, 1, 0,-1,-2, 0,-3, 0},
/* T */  { 1, 0,-2, 0, 0,-3, 0,-1, 0, 0, 0,-1,-1, 0,_M, 0,-1,-1, 1, 3, 0, 0,-5, 0,-3, 0},
/* U */  { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,_M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */  { 0,-2,-2,-2,-2,-1,-1,-2, 4, 0,-2, 2, 2,-2,_M,-1,-2,-2,-1, 0, 0, 4,-6, 0,-2,-2},
/* W */  {-6,-5,-8,-7,-7, 0,-7,-3,-5, 0,-3,-2,-4,-4,_M,-6,-5, 2,-2,-5, 0,-6,17, 0, 0,-6},
/* X */  { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,_M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */  {-3,-3, 0,-4,-4, 7,-5, 0,-1, 0,-4,-1,-2,-2,_M,-5,-4,-4,-3,-3, 0,-2, 0, 0,10,-4},
/* Z */  { 0, 1,-5, 2, 3,-5, 0, 2,-2, 0, 0,-2,-1, 1,_M, 0, 3, 0, 0, 0, 0,-2,-6, 0,-4, 4}
};
```

Table 1 (cont')

```
/*
 */
#include <stdio.h>
#include <ctype.h>

#define  MAXJMP     16      /* max jumps in a diag */
#define  MAXGAP     24      /* don't continue to penalize gaps larger than this */
#define  JMPS       1024    /* max jmps in an path */
#define  MX         4       /* save if there's at least MX-1 bases since last jmp */

#define  DMAT       3       /* value of matching bases */
#define  DMIS       0       /* penalty for mismatched bases */
#define  DINS0      8       /* penalty for a gap */
#define  DINS1      1       /* penalty per base */
#define  PINS0      8       /* penalty for a gap */
#define  PINS1      4       /* penalty per residue */

struct jmp {
        short           n[MAXJMP];      /* size of jmp (neg for dely) */
        unsigned short  x[MAXJMP];      /* base no. of jmp in seq x */
};                                      /* limits seq to 2^16 -1 */

struct diag {
        int             score;          /* score at last jmp */
        long            offset;         /* offset of prev block */
        short           ijmp;           /* current jmp index */
        struct jmp      jp;             /* list of jmps */
};

struct path {
        int     spc;            /* number of leading spaces */
        short   n[JMPS];/* size of jmp (gap) */
        int     x[JMPS];/* loc of jmp (last elem before gap) */
};

char            *ofile;                 /* output file name */
char            *namex[2];              /* seq names: getseqs() */
char            *prog;                  /* prog name for err msgs */
char            *seqx[2];               /* seqs: getseqs() */
int             dmax;                   /* best diag: nw() */
```

```
        int             dmax0;          /* final diag */
        int             dna;            /* set if dna: main() */
        int             endgaps;        /* set if penalizing end gaps */
        int             gapx, gapy;     /* total gaps in seqs */
        int             len0, len1;     /* seq lens */
        int             ngapx, ngapy;   /* total size of gaps */
        int             smax;           /* max score: nw() */
        int             *xbm;           /* bitmap for matching */
        long            offset;         /* current offset in jmp file */
        struct  diag    *dx;            /* holds diagonals */
        struct  path    pp[2];          /* holds path for seqs */

        char            *calloc(), *malloc(), *index(), *strcpy();
        char            *getseq(), *g_calloc();
```

Table 1 (cont')

```
/* Needleman-Wunsch alignment program
 *
 * usage: progs file1 file2
 *   where file1 and file2 are two dna or two protein sequences.
 *   The sequences can be in upper- or lower-case an may contain ambiguity
 *   Any lines beginning with ';', '>' or '<' are ignored
 *   Max file length is 65535 (limited by unsigned short x in the jmp struct)
 *   A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
 *   Output is in the file "align.out"
 *
 * The program may create a tmp file in /tmp to hold info about traceback.
 * Original version developed under BSD 4.3 on a vax 8650
 */
#include "nw.h"
#include "day.h"

static  _dbval[26] = {
        1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};

static  _pbval[26] = {
        1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
        128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
        1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
        1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)                                                    main
        int     ac;
        char    *av[];
{
        prog = av[0];
        if (ac != 3) {
                fprintf(stderr,"usage: %s file1 file2\n", prog);
                fprintf(stderr,"where file1 and file2 are two dna or two protein sequences.\n");
                fprintf(stderr,"The sequences can be in upper- or lower-case\n");
                fprintf(stderr,"Any lines beginning with ';' or '<' are ignored\n");
                fprintf(stderr,"Output is in the file \"align.out\"\n");
                exit(1);
        }
```

```
    namex[0] = av[1];
    namex[1] = av[2];
    seqx[0] = getseq(namex[0], &len0);
    seqx[1] = getseq(namex[1], &len1);
    xbm = (dna)? _dbval : _pbval;

    endgaps = 0;                    /* 1 to penalize endgaps */
    ofile = "align.out";            /* output file */

    nw();           /* fill in the matrix, get the possible jmps */
    readjmps();     /* get the actual jmps */
    print();        /* print stats, alignment */

    cleanup(0);     /* unlink any tmp files */
}
```

Table 1 (cont')

```
/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */
nw()                                                                    nw
{
        char            *px, *py;       /* seqs and ptrs */
        int             *ndely, *dely;  /* keep track of dely */
        int             ndelx, delx;    /* keep track of delx */
        int             *tmp;           /* for swapping row0, row1 */
        int             mis;            /* score for each type */
        int             ins0, ins1;     /* insertion penalties */
        register        id;             /* diagonal index */
        register        ij;             /* jmp index */
        register        *col0, *col1;   /* score for curr, last row */
        register        xx, yy;         /* index into seqs */

        dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

        ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
        dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
        col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
        col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
        ins0 = (dna)? DINS0 : PINS0;
        ins1 = (dna)? DINS1 : PINS1;

        smax = -10000;
        if (endgaps) {
                for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
                        col0[yy] = dely[yy] = col0[yy-1] - ins1;
                        ndely[yy] = yy;
                }
                col0[0] = 0;            /* Waterman Bull Math Biol 84 */
        }
        else
                for (yy = 1; yy <= len1; yy++)
                        dely[yy] = -ins0;
```

```
/* fill in match matrix
 */
for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
         */
        if (endgaps) {
                if (xx == 1)
                        col1[0] = delx = -(ins0+ins1);
                else
                        col1[0] = delx = col0[0] - ins1;
                ndelx = xx;
        }
        else {
                col1[0] = 0;
                delx = -ins0;
                ndelx = 0;
        }
```

Table 1 (cont')

```
...IIW
for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
        mis = col0[yy-1];
        if (dna)
                mis += (xbm['px-'A']&xbm['py-'A'])? DMAT : DMIS;
        else
                mis += _day['px-'A']['py-'A'];

        /* update penalty for del in x seq;
         * favor new del over ongong del
         * ignore MAXGAP if weighting endgaps
         */
        if (endgaps || ndely[yy] < MAXGAP) {
                if (col0[yy] - ins0 >= dely[yy]) {
                        dely[yy] = col0[yy] - (ins0+ins1);
                        ndely[yy] = 1;
                } else {
                        dely[yy] -= ins1;
                        ndely[yy]++;
                }
        } else {
                if (col0[yy] - (ins0+ins1) >= dely[yy]) {
                        dely[yy] = col0[yy] - (ins0+ins1);
                        ndely[yy] = 1;
                } else
                        ndely[yy]++;
        }

        /* update penalty for del in y seq;
         * favor new del over ongong del
         */
        if (endgaps || ndelx < MAXGAP) {
                if (col1[yy-1] - ins0 >= delx) {
                        delx = col1[yy-1] - (ins0+ins1);
                        ndelx = 1;
                } else {
                        delx -= ins1;
                        ndelx++;
                }
        } else {
```

```
                if (col1[yy-1] - (ins0+ins1) >= delx) {
                        delx = col1[yy-1] - (ins0+ins1);
                        ndelx = 1;
                } else
                        ndelx++;
        }

        /* pick the maximum score; we're favoring
         * mis over any del and delx over dely
         */
```

81

Table 1 (cont')

```
                                                                    ...IIW
                    id = xx - yy + len1 - 1;
                    if (mis >= delx && mis >= dely[yy])
                            col1[yy] = mis;
                    else if (delx >= dely[yy]) {
                            col1[yy] = delx;
                            ij = dx[id].ijmp;
                            if (dx[id].jp.n[0] && (idna || (ndelx >= MAXJMP
                            && xx > dx[id].jp.x[ij+MX) || mis > dx[id].score+DINS0)) {
                                    dx[id].ijmp++;
                                    if (++ij >= MAXJMP) {
                                            writejmps(id);
                                            ij = dx[id].ijmp = 0;
                                            dx[id].offset = offset;
                                            offset += sizeof(struct jmp) + sizeof(offset);
                                    }
                            }
                            dx[id].jp.n[ij] = ndelx;
                            dx[id].jp.x[ij] = xx;
                            dx[id].score = delx;
                    }
                    else {
                            col1[yy] = dely[yy];
                            ij = dx[id].ijmp;
            if (dx[id].jp.n[0] && (idna || (ndely[yy] >= MAXJMP
                            && xx > dx[id].jp.x[ij+MX) || mis > dx[id].score+DINS0)) {
                                    dx[id].ijmp++;
                                    if (++ij >= MAXJMP) {
                                            writejmps(id);
                                            ij = dx[id].ijmp = 0;
                                            dx[id].offset = offset;
                                            offset += sizeof(struct jmp) + sizeof(offset);
                                    }
                            }
                            dx[id].jp.n[ij] = -ndely[yy];
                            dx[id].jp.x[ij] = xx;
                            dx[id].score = dely[yy];
                    }
                    if (xx == len0 && yy < len1) {
                            /* last col
```

```
                        */
                        if (endgaps)
                                col1[yy] -= ins0+ins1*(len1-yy);
                        if (col1[yy] > smax) {
                                smax = col1[yy];
                                dmax = id;
                        }
                }
        }
        if (endgaps && xx < len0)
                col1[yy-1] -= ins0+ins1*(len0-xx);
        if (col1[yy-1] > smax) {
                smax = col1[yy-1];
                dmax = id;
        }
        tmp = col0; col0 = col1; col1 = tmp;
}
(void) free((char *)ndely);
(void) free((char *)dely);
(void) free((char *)col0);
(void) free((char *)col1);                }
```

Table 1 (cont')

```
/*
*
* print() -- only routine visible outside this module
*
* static:
* getmat() -- trace back best path, count matches: print()
* pr_align() -- print alignment of described in array p[]: print()
* dumpblock() -- dump a block of lines with numbers, stars: pr_align()
* nums() -- put out a number line: dumpblock()
* putline() -- put out a line (name, [num], seq, [num]): dumpblock()
* stars() -- put a line of stars: dumpblock()
* stripname() -- strip any path and prefix from a seqname
*/

#include "nw.h"

#define SPC     3
#define P_LINE  256     /* maximum output line */
#define P_SPC   3       /* space between name or num and seq */

extern  _day[26][26];
int     olen;           /* set output line length */
FILE    *fx;            /* output file */

print()                                                          print
{
        int     lx, ly, firstgap, lastgap;      /* overlap */

        if ((fx = fopen(ofile, "w")) == 0) {
                fprintf(stderr,"%s: can't write %s\n", prog, ofile);
                cleanup(1);
        }
        fprintf(fx, "<first sequence: %s (length = %d)\n", namex[0], len0);
        fprintf(fx, "<second sequence: %s (length = %d)\n", namex[1], len1);
        olen = 60;
        lx = len0;
        ly = len1;
        firstgap = lastgap = 0;
        if (dmax < len1 - 1) {          /* leading gap in x */
                pp[0].spc = firstgap = len1 - dmax - 1;
```

```
            ly -= pp[0].spc;
        }
        else if (dmax > len1 - 1) {     /* leading gap in y */
            pp[1].spc = firstgap = dmax - (len1 - 1);
            lx -= pp[1].spc;
        }
        if (dmax0 < len0 - 1) {         /* trailing gap in x */
            lastgap = len0 - dmax0 -1;
            lx -= lastgap;
        }
        else if (dmax0 > len0 - 1) {    /* trailing gap in y */
            lastgap = dmax0 - (len0 - 1);
            ly -= lastgap;
        }
        getmat(lx, ly, firstgap, lastgap);
        pr_align();
    }
```

Table 1 (cont')

```
/*
 * trace back the best path, count matches
 */
static
getmat(lx, ly, firstgap, lastgap)                                    getmat
        int     lx, ly;                 /* "core" (minus endgaps) */
        int     firstgap, lastgap;      /* leading trailing overlap */
{
        int             nm, i0, i1, siz0, siz1;
        char            outx[32];
        double          pct;
        register        n0, n1;
        register char   *p0, *p1;

        /* get total matches, score
         */
        i0 = i1 = siz0 = siz1 = 0;
        p0 = seqx[0] + pp[1].spc;
        p1 = seqx[1] + pp[0].spc;
        n0 = pp[1].spc + 1;
        n1 = pp[0].spc + 1;

        nm = 0;
        while ( *p0 && *p1 ) {
                if (siz0) {
                        p1++;
                        n1++;
                        siz0--;
                }
                else if (siz1) {
                        p0++;
                        n0++;
                        siz1--;
                }
                else {
                        if (xbm[*p0-'A']&xbm[*p1-'A'])
                                nm++;
                        if (n0++ == pp[0].x[i0])
                                siz0 = pp[0].n[i0++];
                        if (n1++ == pp[1].x[i1])
```

```
                    aix1 = pp[1].n[i1++];
                p0++;
                p1++;
            }
        }

        /* pct homology:
         * if penalizing endgaps, base is the shorter seq
         * else, knock off overhangs and take shorter core
         */
        if (endgaps)
            lx = (len0 < len1)? len0 : len1;
        else
            lx = (lx < ly)? lx : ly;
        pct = 100.*(double)nm/(double)lx;
        fprintf(fx, "\n");
        fprintf(fx, "<%d match%s in an overlap of %d: %.2f percent similarity\n",
            nm, (nm == 1)? "" : "es", lx, pct);
```

Table 1 (cont')

```
        fprintf(fx, "<gaps in first sequence: %d", gapx);                    --getmat
        if (gapx) {
                (void) sprintf(outx, " (%d %s%s)",
                        ngapx, (dna)? "base":"residue", (ngapx == 1)? "":"s");
                fprintf(fx,"%s", outx);

        fprintf(fx, ", gaps in second sequence: %d", gapy);
        if (gapy) {
                (void) sprintf(outx, " (%d %s%s)",
                        ngapy, (dna)? "base":"residue", (ngapy == 1)? "":"s");
                fprintf(fx,"%s", outx);
        }
        if (dna)
                fprintf(fx,
                "\n<score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
                smax, DMAT, DMIS, DINS0, DINS1);
        else
                fprintf(fx,
                "\n<score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
                smax, PINS0, PINS1);
        if (endgaps)
                fprintf(fx,
                "<endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
                firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
                lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
        else
                fprintf(fx, "<endgaps not penalized\n");
}

static          nm;             /* matches in core -- for checking */
static          lmax;           /* lengths of stripped file names */
static          ij[2];          /* jmp index for a path */
static          nc[2];          /* number at start of current line */
static          ni[2];          /* current elem number -- for gapping */
static          siz[2];
static char     *ps[2];         /* ptr to current element */
static char     *po[2];         /* ptr to next output char slot */
static char     out[2][P_LINE]; /* output line */
static char     star[P_LINE];   /* set by stars() */
```

pr_align

```
/*
 * print alignment of described in struct path pp[]
 */
static
pr_align()
{
        int             nn;     /* char count */
        int             more;
        register        i;

        for (i = 0, lmax = 0; i < 2; i++) {
                nn = stripname(namex[i]);
                if (nn > lmax)
                        lmax = nn;

                nc[i] = 1;
                nl[i] = 1;
                siz[i] = ij[i] = 0;
                pa[i] = seqx[i];
                po[i] = out[i];                 }
```

Table 1 (cont')

...pr_align

```
for (nn = nm = 0, more = 1; more; ) {
        for (i = more = 0; i < 2; i++) {
                /*
                 * do we have more of this sequence?
                 */
                if (!*ps[i])
                        continue;

                more++;

                if (pp[i].spc) {        /* leading space */
                        *po[i]++ = ' ';
                        pp[i].spc--;
                }
                else if (siz[i]) {      /* in a gap */
                        *po[i]++ = '-';
                        siz[i]--;
                }
                else {                  /* we're putting a seq element
                                         */
                        *po[i] = *ps[i];
                        if (islower(*ps[i]))
                                *ps[i] = toupper(*ps[i]);
                        po[i]++;
                        ps[i]++;

                        /*
                         * are we at next gap for this seq?
                         */
                        if (ni[i] == pp[i].x[ij[i]]) {
                                /*
                                 * we need to merge all gaps
                                 * at this location
                                 */
                                siz[i] = pp[i].n[ij[i]++];
                                while (ni[i] == pp[i].x[ij[i]])
                                        siz[i] += pp[i].n[ij[i]++];
                        }
                        ni[i]++;
```

```
                        }
                }
                if (++nn == olen[] !more && nn) {
                        dumpblock();
                        for (i = 0; i < 2; i++)
                                po[i] = out[i];
                        nn = 0;
                }
        }
}

/*
 * dump a block of lines, including numbers, stars: pr_align()
 */
static
dumpblock()
{
        register  i;

        for (i = 0; i < 2; i++)
                *po[i]-- = '\0';
```

dumpblock

Table 1 (cont')

...dumpblock

```
        (void) putc('\n', fx);
        for (i = 0; i < 2; i++) {
5               if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
                        if (i == 0)
                                nums(i);
                        if (i == 0 && *out[1])
                                stars();
10                      putline(i);
                        if (i == 0 && *out[1])
                                fprintf(fx, star);
                        if (i == 1)
                                nums(i);
15              }
        }
}

/*
20  * put out a number line: dumpblock()
 */
static
nums(ix)                                                        nums
        int     ix;     /* index in out[] holding seq line */
25  {
        char            nline[P_LINE];
        register        i, j;
        register char   *pn, *px, *py;

30      for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
                *pn = ' ';
        for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
                if (*py == ' ' || *py == '-')
                        *pn = ' ';
35              else {
                        if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                                j = (i < 0)? -i : i;
                                for (px = pn; j; j /= 10, px--)
                                        *px = j%10 + '0';
```

```
                        if (i < 0)
                                *px = '-';
                }
                else
                        *pn = ' ';
                i++;
        }
    }
    *pn = '\0';
    nc[lx] = l;
    for (pn = nline; *pn; pn++)
        (void) putc(*pn, fx);
    (void) putc('\n', fx);
}

/*
 * put out a line (name, [num], seq, [num]); dumpblock()
 */
static
putline(lx)                                                    putline
    int     ix;                     {
```

Table 1 (cont')

...putline

```
        int             i;
        register char   *px;

        for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
                (void) putc(*px, fx);
        for (; i < lmax+P_SPC; i++)
                (void) putc(' ', fx);

        /* these count from 1:
         * ni[] is current element (from 1)
         * nc[] is number at start of current line
         */
        for (px = out[ix]; *px; px++)
                (void) putc(*px&0x7F, fx);
        (void) putc('\n', fx);
}

/*
 * put a line of stars (seqs always in out[0], out[1]): dumpblock()
 */
static
stars()                                                          stars
{
        int             i;
        register char   *p0, *p1, cx, *px;

        if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
            !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
                return;
        px = star;
        for (i = lmax+P_SPC; i; i--)
                *px++ = ' ';

        for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
                if (isalpha(*p0) && isalpha(*p1)) {
```

```
                    if (xbm[*p0-'A']&xbm[*p1-'A']) {
                        cx = '*';
                        nm++;
                    }
                    else if (kbm && _day[*p0-'A'][*p1-'A'] > 0)
                        cx = '.';
                    else
                        cx = ' ';
                }
                else
                    cx = ' ';
                *px++ = cx;
            }
            *px++ = '\n';
            *px = '\0';
        }
```

Table 1 (cont')

```
/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
stripname(pn)                                                    stripname
        char    *pn;    /* file name (may be path) */
{
        register char   *px, *py;

        py = 0;
        for (px = pn; *px; px++)
                if (*px == '/')
                        py = px + 1;
        if (py)
                (void) strcpy(pn, py);
        return(strlen(pn));

}
```

Table 1 (cont')

```
/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_calloc() -- calloc() with error checkin
 * readjmps() -- get the good jmps, from tmp file if necessary
 * writejmps() -- write a filled array of jmps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>

char    *jname = "/tmp/homgXXXXXX";             /* tmp file for jmps */
FILE    *fj;

int     cleanup();                              /* cleanup tmp file */
long    lseek();

/*
 * remove any tmp file if we blow
 */
cleanup()                                                    cleanup
        int     i;
{
        if (fj)
                (void) unlink(jname);
        exit(i);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
char    *
getseq(file, len)                                            getseq
        char    *file;  /* file name */
        int     *len;   /* seq len */
{
        char            line[1024], *pseq;
        register char   *px, *py;
```

```
int             natgc, tlen;
FILE            *fp;

if ((fp = fopen(file,"r")) == 0) {
        fprintf(stderr,"%s: can't read %s\n", prog, file);
        exit(1);
}
tlen = natgc = 0;
while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
                continue;
        for (px = line; *px != '\n'; px++)
                if (isupper(*px) || islower(*px))
                        tlen++;
}
if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr,"%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
}
pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
```

Table 1 (cont')

...getseq

```
        py = pseq + 4;
        *len = tlen;
        rewind(fp);

        while (fgets(line, 1024, fp)) {
                if (*line == ';' || *line == '<' || *line == '>')
                        continue;
                for (px = line; *px != '\n'; px++) {
                        if (isupper(*px))
                                *py++ = *px;
                        else if (islower(*px))
                                *py++ = toupper(*px);
                        if (index("ATGCU",*(py-1)))
                                natgc++;
                }
        }
        *py++ = '\0';
        *py = '\0';
        (void) fclose(fp);
        dna = natgc > (tlen/3);
        return(pseq+4);
}

char    *
g_calloc(msg, nx, sz)                                                    g_calloc
        char    *msg;           /* program, calling routine */
        int     nx, sz;         /* number and size of elements */
{
        char            *px, *calloc();

        if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
                if (*msg) {
                        fprintf(stderr, "%s: g_calloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
                        exit(1);
                }
        }
        return(px);
}
```

```
/*
 * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
 */
readjmps()                                                          readjmps
{
        int             fd = -1;
        int             siz, i0, i1;
        register        i, j, xx;

        if (fj) {
                (void) fclose(fj);
                if ((fd = open(jname, O_RDONLY, 0)) < 0) {
                        fprintf(stderr, "%s: can't open() %s\n", prog, jname);
                        cleanup(1);
                }
        }
        for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {
                while (1) {
                        for (j = dx[dmax].jmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                                ;
```

Table 1 (cont')

...readjmps

```
                if (j < 0 && dx[dmax].offset && fj) {
                        (void) lseek(fd, dx[dmax].offset, 0);
                        (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
                        (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
                        dx[dmax].jmp = MAXJMP-1;
                }
                else
                        break;
        }
        if (j >= JMPS) {
                fprintf(stderr, "%s: too many gaps in alignment\n", prog);
                cleanup(1);
        }
        if (j >= 0) {
                siz = dx[dmax].jp.n[j];
                xx = dx[dmax].jp.x[j];
                dmax += siz;
                if (siz < 0) {                  /* gap in second seq */
                        pp[1].n[i1] = -siz;
                        xx += siz;
                        /* id = xx - yy + len1 - 1
                         */
                        pp[1].x[i1] = xx - dmax + len1 - 1;
                        gapy++;
                        ngapy -= siz;
/* ignore MAXGAP when doing endgaps */
                        siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
                        i1++;
                }
                else if (siz > 0) {     /* gap in first seq */
                        pp[0].n[i0] = siz;
                        pp[0].x[i0] = xx;
                        gapx++;
                        ngapx += siz;
/* ignore MAXGAP when doing endgaps */
                        siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
                        i0++;
                }
        }
```

```
            else
                    break;
    }

    /* reverse the order of jmps
     */
    for (j = 0, i0--; j < i0; j++, i0--) {
            i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
            i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
    }
    for (j = 0, i1--; j < i1; j++, i1--) {
            i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
            i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
    }
    if (fd >= 0)
            (void) close(fd);
    if (fj) {
            (void) unlink(jname);
            fj = 0;
            offset = 0;
    }                                       }
```

Table 1 (cont')

```
/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
writejmps(ix)                                                    writejmps
        int     ix;
{
        char    *mktemp();

        if (!fj) {
                if (mktemp(jname) < 0) {
                        fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
                        cleanup(1);
                }
                if ((fj = fopen(jname, "w")) == 0) {
                        fprintf(stderr, "%s: can't write %s\n", prog, jname);
                        exit(1);
                }
        }
        (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
        (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}
```

Table 2

TAT	XXXXXXXXXXXXXXX	(Length = 15 amino acids)
Comparison Protein	XXXXXYYYYYYY	(Length = 12 amino acids)

% amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the TAT polypeptide) =

5 divided by 15 = 33.3%

Table 3

TAT	XXXXXXXXXX	(Length = 10 amino acids)
Comparison Protein	XXXXXYYYYYYZZYZ	(Length = 15 amino acids)

% amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the TAT polypeptide) =

5 divided by 10 = 50%

Table 4

TAT-DNA	NNNNNNNNNNNNNN	(Length = 14 nucleotides)
Comparison DNA	NNNNNNLLLLLLLLLL	(Length = 16 nucleotides)

% nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the TAT-DNA nucleic acid sequence) =

6 divided by 14 = 42.9%

Table 5

TAT-DNA	NNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNLLLVV	(Length = 9 nucleotides)

% nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the TAT-DNA nucleic acid sequence) =

4 divided by 12 = 33.3%.

II. Compositions and Methods of the Invention

A. Anti-TAT Antibodies

In one embodiment, the present invention provides anti-TAT antibodies which may find use herein as therapeutic and/or diagnostic agents. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

1. Polyclonal Antibodies

Polyclonal antibodies are preferably raised in animals by multiple subcutaneous (sc) or intraperitoneal (ip) injections of the relevant antigen and an adjuvant. It may be useful to conjugate the relevant antigen (especially when synthetic peptides are used) to a protein that is immunogenic in the species to be immunized. For example, the antigen can be conjugated to keyhole limpet hemocyanin (KLH), serum albumin, bovine thyroglobulin, or soybean trypsin inhibitor, using a bifunctional or derivatizing agent, e.g., maleimidobenzoyl sulfosuccinimide ester (conjugation through cysteine residues), N-hydroxysuccinimide (through lysine residues), glutaraldehyde, succinic anhydride, $SOCl_2$, or $R^1N=C=NR$, where R and R^1 are different alkyl groups.

Animals are immunized against the antigen, immunogenic conjugates, or derivatives by combining, e.g., 100 µg or 5 µg of the protein or conjugate (for rabbits or mice, respectively) with 3 volumes of Freund's complete adjuvant and injecting the solution intradermally at multiple sites. One month later, the animals are boosted with 1/5 to 1/10 the original amount of peptide or conjugate in Freund's complete adjuvant by subcutaneous injection at multiple sites. Seven to 14 days later, the animals are bled and the serum is assayed for antibody titer. Animals are boosted until the titer plateaus. Conjugates also can be made in recombinant cell culture as protein fusions. Also, aggregating agents such as alum are suitably used to enhance the immune response.

2. Monoclonal Antibodies

Monoclonal antibodies may be made using the hybridoma method first described by Kohler et al., Nature, 256:495 (1975), or may be made by recombinant DNA methods (U.S. Patent No. 4,816,567).

In the hybridoma method, a mouse or other appropriate host animal, such as a hamster, is immunized as described above to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the protein used for immunization. Alternatively, lymphocytes may be immunized *in vitro*. After immunization, lymphocytes are isolated and then fused with a myeloma cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, pp.59-103 (Academic Press, 1986)).

The hybridoma cells thus prepared are seeded and grown in a suitable culture medium which medium preferably contains one or more substances that inhibit the growth or survival of the unfused, parental myeloma cells (also referred to as fusion partner). For example, if the parental myeloma cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the selective culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine (HAT medium), which substances prevent the growth of HGPRT-deficient cells.

Preferred fusion partner myeloma cells are those that fuse efficiently, support stable high-level production of antibody by the selected antibody-producing cells, and are sensitive to a selective medium that selects against the unfused parental cells. Preferred myeloma cell lines are murine myeloma lines, such as those derived from MOPC-21 and MPC-11 mouse tumors available from the Salk Institute Cell Distribution Center, San Diego, California USA, and SP-2 and derivatives e.g., X63-Ag8-653 cells available from the American Type Culture Collection, Manassas, Virginia, USA. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); and Brodeur et al., Monoclonal Antibody Production Techniques and Applications, pp. 51-63 (Marcel Dekker, Inc., New York, 1987)).

Culture medium in which hybridoma cells are growing is assayed for production of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunosorbent assay (ELISA).

The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis described in Munson et al., Anal. Biochem., 107:220 (1980).

Once hybridoma cells that produce antibodies of the desired specificity, affinity, and/or activity are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods (Goding, Monoclonal Antibodies: Principles and Practice, pp.59-103 (Academic Press, 1986)). Suitable culture media for this purpose include, for example, D-MEM or RPMI-1640 medium. In addition, the hybridoma cells may be grown *in vivo* as ascites tumors in an animal e.g., by i.p. injection of the cells into mice.

The monoclonal antibodies secreted by the subclones are suitably separated from the culture medium, ascites fluid, or serum by conventional antibody purification procedures such as, for example, affinity chromatography (e.g., using protein A or protein G-Sepharose) or ion-exchange chromatography, hydroxylapatite chromatography, gel electrophoresis, dialysis, etc.

DNA encoding the monoclonal antibodies is readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as *E. coli* cells, simian COS cells, Chinese Hamster Ovary (CHO) cells, or myeloma cells that do not otherwise produce antibody protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. Review articles on recombinant expression in bacteria of DNA encoding the antibody include Skerra et al., Curr. Opinion in Immunol., 5:256-262 (1993) and Plückthun, Immunol. Revs. 130:151-188 (1992).

In a further embodiment, monoclonal antibodies or antibody fragments can be isolated from antibody phage libraries generated using the techniques described in McCafferty et al., Nature, 348:552-554 (1990). Clackson et al., Nature, 352:624-628 (1991) and Marks et al., J. Mol. Biol., 222:581-597 (1991) describe the isolation of murine and human antibodies, respectively, using phage libraries. Subsequent publications describe the production of high affinity (nM range) human antibodies by chain shuffling (Marks et al., Bio/Technology, 10:779-783 (1992)), as well as combinatorial infection and *in vivo* recombination as a strategy for constructing very large phage libraries (Waterhouse et al., Nuc. Acids. Res. 21:2265-2266 (1993)). Thus, these techniques are viable alternatives to traditional monoclonal antibody hybridoma techniques for isolation of monoclonal antibodies.

The DNA that encodes the antibody may be modified to produce chimeric or fusion antibody polypeptides, for example, by substituting human heavy chain and light chain constant domain (C_H and C_L) sequences for the homologous murine sequences (U.S. Patent No. 4,816,567; and Morrison, et al., Proc. Natl Acad. Sci. USA, 81:6851 (1984)), or by fusing the immunoglobulin coding sequence with all or part of the coding sequence for a non-immunoglobulin polypeptide (heterologous polypeptide). The non-immunoglobulin polypeptide sequences can substitute for the constant domains of an antibody, or they are substituted for the variable domains of one antigen-combining site of an antibody to create a chimeric bivalent antibody comprising one antigen-combining site having specificity for an antigen and another antigen-combining site having specificity for a different antigen.

3. Human and Humanized Antibodies

The anti-TAT antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized

antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-329 (1988); and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

The choice of human variable domains, both light and heavy, to be used in making the humanized antibodies is very important to reduce antigenicity and HAMA response (human anti-mouse antibody) when the antibody is intended for human therapeutic use. According to the so-called "best-fit" method, the sequence of the variable domain of a rodent antibody is screened against the entire library of known human variable domain sequences. The human V domain sequence which is closest to that of the rodent is identified and the human framework region (FR) within it accepted for the humanized antibody (Sims et al., J. Immunol. 151:2296 (1993); Chothia et al., J. Mol. Biol., 196:901 (1987)). Another method uses a particular framework region derived from the consensus sequence of all human antibodies of a particular subgroup of light or heavy chains. The same framework may be used for several different humanized antibodies (Carter et al., Proc. Natl. Acad. Sci. USA, 89:4285 (1992); Presta et al., J. Immunol. 151:2623 (1993)).

It is further important that antibodies be humanized with retention of high binding affinity for the antigen and other favorable biological properties. To achieve this goal, according to a preferred method, humanized antibodies are prepared by a process of analysis of the parental sequences and various conceptual humanized products using three-dimensional models of the parental and humanized sequences. Three-dimensional immunoglobulin models are commonly available and are familiar to those skilled in the art. Computer programs are available which illustrate and display probable three-dimensional conformational structures of selected candidate immunoglobulin sequences. Inspection of these displays permits analysis of the likely role of the residues in the functioning of the candidate immunoglobulin sequence, i.e., the analysis of residues that influence the ability of the candidate immunoglobulin to bind its antigen. In this way, FR residues can be selected and combined from the recipient and import sequences so that the desired antibody

characteristic, such as increased affinity for the target antigen(s), is achieved. In general, the hypervariable region residues are directly and most substantially involved in influencing antigen binding.

Various forms of a humanized anti-TAT antibody are contemplated. For example, the humanized antibody may be an antibody fragment, such as a Fab, which is optionally conjugated with one or more cytotoxic agent(s) in order to generate an immunoconjugate. Alternatively, the humanized antibody may be an intact antibody, such as an intact IgG1 antibody.

As an alternative to humanization, human antibodies can be generated. For example, it is now possible to produce transgenic animals (e.g., mice) that are capable, upon immunization, of producing a full repertoire of human antibodies in the absence of endogenous immunoglobulin production. For example, it has been described that the homozygous deletion of the antibody heavy-chain joining region (J_H) gene in chimeric and germ-line mutant mice results in complete inhibition of endogenous antibody production. Transfer of the human germ-line immunoglobulin gene array into such germ-line mutant mice will result in the production of human antibodies upon antigen challenge. See, e.g., Jakobovits et al., Proc. Natl. Acad. Sci. USA, 90:2551 (1993); Jakobovits et al., Nature, 362:255-258 (1993); Bruggemann et al., Year in Immuno. 7:33 (1993); U.S. Patent Nos. 5,545,806, 5,569,825, 5,591,669 (all of GenPharm); 5,545,807; and WO 97/17852.

Alternatively, phage display technology (McCafferty et al., Nature 348:552-553 [1990]) can be used to produce human antibodies and antibody fragments *in vitro*, from immunoglobulin variable (V) domain gene repertoires from unimmunized donors. According to this technique, antibody V domain genes are cloned in-frame into either a major or minor coat protein gene of a filamentous bacteriophage, such as M13 or fd, and displayed as functional antibody fragments on the surface of the phage particle. Because the filamentous particle contains a single-stranded DNA copy of the phage genome, selections based on the functional properties of the antibody also result in selection of the gene encoding the antibody exhibiting those properties. Thus, the phage mimics some of the properties of the B-cell. Phage display can be performed in a variety of formats, reviewed in, e.g., Johnson, Kevin S. and Chiswell, David J., Current Opinion in Structural Biology 3:564-571 (1993). Several sources of V-gene segments can be used for phage display. Clackson et al., Nature, 352:624-628 (1991) isolated a diverse array of anti-oxazolone antibodies from a small random combinatorial library of V genes derived from the spleens of immunized mice. A repertoire of V genes from unimmunized human donors can be constructed and antibodies to a diverse array of antigens (including self-antigens) can be isolated essentially following the techniques described by Marks et al., J. Mol. Biol. 222:581-597 (1991), or Griffith et al., EMBO J. 12:725-734 (1993). See, also, U.S. Patent Nos. 5,565,332 and 5,573,905.

As discussed above, human antibodies may also be generated by *in vitro* activated B cells (see U.S. Patents 5,567,610 and 5,229,275).

4. Antibody fragments

In certain circumstances there are advantages of using antibody fragments, rather than whole antibodies. The smaller size of the fragments allows for rapid clearance, and may lead to improved access to solid tumors.

Various techniques have been developed for the production of antibody fragments. Traditionally, these fragments were derived via proteolytic digestion of intact antibodies (see, e.g., Morimoto et al., Journal of Biochemical and Biophysical Methods 24:107-117 (1992); and Brennan et al., Science, 229:81 (1985)). However, these fragments can now be produced directly by recombinant host cells. Fab, Fv and ScFv antibody fragments can all be expressed in and secreted from *E. coli*, thus allowing the facile production of large amounts of these fragments. Antibody fragments can be isolated from the antibody phage libraries discussed above. Alternatively, Fab'-SH fragments can be directly recovered from *E. coli* and chemically coupled to form $F(ab')_2$ fragments (Carter et al., Bio/Technology 10:163-167 (1992)). According to another approach, $F(ab')_2$ fragments can be isolated directly from recombinant host cell culture. Fab and $F(ab')_2$ fragment with increased in vivo half-life comprising a salvage receptor binding epitope residues are described in U.S. Patent No. 5,869,046. Other techniques for the production of antibody fragments will be apparent to the skilled practitioner. In other embodiments, the antibody of choice is a single chain Fv fragment (scFv). See WO 93/16185; U.S. Patent No. 5,571,894; and U.S. Patent No. 5,587,458. Fv and sFv are the only species with intact combining sites that are devoid of constant regions; thus, they are suitable for reduced nonspecific binding during in vivo use. sFv fusion proteins may be constructed to yield fusion of an effector protein at either the amino or the carboxy terminus of an sFv. See Antibody Engineering, ed. Borrebaeck, supra. The antibody fragment may also be a "linear antibody", e.g., as described in U.S. Patent 5,641,870 for example. Such linear antibody fragments may be monospecific or bispecific.

5. Bispecific Antibodies

Bispecific antibodies are antibodies that have binding specificities for at least two different epitopes. Exemplary bispecific antibodies may bind to two different epitopes of a TAT protein as described herein. Other such antibodies may combine a TAT binding site with a binding site for another protein. Alternatively, an anti-TAT arm may be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD3), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16), so as to focus and localize cellular defense mechanisms to the TAT-expressing cell. Bispecific antibodies may also be used to localize cytotoxic agents to cells which express TAT. These antibodies possess a TAT-binding arm and an arm which binds the cytotoxic agent (e.g., saporin, anti-interferon-α, vinca alkaloid, ricin A chain, methotrexate or radioactive isotope hapten). Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g., $F(ab')_2$ bispecific antibodies).

WO 96/16673 describes a bispecific anti-ErbB2/anti-FcγRIII antibody and U.S. Patent No. 5,837,234 discloses a bispecific anti-ErbB2/anti-FcγRI antibody. A bispecific anti-ErbB2/Fcα antibody is shown in WO98/02463. U.S. Patent No. 5,821,337 teaches a bispecific anti-ErbB2/anti-CD3 antibody.

Methods for making bispecific antibodies are known in the art. Traditional production of full length bispecific antibodies is based on the co-expression of two immunoglobulin heavy chain-light chain pairs, where the two chains have different specificities (Millstein et al., Nature 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of 10 different antibody molecules, of which only one has the correct bispecific structure.

Purification of the correct molecule, which is usually done by affinity chromatography steps, is rather cumbersome, and the product yields are low. Similar procedures are disclosed in WO 93/08829, and in Traunecker et al., EMBO J. 10:3655-3659 (1991).

According to a different approach, antibody variable domains with the desired binding specificities (antibody-antigen combining sites) are fused to immunoglobulin constant domain sequences. Preferably, the fusion is with an Ig heavy chain constant domain, comprising at least part of the hinge, C_H2, and C_H3 regions. It is preferred to have the first heavy-chain constant region (C_H1) containing the site necessary for light chain bonding, present in at least one of the fusions. DNAs encoding the immunoglobulin heavy chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host cell. This provides for greater flexibility in adjusting the mutual proportions of the three polypeptide fragments in embodiments when unequal ratios of the three polypeptide chains used in the construction provide the optimum yield of the desired bispecific antibody. It is, however, possible to insert the coding sequences for two or all three polypeptide chains into a single expression vector when the expression of at least two polypeptide chains in equal ratios results in high yields or when the ratios have no significant affect on the yield of the desired chain combination.

In a preferred embodiment of this approach, the bispecific antibodies are composed of a hybrid immunoglobulin heavy chain with a first binding specificity in one arm, and a hybrid immunoglobulin heavy chain-light chain pair (providing a second binding specificity) in the other arm. It was found that this asymmetric structure facilitates the separation of the desired bispecific compound from unwanted immunoglobulin chain combinations, as the presence of an immunoglobulin light chain in only one half of the bispecific molecule provides for a facile way of separation. This approach is disclosed in WO 94/04690. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology 121:210 (1986).

According to another approach described in U.S. Patent No. 5,731,168, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the C_H3 domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g., tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g., alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies include cross-linked or "heteroconjugate" antibodies. For example, one of the antibodies in the heteroconjugate can be coupled to avidin, the other to biotin. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360, WO 92/200373, and EP 03089). Heteroconjugate antibodies may be made using any convenient cross-linking methods. Suitable cross-linking agents are well known in the art, and are disclosed in U.S. Patent No. 4,676,980, along with a number of cross-linking techniques.

Techniques for generating bispecific antibodies from antibody fragments have also been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate $F(ab')_2$ fragments. These fragments are reduced in the presence of the dithiol complexing agent, sodium arsenite, to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Recent progress has facilitated the direct recovery of Fab'-SH fragments from *E. coli*, which can be chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175: 217-225 (1992) describe the production of a fully humanized bispecific antibody $F(ab')_2$ molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling *in vitro* to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a V_H connected to a V_L by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., J. Immunol. 147:60 (1991).

6. Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a

thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

7. Multivalent Antibodies

A multivalent antibody may be internalized (and/or catabolized) faster than a bivalent antibody by a cell expressing an antigen to which the antibodies bind. The antibodies of the present invention can be multivalent antibodies (which are other than of the IgM class) with three or more antigen binding sites (e.g. tetravalent antibodies), which can be readily produced by recombinant expression of nucleic acid encoding the polypeptide chains of the antibody. The multivalent antibody can comprise a dimerization domain and three or more antigen binding sites. The preferred dimerization domain comprises (or consists of) an Fc region or a hinge region. In this scenario, the antibody will comprise an Fc region and three or more antigen binding sites amino-terminal to the Fc region. The preferred multivalent antibody herein comprises (or consists of) three to about eight, but preferably four, antigen binding sites. The multivalent antibody comprises at least one polypeptide chain (and preferably two polypeptide chains), wherein the polypeptide chain(s) comprise two or more variable domains. For instance, the polypeptide chain(s) may comprise VD1-$(X1)_n$-VD2-$(X2)_n$-Fc, wherein VD1 is a first variable domain, VD2 is a second variable domain, Fc is one polypeptide chain of an Fc region, X1 and X2 represent an amino acid or polypeptide, and n is 0 or 1. For instance, the polypeptide chain(s) may comprise: VH-CH1-flexible linker-VH-CH1-Fc region chain; or VH-CH1-VH-CH1-Fc region chain. The multivalent antibody herein preferably further comprises at least two (and preferably four) light chain variable domain polypeptides. The multivalent antibody herein may, for instance, comprise from about two to about eight light chain variable domain polypeptides. The light chain variable domain polypeptides contemplated here comprise a light chain variable domain and, optionally, further comprise a CL domain.

8. Effector Function Engineering

It may be desirable to modify the antibody of the invention with respect to effector function, e.g., so as to enhance antigen-dependent cell-mediated cytotoxicity (ADCC) and/or complement dependent cytotoxicity (CDC) of the antibody. This may be achieved by introducing one or more amino acid substitutions in an Fc region of the antibody. Alternatively or additionally, cysteine residue(s) may be introduced in the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med. 176:1191-1195 (1992) and Shopes, B. J. Immunol. 148:2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff et al., Cancer Research 53:2560-2565 (1993). Alternatively, an antibody can be engineered which has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design 3:219-230 (1989). To increase the serum half life of the antibody, one may incorporate a salvage receptor binding epitope into the antibody (especially an antibody fragment) as described in U.S. Patent 5,739,277, for example. As used herein, the term "salvage receptor binding epitope" refers to an epitope of the Fc region of an IgG

molecule (e.g., IgG_1, IgG_2, IgG_3, or IgG_4) that is responsible for increasing the *in vivo* serum half-life of the IgG molecule.

9. Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, a growth inhibitory agent, a toxin (*e.g.*, an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (*i.e.*, a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi, ^{131}I, ^{131}In, ^{90}Y, and ^{186}Re. Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al.*, Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

Conjugates of an antibody and one or more small molecule toxins, such as a calicheamicin, maytansinoids, a trichothene, and CC1065, and the derivatives of these toxins that have toxin activity, are also contemplated herein.

Maytansine and maytansinoids

In one preferred embodiment, an anti-TAT antibody (full length or fragments) of the invention is conjugated to one or more maytansinoid molecules.

Maytansinoids are mitototic inhibitors which act by inhibiting tubulin polymerization. Maytansine was first isolated from the east African shrub *Maytenus serrata* (U.S. Patent No. 3,896,111). Subsequently, it was discovered that certain microbes also produce maytansinoids, such as maytansinol and C-3 maytansinol esters (U.S. Patent No. 4,151,042). Synthetic maytansinol and derivatives and analogues thereof are disclosed, for example, in U.S. Patent Nos. 4,137,230; 4,248,870; 4,256,746; 4,260,608; 4,265,814; 4,294,757; 4,307,016; 4,308,268; 4,308,269; 4,309,428; 4,313,946; 4,315,929; 4,317,821; 4,322,348; 4,331,598;

4,361,650; 4,364,866; 4,424,219; 4,450,254; 4,362,663; and 4,371,533, the disclosures of which are hereby expressly incorporated by reference.

Maytansinoid-antibody conjugates

In an attempt to improve their therapeutic index, maytansine and maytansinoids have been conjugated to antibodies specifically binding to tumor cell antigens. Immunoconjugates containing maytansinoids and their therapeutic use are disclosed, for example, in U.S. Patent Nos. 5,208,020, 5,416,064 and European Patent EP 0 425 235 B1, the disclosures of which are hereby expressly incorporated by reference. Liu et al., Proc. Natl. Acad. Sci. USA 93:8618-8623 (1996) described immunoconjugates comprising a maytansinoid designated DM1 linked to the monoclonal antibody C242 directed against human colorectal cancer. The conjugate was found to be highly cytotoxic towards cultured colon cancer cells, and showed antitumor activity in an *in vivo* tumor growth assay. Chari et al., Cancer Research 52:127-131 (1992) describe immunoconjugates in which a maytansinoid was conjugated via a disulfide linker to the murine antibody A7 binding to an antigen on human colon cancer cell lines, or to another murine monoclonal antibody TA.1 that binds the HER-2/*neu* oncogene. The cytotoxicity of the TA.1-maytansonoid conjugate was tested *in vitro* on the human breast cancer cell line SK-BR-3, which expresses 3×10^5 HER-2 surface antigens per cell. The drug conjugate achieved a degree of cytotoxicity similar to the free maytansonid drug, which could be increased by increasing the number of maytansinoid molecules per antibody molecule. The A7-maytansinoid conjugate showed low systemic cytotoxicity in mice.

Anti-TAT polypeptide antibody-maytansinoid conjugates (immunoconjugates)

Anti-TAT antibody-maytansinoid conjugates are prepared by chemically linking an anti-TAT antibody to a maytansinoid molecule without significantly diminishing the biological activity of either the antibody or the maytansinoid molecule. An average of 3-4 maytansinoid molecules conjugated per antibody molecule has shown efficacy in enhancing cytotoxicity of target cells without negatively affecting the function or solubility of the antibody, although even one molecule of toxin/antibody would be expected to enhance cytotoxicity over the use of naked antibody. Maytansinoids are well known in the art and can be synthesized by known techniques or isolated from natural sources. Suitable maytansinoids are disclosed, for example, in U.S. Patent No. 5,208,020 and in the other patents and nonpatent publications referred to hereinabove. Preferred maytansinoids are maytansinol and maytansinol analogues modified in the aromatic ring or at other positions of the maytansinol molecule, such as various maytansinol esters.

There are many linking groups known in the art for making antibody-maytansinoid conjugates, including, for example, those disclosed in U.S. Patent No. 5,208,020 or EP Patent 0 425 235 B1, and Chari et al., Cancer Research 52:127-131 (1992). The linking groups include disulfide groups, thioether groups, acid labile groups, photolabile groups, peptidase labile groups, or esterase labile groups, as disclosed in the above-identified patents, disulfide and thioether groups being preferred.

Conjugates of the antibody and maytansinoid may be made using a variety of bifunctional protein coupling agents such as N-succinimidyl-3-(2-pyridyldithio) propionate (SPDP), succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate, iminothiolane (IT), bifunctional derivatives of imidoesters (such

as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as toluene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). Particularly preferred coupling agents include N-succinimidyl-3-(2-pyridyldithio) propionate (SPDP) (Carlsson et al., Biochem. J. 173:723-737 [1978]) and N-succinimidyl-4-(2-pyridylthio)pentanoate (SPP) to provide for a disulfide linkage.

The linker may be attached to the maytansinoid molecule at various positions, depending on the type of the link. For example, an ester linkage may be formed by reaction with a hydroxyl group using conventional coupling techniques. The reaction may occur at the C-3 position having a hydroxyl group, the C-14 position modified with hyrdoxymethyl, the C-15 position modified with a hydroxyl group, and the C-20 position having a hydroxyl group. In a preferred embodiment, the linkage is formed at the C-3 position of maytansinol or a maytansinol analogue.

Calicheamicin

Another immunoconjugate of interest comprises an anti-TAT antibody conjugated to one or more calicheamicin molecules. The calicheamicin family of antibiotics are capable of producing double-stranded DNA breaks at sub-picomolar concentrations. For the preparation of conjugates of the calicheamicin family, see U.S. patents 5,712,374, 5,714,586, 5,739,116, 5,767,285, 5,770,701, 5,770,710, 5,773,001, 5,877,296 (all to American Cyanamid Company). Structural analogues of calicheamicin which may be used include, but are not limited to, γ_1^I, α_2^I, α_3^I, N-acetyl-γ_1^I, PSAG and θ^I_1 (Hinman et al., Cancer Research 53:3336-3342 (1993), Lode et al., Cancer Research 58:2925-2928 (1998) and the aforementioned U.S. patents to American Cyanamid). Another anti-tumor drug that the antibody can be conjugated is QFA which is an antifolate. Both calicheamicin and QFA have intracellular sites of action and do not readily cross the plasma membrane. Therefore, cellular uptake of these agents through antibody mediated internalization greatly enhances their cytotoxic effects.

Other cytotoxic agents

Other antitumor agents that can be conjugated to the anti-TAT antibodies of the invention include BCNU, streptozoicin, vincristine and 5-fluorouracil, the family of agents known collectively LL-E33288 complex described in U.S. patents 5,053,394, 5,770,710, as well as esperamicins (U.S. patent 5,877,296).

Enzymatically active toxins and fragments thereof which can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin and the tricothecenes. See, for example, WO 93/21232 published October 28, 1993.

The present invention further contemplates an immunoconjugate formed between an antibody and a compound with nucleolytic activity (e.g., a ribonuclease or a DNA endonuclease such as a deoxyribonuclease; DNase).

For selective destruction of the tumor, the antibody may comprise a highly radioactive atom. A variety of radioactive isotopes are available for the production of radioconjugated anti-TAT antibodies. Examples include At^{211}, I^{131}, I^{125}, Y^{90}, Re^{186}, Re^{188}, Sm^{153}, Bi^{212}, P^{32}, Pb^{212} and radioactive isotopes of Lu. When the conjugate is used for diagnosis, it may comprise a radioactive atom for scintigraphic studies, for example tc^{99m} or I^{123}, or a spin label for nuclear magnetic resonance (NMR) imaging (also known as magnetic resonance imaging, mri), such as iodine-123 again, iodine-131, indium-111, fluorine-19, carbon-13, nitrogen-15, oxygen-17, gadolinium, manganese or iron.

The radio- or other labels may be incorporated in the conjugate in known ways. For example, the peptide may be biosynthesized or may be synthesized by chemical amino acid synthesis using suitable amino acid precursors involving, for example, fluorine-19 in place of hydrogen. Labels such as tc^{99m} or I^{123}, Re^{186}, Re^{188} and In^{111} can be attached via a cysteine residue in the peptide. Yttrium-90 can be attached via a lysine residue. The IODOGEN method (Fraker et al (1978) Biochem. Biophys. Res. Commun. 80: 49-57 can be used to incorporate iodine-123. "Monoclonal Antibodies in Immunoscintigraphy" (Chatal,CRC Press 1989) describes other methods in detail.

Conjugates of the antibody and cytotoxic agent may be made using a variety of bifunctional protein coupling agents such as N-succinimidyl-3-(2-pyridyldithio) propionate (SPDP), succinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxylate, iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science 238:1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026. The linker may be a "cleavable linker" facilitating release of the cytotoxic drug in the cell. For example, an acid-labile linker, peptidase-sensitive linker, photolabile linker, dimethyl linker or disulfide-containing linker (Chari et al., Cancer Research 52:127-131 (1992); U.S. Patent No. 5,208,020) may be used.

Alternatively, a fusion protein comprising the anti-TAT antibody and cytotoxic agent may be made, e.g., by recombinant techniques or peptide synthesis. The length of DNA may comprise respective regions encoding the two portions of the conjugate either adjacent one another or separated by a region encoding a linker peptide which does not destroy the desired properties of the conjugate.

In yet another embodiment, the antibody may be conjugated to a "receptor" (such streptavidin) for utilization in tumor pre-targeting wherein the antibody-receptor conjugate is administered to the patient,

followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) which is conjugated to a cytotoxic agent (e.g., a radionucleotide).

10. Immunoliposomes

The anti-TAT antibodies disclosed herein may also be formulated as immunoliposomes. A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein et al., Proc. Natl. Acad. Sci. USA 82:3688 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:4030 (1980); U.S. Pat. Nos. 4,485,045 and 4,544,545; and WO97/38731 published October 23, 1997. Liposomes with enhanced circulation time are disclosed in U.S. Patent No. 5,013,556.

Particularly useful liposomes can be generated by the reverse phase evaporation method with a lipid composition comprising phosphatidylcholine, cholesterol and PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of the antibody of the present invention can be conjugated to the liposomes as described in Martin et al., J. Biol. Chem. 257:286-288 (1982) via a disulfide interchange reaction. A chemotherapeutic agent is optionally contained within the liposome. See Gabizon et al., J. National Cancer Inst. 81(19):1484 (1989).

B. TAT Binding Oligopeptides

TAT binding oligopeptides of the present invention are oligopeptides that bind, preferably specifically, to a TAT polypeptide as described herein. TAT binding oligopeptides may be chemically synthesized using known oligopeptide synthesis methodology or may be prepared and purified using recombinant technology. TAT binding oligopeptides are usually at least about 5 amino acids in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 amino acids in length or more, wherein such oligopeptides that are capable of binding, preferably specifically, to a TAT polypeptide as described herein. TAT binding oligopeptides may be identified without undue experimentation using well known techniques. In this regard, it is noted that techniques for screening oligopeptide libraries for oligopeptides that are capable of specifically binding to a polypeptide target are well known in the art (see, e.g., U.S. Patent Nos. 5,556,762, 5,750,373, 4,708,871, 4,833,092, 5,223,409, 5,403,484, 5,571,689, 5,663,143; PCT Publication Nos. WO 84/03506 and WO84/03564; Geysen et al., Proc. Natl. Acad. Sci. U.S.A., 81:3998-4002 (1984); Geysen et al., Proc. Natl. Acad. Sci. U.S.A., 82:178-182 (1985); Geysen et al., in Synthetic Peptides as Antigens, 130-149 (1986); Geysen et al., J. Immunol. Meth., 102:259-274 (1987); Schoofs et al., J. Immunol., 140:611-616 (1988), Cwirla, S. E. et al. (1990) Proc. Natl. Acad. Sci. USA, 87:6378; Lowman, H.B. et al. (1991) Biochemistry, 30:10832; Clackson, T. et al. (1991) Nature, 352: 624; Marks, J. D. et al. (1991), J. Mol. Biol.,

222:581; Kang, A.S. et al. (1991) Proc. Natl. Acad. Sci. USA, 88:8363, and Smith, G. P. (1991) Current Opin. Biotechnol., 2:668).

In this regard, bacteriophage (phage) display is one well known technique which allows one to screen large oligopeptide libraries to identify member(s) of those libraries which are capable of specifically binding to a polypeptide target. Phage display is a technique by which variant polypeptides are displayed as fusion proteins to the coat protein on the surface of bacteriophage particles (Scott, J.K. and Smith, G. P. (1990) Science 249: 386). The utility of phage display lies in the fact that large libraries of selectively randomized protein variants (or randomly cloned cDNAs) can be rapidly and efficiently sorted for those sequences that bind to a target molecule with high affinity. Display of peptide (Cwirla, S. E. et al. (1990) Proc. Natl. Acad. Sci. USA, 87:6378) or protein (Lowman, H.B. et al. (1991) Biochemistry, 30:10832; Clackson, T. et al. (1991) Nature, 352: 624; Marks, J. D. et al. (1991), J. Mol. Biol., 222:581; Kang, A.S. et al. (1991) Proc. Natl. Acad. Sci. USA, 88:8363) libraries on phage have been used for screening millions of polypeptides or oligopeptides for ones with specific binding properties (Smith, G. P. (1991) Current Opin. Biotechnol., 2:668). Sorting phage libraries of random mutants requires a strategy for constructing and propagating a large number of variants, a procedure for affinity purification using the target receptor, and a means of evaluating the results of binding enrichments. U.S. Patent Nos. 5,223,409, 5,403,484, 5,571,689, and 5,663,143.

Although most phage display methods have used filamentous phage, lambdoid phage display systems (WO 95/34683; U.S. 5,627,024), T4 phage display systems (Ren, Z-J. et al. (1998) Gene 215:439; Zhu, Z. (1997) CAN 33:534; Jiang, J. et al. (1997) can 128:44380; Ren, Z-J. et al. (1997) CAN 127:215644; Ren, Z-J. (1996) Protein Sci. 5:1833; Efimov, V. P. et al. (1995) Virus Genes 10:173) and T7 phage display systems (Smith, G. P. and Scott, J.K. (1993) Methods in Enzymology, 217, 228–257; U.S. 5,766,905) are also known.

Many other improvements and variations of the basic phage display concept have now been developed. These improvements enhance the ability of display systems to screen peptide libraries for binding to selected target molecules and to display functional proteins with the potential of screening these proteins for desired properties. Combinatorial reaction devices for phage display reactions have been developed (WO 98/14277) and phage display libraries have been used to analyze and control bimolecular interactions (WO 98/20169; WO 98/20159) and properties of constrained helical peptides (WO 98/20036). WO 97/35196 describes a method of isolating an affinity ligand in which a phage display library is contacted with one solution in which the ligand will bind to a target molecule and a second solution in which the affinity ligand will not bind to the target molecule, to selectively isolate binding ligands. WO 97/46251 describes a method of biopanning a random phage display library with an affinity purified antibody and then isolating binding phage, followed by a micropanning process using microplate wells to isolate high affinity binding phage. The use of *Staphylococcus aureus* protein A as an affinity tag has also been reported (Li et al. (1998) Mol Biotech., 9:187). WO 97/47314 describes the use of substrate subtraction libraries to distinguish enzyme specificities using a combinatorial library which may be a phage display library. A method for selecting enzymes suitable for use in detergents using phage display is described in WO 97/09446. Additional methods of selecting specific binding proteins are described in U.S. Patent Nos. 5,498,538, 5,432,018, and WO 98/15833.

Methods of generating peptide libraries and screening these libraries are also disclosed in U.S. Patent Nos. 5,723,286, 5,432,018, 5,580,717, 5,427,908, 5,498,530, 5,770,434, 5,734,018, 5,698,426, 5,763,192, and 5,723,323.

C. TAT Binding Organic Molecules

TAT binding organic molecules are organic molecules other than oligopeptides or antibodies as defined herein that bind, preferably specifically, to a TAT polypeptide as described herein. TAT binding organic molecules may be identified and chemically synthesized using known methodology (see, e.g., PCT Publication Nos. WO00/00823 and WO00/39585). TAT binding organic molecules are usually less than about 2000 daltons in size, alternatively less than about 1500, 750, 500, 250 or 200 daltons in size, wherein such organic molecules that are capable of binding, preferably specifically, to a TAT polypeptide as described herein may be identified without undue experimentation using well known techniques. In this regard, it is noted that techniques for screening organic molecule libraries for molecules that are capable of binding to a polypeptide target are well known in the art (see, e.g., PCT Publication Nos. WO00/00823 and WO00/39585). TAT binding organic molecules may be, for example, aldehydes, ketones, oximes, hydrazones, semicarbazones, carbazides, primary amines, secondary amines, tertiary amines, N-substituted hydrazines, hydrazides, alcohols, ethers, thiols, thioethers, disulfides, carboxylic acids, esters, amides, ureas, carbamates, carbonates, ketals, thioketals, acetals, thioacetals, aryl halides, aryl sulfonates, alkyl halides, alkyl sulfonates, aromatic compounds, heterocyclic compounds, anilines, alkenes, alkynes, diols, amino alcohols, oxazolidines, oxazolines, thiazolidines, thiazolines, enamines, sulfonamides, epoxides, aziridines, isocyanates, sulfonyl chlorides, diazo compounds, acid chlorides, or the like.

D. Screening for Anti-TAT Antibodies, TAT Binding Oligopeptides and TAT Binding Organic Molecules With the Desired Properties

Techniques for generating antibodies, oligopeptides and organic molecules that bind to TAT polypeptides have been described above. One may further select antibodies, oligopeptides or other organic molecules with certain biological characteristics, as desired.

The growth inhibitory effects of an anti-TAT antibody, oligopeptide or other organic molecule of the invention may be assessed by methods known in the art, e.g., using cells which express a TAT polypeptide either endogenously or following transfection with the TAT gene. For example, appropriate tumor cell lines and TAT-transfected cells may treated with an anti-TAT monoclonal antibody, oligopeptide or other organic molecule of the invention at various concentrations for a few days (e.g., 2-7) days and stained with crystal violet or MTT or analyzed by some other colorimetric assay. Another method of measuring proliferation would be by comparing ^{3}H-thymidine uptake by the cells treated in the presence or absence an anti-TAT antibody, TAT binding oligopeptide or TAT binding organic molecule of the invention. After treatment, the cells are harvested and the amount of radioactivity incorporated into the DNA quantitated in a scintillation counter. Appropriate positive controls include treatment of a selected cell line with a growth inhibitory antibody known to inhibit growth of that cell line. Growth inhibition of tumor cells *in vivo* can be determined in various ways known in the art. Preferably, the tumor cell is one that overexpresses a TAT polypeptide.

Preferably, the anti-TAT antibody, TAT binding oligopeptide or TAT binding organic molecule will inhibit cell proliferation of a TAT-expressing tumor cell *in vitro* or *in vivo* by about 25-100% compared to the untreated tumor cell, more preferably, by about 30-100%, and even more preferably by about 50-100% or 70-100%, in one embodiment, at an antibody concentration of about 0.5 to 30 μg/ml. Growth inhibition can be measured at an antibody concentration of about 0.5 to 30 μg/ml or about 0.5 nM to 200 nM in cell culture, where the growth inhibition is determined 1-10 days after exposure of the tumor cells to the antibody. The antibody is growth inhibitory *in vivo* if administration of the anti-TAT antibody at about 1 μg/kg to about 100 mg/kg body weight results in reduction in tumor size or reduction of tumor cell proliferation within about 5 days to 3 months from the first administration of the antibody, preferably within about 5 to 30 days.

To select for an anti-TAT antibody, TAT binding oligopeptide or TAT binding organic molecule which induces cell death, loss of membrane integrity as indicated by, e.g., propidium iodide (PI), trypan blue or 7AAD uptake may be assessed relative to control. A PI uptake assay can be performed in the absence of complement and immune effector cells. TAT polypeptide-expressing tumor cells are incubated with medium alone or medium containing the appropriate anti-TAT antibody (e.g. at about 10μg/ml), TAT binding oligopeptide or TAT binding organic molecule. The cells are incubated for a 3 day time period. Following each treatment, cells are washed and aliquoted into 35 mm strainer-capped 12 x 75 tubes (1ml per tube, 3 tubes per treatment group) for removal of cell clumps. Tubes then receive PI (10μg/ml). Samples may be analyzed using a FACSCAN® flow cytometer and FACSCONVERT® CellQuest software (Becton Dickinson). Those anti-TAT antibodies, TAT binding oligopeptides or TAT binding organic molecules that induce statistically significant levels of cell death as determined by PI uptake may be selected as cell death-inducing anti-TAT antibodies, TAT binding oligopeptides or TAT binding organic molecules.

To screen for antibodies, oligopeptides or other organic molecules which bind to an epitope on a TAT polypeptide bound by an antibody of interest, a routine cross-blocking assay such as that described in Antibodies, A Laboratory Manual, Cold Spring Harbor Laboratory, Ed Harlow and David Lane (1988), can be performed. This assay can be used to determine if a test antibody, oligopeptide or other organic molecule binds the same site or epitope as a known anti-TAT antibody. Alternatively, or additionally, epitope mapping can be performed by methods known in the art . For example, the antibody sequence can be mutagenized such as by alanine scanning, to identify contact residues. The mutant antibody is initially tested for binding with polyclonal antibody to ensure proper folding. In a different method, peptides corresponding to different regions of a TAT polypeptide can be used in competition assays with the test antibodies or with a test antibody and an antibody with a characterized or known epitope.

E. Antibody Dependent Enzyme Mediated Prodrug Therapy (ADEPT)

The antibodies of the present invention may also be used in ADEPT by conjugating the antibody to a prodrug-activating enzyme which converts a prodrug (e.g., a peptidyl chemotherapeutic agent, see WO81/01145) to an active anti-cancer drug. See, for example, WO 88/07378 and U.S. Patent No. 4,975,278.

The enzyme component of the immunoconjugate useful for ADEPT includes any enzyme capable of acting on a prodrug in such a way so as to covert it into its more active, cytotoxic form.

Enzymes that are useful in the method of this invention include, but are not limited to, alkaline phosphatase useful for converting phosphate-containing prodrugs into free drugs; arylsulfatase useful for converting sulfate-containing prodrugs into free drugs; cytosine deaminase useful for converting non-toxic 5-fluorocytosine into the anti-cancer drug, 5-fluorouracil; proteases, such as serratia protease, thermolysin, subtilisin, carboxypeptidases and cathepsins (such as cathepsins B and L), that are useful for converting peptide-containing prodrugs into free drugs; D-alanylcarboxypeptidases, useful for converting prodrugs that contain D-amino acid substituents; carbohydrate-cleaving enzymes such as β-galactosidase and neuraminidase useful for converting glycosylated prodrugs into free drugs; β-lactamase useful for converting drugs derivatized with β-lactams into free drugs; and penicillin amidases, such as penicillin V amidase or penicillin G amidase, useful for converting drugs derivatized at their amine nitrogens with phenoxyacetyl or phenylacetyl groups, respectively, into free drugs. Alternatively, antibodies with enzymatic activity, also known in the art as "abzymes", can be used to convert the prodrugs of the invention into free active drugs (see, e.g., Massey, Nature 328:457-458 (1987)). Antibody-abzyme conjugates can be prepared as described herein for delivery of the abzyme to a tumor cell population.

The enzymes of this invention can be covalently bound to the anti-TAT antibodies by techniques well known in the art such as the use of the heterobifunctional crosslinking reagents discussed above. Alternatively, fusion proteins comprising at least the antigen binding region of an antibody of the invention linked to at least a functionally active portion of an enzyme of the invention can be constructed using recombinant DNA techniques well known in the art (see, e.g., Neuberger et al., Nature 312:604-608 (1984).

F. Full-Length TAT Polypeptides

The present invention also provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as TAT polypeptides. In particular, cDNAs (partial and full-length) encoding various TAT polypeptides have been identified and isolated, as disclosed in further detail in the Examples below.

As disclosed in the Examples below, various cDNA clones have been deposited with the ATCC. The actual nucleotide sequences of those clones can readily be determined by the skilled artisan by sequencing of the deposited clone using routine methods in the art. The predicted amino acid sequence can be determined from the nucleotide sequence using routine skill. For the TAT polypeptides and encoding nucleic acids described herein, in some cases, Applicants have identified what is believed to be the reading frame best identifiable with the sequence information available at the time.

G. Anti-TAT Antibody and TAT Polypeptide Variants

In addition to the anti-TAT antibodies and full-length native sequence TAT polypeptides described herein, it is contemplated that anti-TAT antibody and TAT polypeptide variants can be prepared. Anti-TAT antibody and TAT polypeptide variants can be prepared by introducing appropriate nucleotide changes into the encoding DNA, and/or by synthesis of the desired antibody or polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the anti-TAT antibody or TAT

polypeptide, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

Variations in the anti-TAT antibodies and TAT polypeptides described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the antibody or polypeptide that results in a change in the amino acid sequence as compared with the native sequence antibody or polypeptide. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the anti-TAT antibody or TAT polypeptide. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the anti-TAT antibody or TAT polypeptide with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

Anti-TAT antibody and TAT polypeptide fragments are provided herein. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full length native antibody or protein. Certain fragments lack amino acid residues that are not essential for a desired biological activity of the anti-TAT antibody or TAT polypeptide.

Anti-TAT antibody and TAT polypeptide fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized. An alternative approach involves generating antibody or polypeptide fragments by enzymatic digestion, e.g., by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the desired fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired antibody or polypeptide fragment, by polymerase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, anti-TAT antibody and TAT polypeptide fragments share at least one biological and/or immunological activity with the native anti-TAT antibody or TAT polypeptide disclosed herein.

In particular embodiments, conservative substitutions of interest are shown in Table 6 under the heading of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions in Table 6, or as further described below in reference to amino acid classes, are introduced and the products screened.

Table 6

Original	Exemplary	Preferred

Residue	Substitutions	Substitutions
Ala (A)	val; leu; ile	val
Arg (R)	lys; gln; asn	lys
Asn (N)	gln; his; lys; arg	gln
Asp (D)	glu	glu
Cys (C)	ser	ser
Gln (Q)	asn	asn
Glu (E)	asp	asp
Gly (G)	pro; ala	ala
His (H)	asn; gln; lys; arg	arg
Ile (I)	leu; val; met; ala; phe; norleucine	leu
Leu (L)	norleucine; ile; val; met; ala; phe	ile
Lys (K)	arg; gln; asn	arg
Met (M)	leu; phe; ile	leu
Phe (F)	leu; val; ile; ala; tyr	leu
Pro (P)	ala	ala
Ser (S)	thr	thr
Thr (T)	ser	ser
Trp (W)	tyr; phe	tyr
Tyr (Y)	trp; phe; thr; ser	phe
Val (V)	ile; leu; met; phe; ala; norleucine	leu

Substantial modifications in function or immunological identity of the anti-TAT antibody or TAT polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

(1) hydrophobic: norleucine, met, ala, val, leu, ile;

(2) neutral hydrophilic: cys, ser, thr;

(3) acidic: asp, glu;

(4) basic: asn, gln, his, lys, arg;

(5) residues that influence chain orientation: gly, pro; and

(6) aromatic: trp, tyr, phe.

Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells et al., Gene, 34:315 (1985)], restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the anti-TAT antibody or TAT polypeptide variant DNA.

Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, Science, 244:1081-1085 (1989)]. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

Any cysteine residue not involved in maintaining the proper conformation of the anti-TAT antibody or TAT polypeptide also may be substituted, generally with serine, to improve the oxidative stability of the molecule and prevent aberrant crosslinking. Conversely, cysteine bond(s) may be added to the anti-TAT antibody or TAT polypeptide to improve its stability (particularly where the antibody is an antibody fragment such as an Fv fragment).

A particularly preferred type of substitutional variant involves substituting one or more hypervariable region residues of a parent antibody (e.g., a humanized or human antibody). Generally, the resulting variant(s) selected for further development will have improved biological properties relative to the parent antibody from which they are generated. A convenient way for generating such substitutional variants involves affinity maturation using phage display. Briefly, several hypervariable region sites (e.g., 6-7 sites) are mutated to generate all possible amino substitutions at each site. The antibody variants thus generated are displayed in a monovalent fashion from filamentous phage particles as fusions to the gene III product of M13 packaged within each particle. The phage-displayed variants are then screened for their biological activity (e.g., binding affinity) as herein disclosed. In order to identify candidate hypervariable region sites for modification, alanine scanning mutagenesis can be performed to identify hypervariable region residues contributing significantly to antigen binding. Alternatively, or additionally, it may be beneficial to analyze a crystal structure of the antigen-antibody complex to identify contact points between the antibody and human TAT polypeptide. Such contact residues and neighboring residues are candidates for substitution according to the techniques elaborated herein. Once such variants are generated, the panel of variants is subjected to screening as

described herein and antibodies with superior properties in one or more relevant assays may be selected for further development.

Nucleic acid molecules encoding amino acid sequence variants of the anti-TAT antibody are prepared by a variety of methods known in the art. These methods include, but are not limited to, isolation from a natural source (in the case of naturally occurring amino acid sequence variants) or preparation by oligonucleotide-mediated (or site-directed) mutagenesis, PCR mutagenesis, and cassette mutagenesis of an earlier prepared variant or a non-variant version of the anti-TAT antibody.

H. Modifications of Anti-TAT Antibodies and TAT Polypeptides

Covalent modifications of anti-TAT antibodies and TAT polypeptides are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of an anti-TAT antibody or TAT polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of the anti-TAT antibody or TAT polypeptide. Derivatization with bifunctional agents is useful, for instance, for crosslinking anti-TAT antibody or TAT polypeptide to a water-insoluble support matrix or surface for use in the method for purifying anti-TAT antibodies, and vice-versa. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimidate.

Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α-amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the anti-TAT antibody or TAT polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the antibody or polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence anti-TAT antibody or TAT polypeptide (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the native sequence anti-TAT antibody or TAT polypeptide. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

Glycosylation of antibodies and other polypeptides is typically either N-linked or O-linked. N-linked refers to the attachment of the carbohydrate moiety to the side chain of an asparagine residue. The tripeptide sequences asparagine-X-serine and asparagine-X-threonine, where X is any amino acid except proline, are the recognition sequences for enzymatic attachment of the carbohydrate moiety to the asparagine side chain. Thus, the presence of either of these tripeptide sequences in a polypeptide creates a potential glycosylation site. O-linked glycosylation refers to the attachment of one of the sugars N-acetylgalactosamine, galactose, or xylose to

a hydroxyamino acid, most commonly serine or threonine, although 5-hydroxyproline or 5-hydroxylysine may also be used.

Addition of glycosylation sites to the anti-TAT antibody or TAT polypeptide is conveniently accomplished by altering the amino acid sequence such that it contains one or more of the above-described tripeptide sequences (for N-linked glycosylation sites). The alteration may also be made by the addition of, or substitution by, one or more serine or threonine residues to the sequence of the original anti-TAT antibody or TAT polypeptide (for O-linked glycosylation sites). The anti-TAT antibody or TAT polypeptide amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the anti-TAT antibody or TAT polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the anti-TAT antibody or TAT polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Removal of carbohydrate moieties present on the anti-TAT antibody or TAT polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of anti-TAT antibody or TAT polypeptide comprises linking the antibody or polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337. The antibody or polypeptide also may be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization (for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacylate) microcapsules, respectively), in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles and nanocapsules), or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences, 16th edition, Oslo, A., Ed., (1980).

The anti-TAT antibody or TAT polypeptide of the present invention may also be modified in a way to form chimeric molecules comprising an anti-TAT antibody or TAT polypeptide fused to another, heterologous polypeptide or amino acid sequence.

In one embodiment, such a chimeric molecule comprises a fusion of the anti-TAT antibody or TAT polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the anti-TAT antibody or TAT polypeptide. The presence of such epitope-tagged forms of the anti-TAT antibody or TAT polypeptide can be

detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the anti-TAT antibody or TAT polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an α-tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

In an alternative embodiment, the chimeric molecule may comprise a fusion of the anti-TAT antibody or TAT polypeptide with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of an anti-TAT antibody or TAT polypeptide in place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH_2 and CH_3, or the hinge, CH_1, CH_2 and CH_3 regions of an IgG1 molecule. For the production of immunoglobulin fusions see also US Patent No. 5,428,130 issued June 27, 1995.

I. Preparation of Anti-TAT Antibodies and TAT Polypeptides

The description below relates primarily to production of anti-TAT antibodies and TAT polypeptides by culturing cells transformed or transfected with a vector containing anti-TAT antibody- and TAT polypeptide-encoding nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare anti-TAT antibodies and TAT polypeptides. For instance, the appropriate amino acid sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of the anti-TAT antibody or TAT polypeptide may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the desired anti-TAT antibody or TAT polypeptide.

1. Isolation of DNA Encoding Anti-TAT Antibody or TAT Polypeptide

DNA encoding anti-TAT antibody or TAT polypeptide may be obtained from a cDNA library prepared from tissue believed to possess the anti-TAT antibody or TAT polypeptide mRNA and to express it at

a detectable level. Accordingly, human anti-TAT antibody or TAT polypeptide DNA can be conveniently obtained from a cDNA library prepared from human tissue. The anti-TAT antibody- or TAT polypeptide-encoding gene may also be obtained from a genomic library or by known synthetic procedures (e.g., automated nucleic acid synthesis).

Libraries can be screened with probes (such as oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding anti-TAT antibody or TAT polypeptide is to use PCR methodology [Sambrook et al., supra; Dieffanbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

Techniques for screening a cDNA library are well known in the art. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ^{32}P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., supra.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined using methods known in the art and as described herein.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for anti-TAT antibody or TAT polypeptide production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

Methods of eukaryotic cell transfection and prokaryotic cell transformation are known to the ordinarily skilled artisan, for example, $CaCl_2$, $CaPO_4$, liposome-mediated and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is

generally used for prokaryotes. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transfections have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Other suitable prokaryotic host cells include Enterobacteriaceae such as *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescans*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (e.g., *B. licheniformis* 41P disclosed in DD 266,710 published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype *tonA* ; *E. coli* W3110 strain 9E4, which has the complete genotype *tonA ptr3*; *E. coli* W3110 strain 27C7 (ATCC 55,244), which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT kan*r; *E. coli* W3110 strain 37D6, which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan*r; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, e.g., PCR or other nucleic acid polymerase reactions, are suitable.

Full length antibody, antibody fragments, and antibody fusion proteins can be produced in bacteria, in particular when glycosylation and Fc effector function are not needed, such as when the therapeutic antibody is conjugated to a cytotoxic agent (e.g., a toxin) and the immunoconjugate by itself shows effectiveness in tumor cell destruction. Full length antibodies have greater half life in circulation. Production in E. coli is faster and more cost efficient. For expression of antibody fragments and polypeptides in bacteria, see, e.g., U.S. 5,648,237 (Carter et. al.), U.S. 5,789,199 (Joly et al.), and U.S. 5,840,523 (Simmons et al.) which describes translation initiation regio (TIR) and signal sequences for optimizing expression and secretion, these patents

incorporated herein by reference. After expression, the antibody is isolated from the E. coli cell paste in a soluble fraction and can be purified through, e.g., a protein A or G column depending on the isotype. Final purification can be carried out similar to the process for purifying antibody expressed e.g., in CHO cells.

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for anti-TAT antibody- or TAT polypeptide-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, Nature, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Fleer et al., Bio/Technology, 9:968-975 (1991)) such as, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al., J. Bacteriol., 154(2):737-742 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickeramii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilarum* (ATCC 36,906; Van den Berg et al., Bio/Technology, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus; yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna et al., J. Basic Microbiol., 28:265-278 [1988]); *Candida; Trichoderma reesia* (EP 244,234); *Neurospora crassa* (Case et al., Proc. Natl. Acad. Sci. USA, 76:5259-5263 [1979]); *Schwanniomyces* such as *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, e.g., *Neurospora, Penicillium, Tolypocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance et al., Biochem. Biophys. Res. Commun., 112:284-289 [1983]; Tilburn et al., Gene, 26:205-221 [1983]; Yelton et al., Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, EMBO J., 4:475-479 [1985]). Methylotropic yeasts are suitable herein and include, but are not limited to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula, Candida, Kloeckera, Pichia, Saccharomyces, Torulopsis,* and *Rhodotorula*. A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, The Biochemistry of Methylotrophs, 269 (1982).

Suitable host cells for the expression of glycosylated anti-TAT antibody or TAT polypeptide are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as Drosophila S2 and Spodoptera Sf9, as well as plant cells, such as cell cultures of cotton, corn, potato, soybean, petunia, tomato, and tobacco. Numerous baculoviral strains and variants and corresponding permissive insect host cells from hosts such as *Spodoptera frugiperda* (caterpillar), *Aedes aegypti* (mosquito), *Aedes albopictus* (mosquito), *Drosophila melanogaster* (fruitfly), and *Bombyx mori* have been identified. A variety of viral strains for transfection are publicly available, e.g., the L-1 variant of *Autographa californica* NPV and the Bm-5 strain of *Bombyx mori* NPV, and such viruses may be used as the virus herein according to the present invention, particularly for transfection of *Spodoptera frugiperda* cells.

However, interest has been greatest in vertebrate cells, and propagation of vertebrate cells in culture (tissue culture) has become a routine procedure. Examples of useful mammalian host cell lines are monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen Virol, 36:59 (1977)); baby hamster kidney cells (BHK, ATCC CCL 10); Chinese hamster ovary cells/-DHFR (CHO, Urlaub et al., Proc. Natl. Acad. Sci. USA 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod. 23:243-251 (1980));

monkey kidney cells (CV1 ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL51); TRI cells (Mather et al., Annals N.Y. Acad. Sci. 383:44-68 (1982)); MRC 5 cells; FS4 cells; and a human hepatoma line (Hep G2).

Host cells are transformed with the above-described expression or cloning vectors for anti-TAT antibody or TAT polypeptide production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences.

3. Selection and Use of a Replicable Vector

The nucleic acid (e.g., cDNA or genomic DNA) encoding anti-TAT antibody or TAT polypeptide may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

The TAT may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the anti-TAT antibody- or TAT polypeptide-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g., the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α-factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g.,

ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*.

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the anti-TAT antibody- or TAT polypeptide-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp*1 gene present in the yeast plasmid YRp7 [Stinchcomb et al., Nature, 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)]. The *trp*1 gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

Expression and cloning vectors usually contain a promoter operably linked to the anti-TAT antibody- or TAT polypeptide-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β-lactamase and lactose promoter systems [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (trp) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid promoters such as the tac promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding anti-TAT antibody or TAT polypeptide.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

Anti-TAT antibody or TAT polypeptide transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a DNA encoding the anti-TAT antibody or TAT polypeptide by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of

DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α-fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the anti-TAT antibody or TAT polypeptide coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding anti-TAT antibody or TAT polypeptide.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of anti-TAT antibody or TAT polypeptide in recombinant vertebrate cell culture are described in Gething et al., Nature, 293:620-625 (1981); Mantei et al., Nature, 281:40-46 (1979); EP 117,060; and EP 117,058.

4. Culturing the Host Cells

The host cells used to produce the anti-TAT antibody or TAT polypeptide of this invention may be cultured in a variety of media. Commercially available media such as Ham's F10 (Sigma), Minimal Essential Medium ((MEM), (Sigma), RPMI-1640 (Sigma), and Dulbecco's Modified Eagle's Medium ((DMEM), Sigma) are suitable for culturing the host cells. In addition, any of the media described in Ham et al., Meth. Enz. 58:44 (1979), Barnes et al., Anal. Biochem.102:255 (1980), U.S. Pat. Nos. 4,767,704; 4,657,866; 4,927,762; 4,560,655; or 5,122,469; WO 90/03430; WO 87/00195; or U.S. Patent Re. 30,985 may be used as culture media for the host cells. Any of these media may be supplemented as necessary with hormones and/or other growth factors (such as insulin, transferrin, or epidermal growth factor), salts (such as sodium chloride, calcium, magnesium, and phosphate), buffers (such as HEPES), nucleotides (such as adenosine and thymidine), antibiotics (such as GENTAMYCIN™ drug), trace elements (defined as inorganic compounds usually present at final concentrations in the micromolar range), and glucose or an equivalent energy source. Any other necessary supplements may also be included at appropriate concentrations that would be known to those skilled in the art. The culture conditions, such as temperature, pH, and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

5. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA

hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence TAT polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to TAT DNA and encoding a specific antibody epitope.

6. Purification of Anti-TAT Antibody and TAT Polypeptide

Forms of anti-TAT antibody and TAT polypeptide may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic cleavage. Cells employed in expression of anti-TAT antibody and TAT polypeptide can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify anti-TAT antibody and TAT polypeptide from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the anti-TAT antibody and TAT polypeptide. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular anti-TAT antibody or TAT polypeptide produced.

When using recombinant techniques, the antibody can be produced intracellularly, in the periplasmic space, or directly secreted into the medium. If the antibody is produced intracellularly, as a first step, the particulate debris, either host cells or lysed fragments, are removed, for example, by centrifugation or ultrafiltration. Carter et al., Bio/Technology 10:163-167 (1992) describe a procedure for isolating antibodies which are secreted to the periplasmic space of *E. coli*. Briefly, cell paste is thawed in the presence of sodium acetate (pH 3.5), EDTA, and phenylmethylsulfonylfluoride (PMSF) over about 30 min. Cell debris can be removed by centrifugation. Where the antibody is secreted into the medium, supernatants from such expression systems are generally first concentrated using a commercially available protein concentration filter, for example, an Amicon or Millipore Pellicon ultrafiltration unit. A protease inhibitor such as PMSF may be included in any of the foregoing steps to inhibit proteolysis and antibiotics may be included to prevent the growth of adventitious contaminants.

The antibody composition prepared from the cells can be purified using, for example, hydroxylapatite chromatography, gel electrophoresis, dialysis, and affinity chromatography, with affinity chromatography being the preferred purification technique. The suitability of protein A as an affinity ligand depends on the species and isotype of any immunoglobulin Fc domain that is present in the antibody. Protein A can be used to purify antibodies that are based on human γ1, γ2 or γ4 heavy chains (Lindmark et al., J. Immunol. Meth. 62:1-13 (1983)). Protein G is recommended for all mouse isotypes and for human γ3 (Guss et al., EMBO J. 5:15671575 (1986)). The matrix to which the affinity ligand is attached is most often agarose, but other matrices are available. Mechanically stable matrices such as controlled pore glass or poly(styrenedivinyl)benzene allow for faster flow rates and shorter processing times than can be achieved with agarose. Where the antibody comprises a C_H3 domain, the Bakerbond ABX™resin (J. T. Baker, Phillipsburg, NJ) is useful for purification. Other techniques for protein purification such as fractionation on an ion-exchange column, ethanol precipitation, Reverse Phase HPLC, chromatography on silica, chromatography on heparin SEPHAROSE™ chromatography on an anion or cation exchange resin (such as a polyaspartic acid column), chromatofocusing, SDS-PAGE, and ammonium sulfate precipitation are also available depending on the antibody to be recovered.

Following any preliminary purification step(s), the mixture comprising the antibody of interest and contaminants may be subjected to low pH hydrophobic interaction chromatography using an elution buffer at a pH between about 2.5-4.5, preferably performed at low salt concentrations (e.g., from about 0-0.25M salt).

J. Pharmaceutical Formulations

Therapeutic formulations of the anti-TAT antibodies, TAT binding oligopeptides, TAT siRNA, TAT binding organic molecules and/or TAT polypeptides used in accordance with the present invention are prepared for storage by mixing the antibody, polypeptide, oligopeptide, siRNA or organic molecule having the desired degree of purity with optional pharmaceutically acceptable carriers, excipients or stabilizers (Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)), in the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients, or stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as acetate, Tris, phosphate, citrate, and other organic acids; antioxidants including ascorbic acid and methionine; preservatives (such as octadecyldimethylbenzyl ammonium chloride; hexamethonium chloride; benzalkonium chloride, benzethonium chloride; phenol, butyl or benzyl alcohol; alkyl parabens such as methyl or propyl paraben; catechol; resorcinol; cyclohexanol; 3-pentanol; and m-cresol); low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, histidine, arginine, or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; tonicifiers such as trehalose and sodium chloride; sugars such as sucrose, mannitol, trehalose or sorbitol; surfactant such as polysorbate; salt-forming counter-ions such as sodium; metal complexes (e.g., Zn-protein complexes); and/or non-ionic surfactants such as TWEEN®, PLURONICS® or

polyethylene glycol (PEG). The antibody preferably comprises the antibody at a concentration of between 5-200 mg/ml, preferably between 10-100 mg/ml.

The formulations herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. For example, in addition to an anti-TAT antibody, TAT binding oligopeptide, TAT siRNA or TAT binding organic molecule, it may be desirable to include in the one formulation, an additional antibody, e.g., a second anti-TAT antibody which binds a different epitope on the TAT polypeptide, or an antibody to some other target such as a growth factor that affects the growth of the particular cancer. Alternatively, or additionally, the composition may further comprise a chemotherapeutic agent, cytotoxic agent, cytokine, growth inhibitory agent, anti-hormonal agent, and/or cardioprotectant. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

The active ingredients may also be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles and nanocapsules) or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences, 16th edition, Osol, A. Ed. (1980).

Sustained-release preparations may be prepared. Suitable examples of sustained-release preparations include semi-permeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, e.g., films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and γ ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOT® (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid.

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

K. Diagnosis and Treatment with Anti-TAT Antibodies, TAT Binding Oligopeptides, TAT siRNA and TAT Binding Organic Molecules

To determine TAT expression in the cancer, various diagnostic assays are available. In one embodiment, TAT polypeptide overexpression may be analyzed by immunohistochemistry (IHC). Parrafin embedded tissue sections from a tumor biopsy may be subjected to the IHC assay and accorded a TAT protein staining intensity criteria as follows:

Score 0 - no staining is observed or membrane staining is observed in less than 10% of tumor cells.

Score 1+ - a faint/barely perceptible membrane staining is detected in more than 10% of the tumor cells. The cells are only stained in part of their membrane.

Score 2+ - a weak to moderate complete membrane staining is observed in more than 10% of the tumor cells.

Score 3+ - a moderate to strong complete membrane staining is observed in more than 10% of the tumor cells.

Those tumors with 0 or 1+ scored for TAT polypeptide expression may be characterized as not overexpressing TAT, whereas those tumors with 2+ or 3+ scores may be characterized as overexpressing TAT.

Alternatively, or additionally, FISH assays such as the INFORM® (sold by Ventana, Arizona) or PATHVISION® (Vysis, Illinois) may be carried out on formalin-fixed, paraffin-embedded tumor tissue to determine the extent (if any) of TAT overexpression in the tumor.

TAT overexpression or amplification may be evaluated using an *in vivo* diagnostic assay, e.g., by administering a molecule (such as an antibody, oligopeptide or organic molecule) which binds the molecule to be detected and is tagged with a detectable label (e.g., a radioactive isotope or a fluorescent label) and externally scanning the patient for localization of the label.

As described above, the anti-TAT antibodies, oligopeptides and organic molecules of the invention have various non-therapeutic applications. The anti-TAT antibodies, oligopeptides and organic molecules of the present invention can be useful for diagnosis and staging of TAT polypeptide-expressing cancers (e.g., in radioimaging). The antibodies, oligopeptides and organic molecules are also useful for purification or immunoprecipitation of TAT polypeptide from cells, for detection and quantitation of TAT polypeptide *in vitro*, e.g., in an ELISA or a Western blot, to kill and eliminate TAT-expressing cells from a population of mixed cells as a step in the purification of other cells.

Currently, depending on the stage of the cancer, cancer treatment involves one or a combination of the following therapies: surgery to remove the cancerous tissue, radiation therapy, and chemotherapy. Anti-TAT antibody, oligopeptide, siRNA or organic molecule therapy may be especially desirable in elderly patients who do not tolerate the toxicity and side effects of chemotherapy well and in metastatic disease where radiation therapy has limited usefulness. The tumor targeting anti-TAT antibodies, oligopeptides, siRNA and organic molecules of the invention are useful to alleviate TAT-expressing cancers upon initial diagnosis of the disease or during relapse. For therapeutic applications, the anti-TAT antibody, oligopeptide, siRNA or organic molecule can be used alone, or in combination therapy with, e.g., hormones, antiangiogens, or radiolabelled compounds, or with surgery, cryotherapy, and/or radiotherapy. Anti-TAT antibody, oligopeptide, siRNA or organic molecule treatment can be administered in conjunction with other forms of conventional therapy, either consecutively with, pre- or post-conventional therapy. Chemotherapeutic drugs such as TAXOTERE® (docetaxel), TAXOL® (palictaxel), estramustine and mitoxantrone are used in treating cancer, in particular, in good risk patients. In the present method of the invention for treating or alleviating cancer, the cancer patient can be administered anti-TAT antibody, oligopeptide, siRNA or organic molecule in conjuction with treatment with the one or more of the preceding chemotherapeutic agents. In particular, combination therapy with palictaxel and modified derivatives (see, e.g., EP0600517) is contemplated. The anti-TAT antibody, oligopeptide, siRNA or organic molecule will be administered with a therapeutically effective dose of the chemotherapeutic agent. In another embodiment, the anti-TAT antibody, oligopeptide, siRNA or organic molecule is administered in conjunction with chemotherapy to enhance the activity and efficacy of the

chemotherapeutic agent, e.g., paclitaxel. The Physicians' Desk Reference (PDR) discloses dosages of these agents that have been used in treatment of various cancers. The dosing regimen and dosages of these aforementioned chemotherapeutic drugs that are therapeutically effective will depend on the particular cancer being treated, the extent of the disease and other factors familiar to the physician of skill in the art and can be determined by the physician.

In one particular embodiment, a conjugate comprising an anti-TAT antibody, oligopeptide, siRNA or organic molecule conjugated with a cytotoxic agent is administered to the patient. Preferably, the immunoconjugate bound to the TAT protein is internalized by the cell, resulting in increased therapeutic efficacy of the immunoconjugate in killing the cancer cell to which it binds. In a preferred embodiment, the cytotoxic agent targets or interferes with the nucleic acid in the cancer cell. Examples of such cytotoxic agents are described above and include maytansinoids, calicheamicins, ribonucleases and DNA endonucleases.

The anti-TAT antibodies, oligopeptides, organic molecules or toxin conjugates thereof are administered to a human patient, in accord with known methods, such as intravenous administration, e.g., as a bolus or by continuous infusion over a period of time, by intramuscular, intraperitoneal, intracerobrospinal, subcutaneous, intra-articular, intrasynovial, intrathecal, oral, topical, or inhalation routes. Intravenous or subcutaneous administration of the antibody, oligopeptide or organic molecule is preferred.

Other therapeutic regimens may be combined with the administration of the anti-TAT antibody, oligopeptide or organic molecule. The combined administration includes co-administration, using separate formulations or a single pharmaceutical formulation, and consecutive administration in either order, wherein preferably there is a time period while both (or all) active agents simultaneously exert their biological activities. Preferably such combined therapy results in a synergistic therapeutic effect.

It may also be desirable to combine administration of the anti-TAT antibody or antibodies, oligopeptides or organic molecules, with administration of an antibody directed against another tumor antigen associated with the particular cancer.

In another embodiment, the therapeutic treatment methods of the present invention involves the combined administration of an anti-TAT antibody (or antibodies), oligopeptides or organic molecules and one or more chemotherapeutic agents or growth inhibitory agents, including co-administration of cocktails of different chemotherapeutic agents. Chemotherapeutic agents include estramustine phosphate, prednimustine, cisplatin, 5-fluorouracil, melphalan, cyclophosphamide, hydroxyurea and hydroxyureataxanes (such as paclitaxel and doxetaxel) and/or anthracycline antibiotics. Preparation and dosing schedules for such chemotherapeutic agents may be used according to manufacturers' instructions or as determined empirically by the skilled practitioner. Preparation and dosing schedules for such chemotherapy are also described in Chemotherapy Service Ed., M.C. Perry, Williams & Wilkins, Baltimore, MD (1992).

The antibody, oligopeptide or organic molecule may be combined with an anti-hormonal compound; e.g., an anti-estrogen compound such as tamoxifen; an anti-progesterone such as onapristone (see, EP 616 812); or an anti-androgen such as flutamide, in dosages known for such molecules. Where the cancer to be treated is androgen independent cancer, the patient may previously have been subjected to anti-androgen

therapy and, after the cancer becomes androgen independent, the anti-TAT antibody, oligopeptide or organic molecule (and optionally other agents as described herein) may be administered to the patient.

Sometimes, it may be beneficial to also co-administer a cardioprotectant (to prevent or reduce myocardial dysfunction associated with the therapy) or one or more cytokines to the patient. In addition to the above therapeutic regimes, the patient may be subjected to surgical removal of cancer cells and/or radiation therapy, before, simultaneously with, or post antibody, oligopeptide or organic molecule therapy. Suitable dosages for any of the above co-administered agents are those presently used and may be lowered due to the combined action (synergy) of the agent and anti-TAT antibody, oligopeptide or organic molecule.

For the prevention or treatment of disease, the dosage and mode of administration will be chosen by the physician according to known criteria. The appropriate dosage of antibody, oligopeptide or organic molecule will depend on the type of disease to be treated, as defined above, the severity and course of the disease, whether the antibody, oligopeptide or organic molecule is administered for preventive or therapeutic purposes, previous therapy, the patient's clinical history and response to the antibody, oligopeptide or organic molecule, and the discretion of the attending physician. The antibody, oligopeptide or organic molecule is suitably administered to the patient at one time or over a series of treatments. Preferably, the antibody, oligopeptide or organic molecule is administered by intravenous infusion or by subcutaneous injections. Depending on the type and severity of the disease, about 1 μg/kg to about 50 mg/kg body weight (e.g., about 0.1-15mg/kg/dose) of antibody can be an initial candidate dosage for administration to the patient, whether, for example, by one or more separate administrations, or by continuous infusion. A dosing regimen can comprise administering an initial loading dose of about 4 mg/kg, followed by a weekly maintenance dose of about 2 mg/kg of the anti-TAT antibody. However, other dosage regimens may be useful. A typical daily dosage might range from about 1 μg/kg to 100 mg/kg or more, depending on the factors mentioned above. For repeated administrations over several days or longer, depending on the condition, the treatment is sustained until a desired suppression of disease symptoms occurs. The progress of this therapy can be readily monitored by conventional methods and assays and based on criteria known to the physician or other persons of skill in the art.

Aside from administration of the antibody protein to the patient, the present application contemplates administration of the antibody by gene therapy. Such administration of nucleic acid encoding the antibody is encompassed by the expression "administering a therapeutically effective amount of an antibody". See, for example, WO96/07321 published March 14, 1996 concerning the use of gene therapy to generate intracellular antibodies.

There are two major approaches to getting the nucleic acid (optionally contained in a vector) into the patient's cells; *in vivo* and *ex vivo*. For *in vivo* delivery the nucleic acid is injected directly into the patient, usually at the site where the antibody is required. For *ex vivo* treatment, the patient's cells are removed, the nucleic acid is introduced into these isolated cells and the modified cells are administered to the patient either directly or, for example, encapsulated within porous membranes which are implanted into the patient (see, e.g., U.S. Patent Nos. 4,892,538 and 5,283,187). There are a variety of techniques available for introducing nucleic

acids into viable cells. The techniques vary depending upon whether the nucleic acid is transferred into cultured cells *in vitro*, or *in vivo* in the cells of the intended host. Techniques suitable for the transfer of nucleic acid into mammalian cells *in vitro* include the use of liposomes, electroporation, microinjection, cell fusion, DEAE-dextran, the calcium phosphate precipitation method, etc. A commonly used vector for *ex vivo* delivery of the gene is a retroviral vector.

The currently preferred *in vivo* nucleic acid transfer techniques include transfection with viral vectors (such as adenovirus, Herpes simplex I virus, or adeno-associated virus) and lipid-based systems (useful lipids for lipid-mediated transfer of the gene are DOTMA, DOPE and DC-Chol, for example). For review of the currently known gene marking and gene therapy protocols see Anderson et al., Science 256:808-813 (1992). See also WO 93/25673 and the references cited therein.

The anti-TAT antibodies of the invention can be in the different forms encompassed by the definition of "antibody" herein. Thus, the antibodies include full length or intact antibody, antibody fragments, native sequence antibody or amino acid variants, humanized, chimeric or fusion antibodies, immunoconjugates, and functional fragments thereof. In fusion antibodies an antibody sequence is fused to a heterologous polypeptide sequence. The antibodies can be modified in the Fc region to provide desired effector functions. As discussed in more detail in the sections herein, with the appropriate Fc regions, the naked antibody bound on the cell surface can induce cytotoxicity, e.g., via antibody-dependent cellular cytotoxicity (ADCC) or by recruiting complement in complement dependent cytotoxicity, or some other mechanism. Alternatively, where it is desirable to eliminate or reduce effector function, so as to minimize side effects or therapeutic complications, certain other Fc regions may be used.

In one embodiment, the antibody competes for binding or bind substantially to, the same epitope as the antibodies of the invention. Antibodies having the biological characteristics of the present anti-TAT antibodies of the invention are also contemplated, specifically including the *in vivo* tumor targeting and any cell proliferation inhibition or cytotoxic characteristics.

Methods of producing the above antibodies are described in detail herein.

The present anti-TAT antibodies, oligopeptides, siRNAs and organic molecules are useful for treating a TAT-expressing cancer or alleviating one or more symptoms of the cancer in a mammal. Such a cancer includes prostate cancer, cancer of the urinary tract, lung cancer, breast cancer, colon cancer, colorectal cancer, and ovarian cancer, more specifically, prostate adenocarcinoma, renal cell carcinomas, colorectal adenocarcinomas, lung adenocarcinomas, lung squamous cell carcinomas, and pleural mesothelioma. The cancers encompass metastatic cancers of any of the preceding. The antibody, oligopeptide, siRNA or organic molecule is able to bind to at least a portion of the cancer cells that express TAT polypeptide in the mammal. In a preferred embodiment, the antibody, oligopeptide, siRNA or organic molecule is effective to destroy or kill TAT-expressing tumor cells or inhibit the growth of such tumor cells, *in vitro* or *in vivo*, upon binding to TAT polypeptide on the cell. Such an antibody includes a naked anti-TAT antibody (not conjugated to any agent). Naked antibodies that have cytotoxic or cell growth inhibition properties can be further harnessed with

a cytotoxic agent to render them even more potent in tumor cell destruction. Cytotoxic properties can be conferred to an anti-TAT antibody by, e.g., conjugating the antibody with a cytotoxic agent, to form an immunoconjugate as described herein. The cytotoxic agent or a growth inhibitory agent is preferably a small molecule. Toxins such as calicheamicin or a maytansinoid and analogs or derivatives thereof, are preferable.

The invention provides a composition comprising an anti-TAT antibody, oligopeptide, siRNA or organic molecule of the invention, and a carrier. For the purposes of treating cancer, compositions can be administered to the patient in need of such treatment, wherein the composition can comprise one or more anti-TAT antibodies present as an immunoconjugate or as the naked antibody. In a further embodiment, the compositions can comprise these antibodies, oligopeptides, siRNAs or organic molecules in combination with other therapeutic agents such as cytotoxic or growth inhibitory agents, including chemotherapeutic agents. The invention also provides formulations comprising an anti-TAT antibody, oligopeptide, siRNA or organic molecule of the invention, and a carrier. In one embodiment, the formulation is a therapeutic formulation comprising a pharmaceutically acceptable carrier.

Another aspect of the invention is isolated nucleic acids encoding the anti-TAT antibodies. Nucleic acids encoding both the H and L chains and especially the hypervariable region residues, chains which encode the native sequence antibody as well as variants, modifications and humanized versions of the antibody, are encompassed.

The invention also provides methods useful for treating a TAT polypeptide-expressing cancer or alleviating one or more symptoms of the cancer in a mammal, comprising administering a therapeutically effective amount of an anti-TAT antibody, oligopeptide, siRNA or organic molecule to the mammal. The antibody, oligopeptide or organic molecule therapeutic compositions can be administered short term (acute) or chronic, or intermittent as directed by physician. Also provided are methods of inhibiting the growth of, and killing a TAT polypeptide-expressing cell.

The invention also provides kits and articles of manufacture comprising at least one anti-TAT antibody, oligopeptide, siRNA or organic molecule. Kits containing anti-TAT antibodies, oligopeptides, siRNAs or organic molecules find use, e.g., for TAT cell killing assays, for purification or immunoprecipitation of TAT polypeptide from cells. For example, for isolation and purification of TAT, the kit can contain an anti-TAT antibody, oligopeptide, siRNA or organic molecule coupled to beads (e.g., sepharose beads). Kits can be provided which contain the antibodies, oligopeptides, siRNA or organic molecules for detection and quantitation of TAT *in vitro*, e.g., in an ELISA or a Western blot. Such antibody, oligopeptide, siRNA or organic molecule useful for detection may be provided with a label such as a fluorescent or radiolabel.

L. Articles of Manufacture and Kits

Another embodiment of the invention is an article of manufacture containing materials useful for the treatment of anti-TAT expressing cancer. The article of manufacture comprises a container and a label or package insert on or associated with the container. Suitable containers include, for example, bottles, vials, syringes, etc. The containers may be formed from a variety of materials such as glass or plastic. The container

holds a composition which is effective for treating the cancer condition and may have a sterile access port (for example the container may be an intravenous solution bag or a vial having a stopper pierceable by a hypodermic injection needle). At least one active agent in the composition is an anti-TAT antibody, oligopeptide, siRNA or organic molecule of the invention. The label or package insert indicates that the composition is used for treating cancer. The label or package insert will further comprise instructions for administering the antibody, oligopeptide, siRNA or organic molecule composition to the cancer patient. Additionally, the article of manufacture may further comprise a second container comprising a pharmaceutically-acceptable buffer, such as bacteriostatic water for injection (BWFI), phosphate-buffered saline, Ringer's solution and dextrose solution. It may further include other materials desirable from a commercial and user standpoint, including other buffers, diluents, filters, needles, and syringes.

Kits are also provided that are useful for various purposes , e.g., for TAT-expressing cell killing assays, for purification or immunoprecipitation of TAT polypeptide from cells. For isolation and purification of TAT polypeptide, the kit can contain an anti-TAT antibody, oligopeptide, siRNA or organic molecule coupled to beads (e.g., sepharose beads). Kits can be provided which contain the antibodies, oligopeptides, siRNAs or organic molecules for detection and quantitation of TAT polypeptide *in vitro*, e.g., in an ELISA or a Western blot. As with the article of manufacture, the kit comprises a container and a label or package insert on or associated with the container. The container holds a composition comprising at least one anti-TAT antibody, oligopeptide, siRNA or organic molecule of the invention. Additional containers may be included that contain, e.g., diluents and buffers, control antibodies. The label or package insert may provide a description of the composition as well as instructions for the intended in vitro or diagnostic use.

M. Uses for TAT Polypeptides and TAT-Polypeptide Encoding Nucleic Acids

Nucleotide sequences (or their complement) encoding TAT polypeptides have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA, siRNA and DNA probes. TAT-encoding nucleic acid will also be useful for the preparation of TAT polypeptides by the recombinant techniques described herein, wherein those TAT polypeptides may find use, for example, in the preparation of anti-TAT antibodies as described herein.

The full-length native sequence TAT gene, or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length TAT cDNA or to isolate still other cDNAs (for instance, those encoding naturally-occurring variants of TAT or TAT from other species) which have a desired sequence identity to the native TAT sequence disclosed herein. Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from at least partially novel regions of the full length native nucleotide sequence wherein those regions may be determined without undue experimentation or from genomic sequences including promoters, enhancer elements and introns of native sequence TAT. By way of example, a screening method will comprise isolating the coding region of the TAT gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ^{32}P or ^{35}S, or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of

the TAT gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below. Any EST sequences disclosed in the present application may similarly be employed as probes, using the methods disclosed herein.

Other useful fragments of the TAT-encoding nucleic acids include antisense or sense oligonucleotides comprising a singe-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target TAT mRNA (sense) or TAT DNA (antisense) sequences. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment of the coding region of TAT DNA. Such a fragment generally comprises at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen (Cancer Res. 48:2659, 1988) and van der Krol et al. (BioTechniques 6:958, 1988).

Binding of antisense or sense oligonucleotides to target nucleic acid sequences results in the formation of duplexes that block transcription or translation of the target sequence by one of several means, including enhanced degradation of the duplexes, premature termination of transcription or translation, or by other means. Such methods are encompassed by the present invention. The antisense oligonucleotides thus may be used to block expression of TAT proteins, wherein those TAT proteins may play a role in the induction of cancer in mammals. Antisense or sense oligonucleotides further comprise oligonucleotides having modified sugar-phosphodiester backbones (or other sugar linkages, such as those described in WO 91/06629) and wherein such sugar linkages are resistant to endogenous nucleases. Such oligonucleotides with resistant sugar linkages are stable *in vivo* (i.e., capable of resisting enzymatic degradation) but retain sequence specificity to be able to bind to target nucleotide sequences.

Preferred intragenic sites for antisense binding include the region incorporating the translation initiation/start codon (5'-AUG / 5'-ATG) or termination/stop codon (5'-UAA, 5'-UAG and 5-UGA / 5'-TAA, 5'-TAG and 5'-TGA) of the open reading frame (ORF) of the gene. These regions refer to a portion of the mRNA or gene that encompasses from about 25 to about 50 contiguous nucleotides in either direction (i.e., 5' or 3') from a translation initiation or termination codon. Other preferred regions for antisense binding include: introns; exons; intron-exon junctions; the open reading frame (ORF) or "coding region," which is the region between the translation initiation codon and the translation termination codon; the 5' cap of an mRNA which comprises an N7-methylated guanosine residue joined to the 5'-most residue of the mRNA via a 5'-5' triphosphate linkage and includes 5' cap structure itself as well as the first 50 nucleotides adjacent to the cap; the 5' untranslated region (5'UTR), the portion of an mRNA in the 5' direction from the translation initiation codon, and thus including nucleotides between the 5' cap site and the translation initiation codon of an mRNA or corresponding nucleotides on the gene; and the 3' untranslated region (3'UTR), the portion of an mRNA in the 3' direction from the translation termination codon, and thus including nucleotides between the translation termination codon and 3' end of an mRNA or corresponding nucleotides on the gene.

Specific examples of preferred antisense compounds useful for inhibiting expression of TAT proteins include oligonucleotides containing modified backbones or non-natural internucleoside linkages. Oligonucleotides having modified backbones include those that retain a phosphorus atom in the backbone and those that do not have a phosphorus atom in the backbone. For the purposes of this specification, and as sometimes referenced in the art, modified oligonucleotides that do not have a phosphorus atom in their internucleoside backbone can also be considered to be oligonucleosides. Preferred modified oligonucleotide backbones include, for example, phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotri-esters, methyl and other alkyl phosphonates including 3'-alkylene phosphonates, 5'-alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates including 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, selenophosphates and borano-phosphates having normal 3'-5' linkages, 2'-5' linked analogs of these, and those having inverted polarity wherein one or more internucleotide linkages is a 3' to 3', 5' to 5' or 2' to 2' linkage. Preferred oligonucleotides having inverted polarity comprise a single 3' to 3' linkage at the 3'-most internucleotide linkage i.e. a single inverted nucleoside residue which may be abasic (the nucleobase is missing or has a hydroxyl group in place thereof). Various salts, mixed salts and free acid forms are also included. Representative United States patents that teach the preparation of phosphorus-containing linkages include, but are not limited to, U.S. Pat. Nos.: 3,687,808; 4,469,863; 4,476,301; 5,023,243; 5,177,196; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361; 5,194,599; 5,565,555; 5,527,899; 5,721,218; 5,672,697 and 5,625,050, each of which is herein incorporated by reference.

Preferred modified oligonucleotide backbones that do not include a phosphorus atom therein have backbones that are formed by short chain alkyl or cycloalkyl internucleoside linkages, mixed heteroatom and alkyl or cycloalkyl internucleoside linkages, or one or more short chain heteroatomic or heterocyclic internucleoside linkages. These include those having morpholino linkages (formed in part from the sugar portion of a nucleoside); siloxane backbones; sulfide, sulfoxide and sulfone backbones; formacetyl and thioformacetyl backbones; methylene formacetyl and thioformacetyl backbones; riboacetyl backbones; alkene containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH.sub.2 component parts. Representative United States patents that teach the preparation of such oligonucleosides include, but are not limited to,. U.S. Pat. Nos.: 5,034,506; 5,166,315; 5,185,444; 5,214,134; 5,216,141; 5,235,033; 5,264,562; 5,264,564; 5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677; 5,541,307; 5,561,225; 5,596,086; 5,602,240; 5,610,289; 5,602,240; 5,608,046; 5,610,289; 5,618,704; 5,623,070; 5,663,312; 5,633,360; 5,677,437; 5,792,608; 5,646,269 and 5,677,439, each of which is herein incorporated by reference.

In other preferred antisense oligonucleotides, both the sugar and the internucleoside linkage, i.e., the backbone, of the nucleotide units are replaced with novel groups. The base units are maintained for hybridization with an appropriate nucleic acid target compound. One such oligomeric compound, an

oligonucleotide mimetic that has been shown to have excellent hybridization properties, is referred to as a peptide nucleic acid (PNA). In PNA compounds, the sugar-backbone of an oligonucleotide is replaced with an amide containing backbone, in particular an aminoethylglycine backbone. The nucleobases are retained and are bound directly or indirectly to aza nitrogen atoms of the amide portion of the backbone. Representative United States patents that teach the preparation of PNA compounds include, but are not limited to, U.S. Pat. Nos.: 5,539,082; 5,714,331; and 5,719,262, each of which is herein incorporated by reference. Further teaching of PNA compounds can be found in Nielsen et al., Science, 1991, 254, 1497-1500.

Preferred antisense oligonucleotides incorporate phosphorothioate backbones and/or heteroatom backbones, and in particular $-CH_2-NH-O-CH_2-$, $-CH_2-N(CH_3)-O-CH_2-$ [known as a methylene (methylimino) or MMI backbone], $-CH_2-O-N(CH_3)-CH_2-$, $-CH_2-N(CH_3)-N(CH_3)-CH_2-$ and $-O-N(CH_3)-CH_2-CH_2-$ [wherein the native phosphodiester backbone is represented as $-O-P-O-CH_2-$] described in the above referenced U.S. Pat. No. 5,489,677, and the amide backbones of the above referenced U.S. Pat. No. 5,602,240. Also preferred are antisense oligonucleotides having morpholino backbone structures of the above-referenced U.S. Pat. No. 5,034,506.

Modified oligonucleotides may also contain one or more substituted sugar moieties. Preferred oligonucleotides comprise one of the following at the 2' position: OH; F; O-alkyl, S-alkyl, or N-alkyl; O-alkenyl, S-alkenyl, or N-alkenyl; O-alkynyl, S-alkynyl or N-alkynyl; or O-alkyl-O-alkyl, wherein the alkyl, alkenyl and alkynyl may be substituted or unsubstituted C_1 to C_{10} alkyl or C_2 to C_{10} alkenyl and alkynyl. Particularly preferred are $O[(CH_2)_nO]_mCH_3$, $O(CH_2)_nOCH_3$, $O(CH_2)_nNH_2$, $O(CH_2)_nCH_3$, $O(CH_2)_nONH_2$, and $O(CH_2)_nON[(CH_2)_nCH_3)]_2$, where n and m are from 1 to about 10. Other preferred antisense oligonucleotides comprise one of the following at the 2' position: C_1 to C_{10} lower alkyl, substituted lower alkyl, alkenyl, alkynyl, alkaryl, aralkyl, O-alkaryl or O-aralkyl, SH, SCH_3, OCN, Cl, Br, CN, CF_3, OCF_3, $SOCH_3$, SO_2CH_3, ONO_2, NO_2, N_3, NH_2, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, an RNA cleaving group, a reporter group, an intercalator, a group for improving the pharmacokinetic properties of an oligonucleotide, or a group for improving the pharmacodynamic properties of an oligonucleotide, and other substituents having similar properties. A preferred modification includes 2'-methoxyethoxy ($2'-O-CH_2CH_2OCH_3$, also known as 2'-O-(2-methoxyethyl) or 2'-MOE) (Martin et al., Helv. Chim. Acta, 1995, 78, 486-504) i.e., an alkoxyalkoxy group. A further preferred modification includes 2'-dimethylaminooxyethoxy, i.e., a $O(CH_2)_2ON(CH_3)_2$ group, also known as 2'-DMAOE, as described in examples hereinbelow, and 2'-dimethylaminoethoxyethoxy (also known in the art as 2'-O-dimethylaminoethoxyethyl or 2'-DMAEOE), i.e., $2'-O-CH_2-O-CH_2-N(CH_3)_2$.

A further preferred modification includes Locked Nucleic Acids (LNAs) in which the 2'-hydroxyl group is linked to the 3' or 4' carbon atom of the sugar ring thereby forming a bicyclic sugar moiety. The linkage is preferably a methelyne ($-CH_2-$)$_n$ group bridging the 2' oxygen atom and the 4' carbon atom wherein n is 1 or 2. LNAs and preparation thereof are described in WO 98/39352 and WO 99/14226.

Other preferred modifications include 2'-methoxy ($2'-O-CH_3$), 2'-aminopropoxy ($2'-OCH_2CH_2CH_2NH_2$), 2'-allyl ($2'-CH_2-CH=CH_2$), 2'-O-allyl ($2'-O-CH_2-CH=CH_2$) and 2'-fluoro (2'-F). The 2'-modification may

be in the arabino (up) position or ribo (down) position. A preferred 2'-arabino modification is 2'-F. Similar modifications may also be made at other positions on the oligonucleotide, particularly the 3' position of the sugar on the 3' terminal nucleotide or in 2'-5' linked oligonucleotides and the 5' position of 5' terminal nucleotide. Oligonucleotides may also have sugar mimetics such as cyclobutyl moieties in place of the pentofuranosyl sugar. Representative United States patents that teach the preparation of such modified sugar structures include, but are not limited to, U.S. Pat. Nos.: 4,981,957; 5,118,800; 5,319,080; 5,359,044; 5,393,878; 5,446,137; 5,466,786; 5,514,785; 5,519,134; 5,567,811; 5,576,427; 5,591,722; 5,597,909; 5,610,300; 5,627,053; 5,639,873; 5,646,265; 5,658,873; 5,670,633; 5,792,747; and 5,700,920, each of which is herein incorporated by reference in its entirety.

Oligonucleotides may also include nucleobase (often referred to in the art simply as "base") modifications or substitutions. As used herein, "unmodified" or "natural" nucleobases include the purine bases adenine (A) and guanine (G), and the pyrimidine bases thymine (T), cytosine (C) and uracil (U). Modified nucleobases include other synthetic and natural nucleobases such as 5-methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-methyl and other alkyl derivatives of adenine and guanine, 2-propyl and other alkyl derivatives of adenine and guanine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-halouracil and cytosine, 5-propynyl (-C≡C-CH_3 or -CH_2-C≡CH) uracil and cytosine and other alkynyl derivatives of pyrimidine bases, 6-azo uracil, cytosine and thymine, 5-uracil (pseudouracil), 4-thiouracil, 8-halo, 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl and other 8-substituted adenines and guanines, 5-halo particularly 5-bromo, 5-trifluoromethyl and other 5-substituted uracils and cytosines, 7-methylguanine and 7-methyladenine, 2-F-adenine, 2-amino-adenine, 8-azaguanine and 8-azaadenine, 7-deazaguanine and 7-deazaadenine and 3-deazaguanine and 3-deazaadenine. Further modified nucleobases include tricyclic pyrimidines such as phenoxazine cytidine(1H-pyrimido[5,4-b][1,4]benzoxazin-2(3H)-one), phenothiazine cytidine (1H-pyrimido[5,4-b][1,4]benzothiazin-2(3H)-one), G-clamps such as a substituted phenoxazine cytidine (e.g. 9-(2-aminoethoxy)-H-pyrimido[5,4-b][1,4]benzoxazin-2(3H)-one), carbazole cytidine (2H-pyrimido[4,5-b]indol-2-one), pyridoindole cytidine (H-pyrido[3',2':4,5]pyrrolo[2,3-d]pyrimidin-2-one). Modified nucleobases may also include those in which the purine or pyrimidine base is replaced with other heterocycles, for example 7-deaza-adenine, 7-deazaguanosine, 2-aminopyridine and 2-pyridone. Further nucleobases include those disclosed in U.S. Pat. No. 3,687,808, those disclosed in The Concise Encyclopedia Of Polymer Science And Engineering, pages 858-859, Kroschwitz, J. I., ed. John Wiley & Sons, 1990, and those disclosed by Englisch et al., Angewandte Chemie, International Edition, 1991, 30, 613. Certain of these nucleobases are particularly useful for increasing the binding affinity of the oligomeric compounds of the invention. These include 5-substituted pyrimidines, 6-azapyrimidines and N-2, N-6 and O-6 substituted purines, including 2-aminopropyladenine, 5-propynyluracil and 5-propynylcytosine. 5-methylcytosine substitutions have been shown to increase nucleic acid duplex stability by 0.6-1.2.degree. C. (Sanghvi et al, Antisense Research and Applications, CRC Press, Boca Raton, 1993, pp. 276-278) and are preferred base substitutions, even more particularly when combined with 2'-O-methoxyethyl sugar modifications. Representative United States patents that teach the preparation of modified nucleobases include, but are not

limited to: U.S. Pat. No. 3,687,808, as well as U.S. Pat. Nos.: 4,845,205; 5,130,302; 5,134,066; 5,175,273; 5,367,066; 5,432,272; 5,457,187; 5,459,255; 5,484,908; 5,502,177; 5,525,711; 5,552,540; 5,587,469; 5,594,121, 5,596,091; 5,614,617; 5,645,985; 5,830,653; 5,763,588; 6,005,096; 5,681,941 and 5,750,692, each of which is herein incorporated by reference.

Another modification of antisense oligonucleotides chemically linking to the oligonucleotide one or more moieties or conjugates which enhance the activity, cellular distribution or cellular uptake of the oligonucleotide. The compounds of the invention can include conjugate groups covalently bound to functional groups such as primary or secondary hydroxyl groups. Conjugate groups of the invention include intercalators, reporter molecules, polyamines, polyamides, polyethylene glycols, polyethers, groups that enhance the pharmacodynamic properties of oligomers, and groups that enhance the pharmacokinetic properties of oligomers. Typical conjugates groups include cholesterols, lipids, cation lipids, phospholipids, cationic phospholipids, biotin, phenazine, folate, phenanthridine, anthraquinone, acridine, fluoresceins, rhodamines, coumarins, and dyes. Groups that enhance the pharmacodynamic properties, in the context of this invention, include groups that improve oligomer uptake, enhance oligomer resistance to degradation, and/or strengthen sequence-specific hybridization with RNA. Groups that enhance the pharmacokinetic properties, in the context of this invention, include groups that improve oligomer uptake, distribution, metabolism or excretion. Conjugate moieties include but are not limited to lipid moieties such as a cholesterol moiety (Letsinger et al., Proc. Natl. Acad. Sci. USA, 1989, 86, 6553-6556), cholic acid (Manoharan et al., Bioorg. Med. Chem. Let., 1994, 4, 1053-1060), a thioether, e.g., hexyl-S-tritylthiol (Manoharan et al., Ann. N.Y. Acad. Sci., 1992, 660, 306-309; Manoharan et al., Bioorg. Med. Chem. Let., 1993, 3, 2765-2770), a thiocholesterol (Oberhauser et al., Nucl. Acids Res., 1992, 20, 533-538), an aliphatic chain, e.g., dodecandiol or undecyl residues (Saison-Behmoaras et al., EMBO J., 1991, 10, 1111-1118; Kabanov et al., FEBS Lett., 1990, 259, 327-330; Svinarchuk et al., Biochimie, 1993, 75, 49-54), a phospholipid, e.g., di-hexadecyl-rac-glycerol or triethyl-ammonium 1,2-di-O-hexadecyl-rac-glycero-3-H-phosphonate (Manoharan et al., Tetrahedron Lett., 1995, 36, 3651-3654; Shea et al., Nucl. Acids Res., 1990, 18, 3777-3783), a polyamine or a polyethylene glycol chain (Manoharan et al., Nucleosides & Nucleotides, 1995, 14, 969-973), or adamantane acetic acid (Manoharan et al., Tetrahedron Lett., 1995, 36, 3651-3654), a palmityl moiety (Mishra et al., Biochim. Biophys. Acta, 1995, 1264, 229-237), or an octadecylamine or hexylamino-carbonyl-oxycholesterol moiety. Oligonucleotides of the invention may also be conjugated to active drug substances, for example, aspirin, warfarin, phenylbutazone, ibuprofen, suprofen, fenbufen, ketoprofen, (S)-(+)-pranoprofen, carprofen, dansylsarcosine, 2,3,5-triiodobenzoic acid, flufenamic acid, folinic acid, a benzothiadiazide, chlorothiazide, a diazepine, indomethicin, a barbiturate, a cephalosporin, a sulfa drug, an antidiabetic, an antibacterial or an antibiotic. Oligonucleotide-drug conjugates and their preparation are described in U.S. patent application Ser. No. 09/334,130 (filed Jun. 15, 1999) and United States patents Nos.: 4,828,979; 4,948,882; 5,218,105; 5,525,465; 5,541,313; 5,545,730; 5,552,538; 5,578,717, 5,580,731; 5,580,731; 5,591,584; 5,109,124; 5,118,802; 5,138,045; 5,414,077; 5,486,603; 5,512,439; 5,578,718; 5,608,046; 4,587,044; 4,605,735; 4,667,025; 4,762,779; 4,789,737; 4,824,941; 4,835,263; 4,876,335; 4,904,582; 4,958,013; 5,082,830;

5,112,963; 5,214,136; 5,082,830; 5,112,963; 5,214,136; 5,245,022; 5,254,469; 5,258,506; 5,262,536; 5,272,250; 5,292,873; 5,317,098; 5,371,241, 5,391,723; 5,416,203, 5,451,463; 5,510,475; 5,512,667; 5,514,785; 5,565,552; 5,567,810; 5,574,142; 5,585,481; 5,587,371; 5,595,726; 5,597,696; 5,599,923; 5,599,928 and 5,688,941, each of which is herein incorporated by reference.

It is not necessary for all positions in a given compound to be uniformly modified, and in fact more than one of the aforementioned modifications may be incorporated in a single compound or even at a single nucleoside within an oligonucleotide. The present invention also includes antisense compounds which are chimeric compounds. "Chimeric" antisense compounds or "chimeras," in the context of this invention, are antisense compounds, particularly oligonucleotides, which contain two or more chemically distinct regions, each made up of at least one monomer unit, i.e., a nucleotide in the case of an oligonucleotide compound. These oligonucleotides typically contain at least one region wherein the oligonucleotide is modified so as to confer upon the oligonucleotide increased resistance to nuclease degradation, increased cellular uptake, and/or increased binding affinity for the target nucleic acid. An additional region of the oligonucleotide may serve as a substrate for enzymes capable of cleaving RNA:DNA or RNA:RNA hybrids. By way of example, RNase H is a cellular endonuclease which cleaves the RNA strand of an RNA:DNA duplex. Activation of RNase H, therefore, results in cleavage of the RNA target, thereby greatly enhancing the efficiency of oligonucleotide inhibition of gene expression. Consequently, comparable results can often be obtained with shorter oligonucleotides when chimeric oligonucleotides are used, compared to phosphorothioate deoxyoligonucleotides hybridizing to the same target region. Chimeric antisense compounds of the invention may be formed as composite structures of two or more oligonucleotides, modified oligonucleotides, oligonucleosides and/or oligonucleotide mimetics as described above. Preferred chimeric antisense oligonucleotides incorporate at least one 2' modified sugar (preferably 2'-O-$(CH_2)_2$-O-CH_3) at the 3' terminal to confer nuclease resistance and a region with at least 4 contiguous 2'-H sugars to confer RNase H activity. Such compounds have also been referred to in the art as hybrids or gapmers. Preferred gapmers have a region of 2' modified sugars (preferably 2'-O-$(CH_2)_2$-O-CH_3) at the 3'-terminal and at the 5' terminal separated by at least one region having at least 4 contiguous 2'-H sugars and preferably incorporate phosphorothioate backbone linkages. Representative United States patents that teach the preparation of such hybrid structures include, but are not limited to, U.S. Pat. Nos. 5,013,830; 5,149,797; 5,220,007; 5,256,775; 5,366,878; 5,403,711; 5,491,133; 5,565,350; 5,623,065; 5,652,355; 5,652,356; and 5,700,922, each of which is herein incorporated by reference in its entirety.

The antisense compounds used in accordance with this invention may be conveniently and routinely made through the well-known technique of solid phase synthesis. Equipment for such synthesis is sold by several vendors including, for example, Applied Biosystems (Foster City, Calif.). Any other means for such synthesis known in the art may additionally or alternatively be employed. It is well known to use similar techniques to prepare oligonucleotides such as the phosphorothioates and alkylated derivatives. The compounds of the invention may also be admixed, encapsulated, conjugated or otherwise associated with other molecules, molecule structures or mixtures of compounds, as for example, liposomes, receptor targeted

molecules, oral, rectal, topical or other formulations, for assisting in uptake, distribution and/or absorption. Representative United States patents that teach the preparation of such uptake, distribution and/or absorption assisting formulations include, but are not limited to, U.S. Pat. Nos. 5,108,921; 5,354,844; 5,416,016; 5,459,127; 5,521,291; 5,543,158; 5,547,932; 5,583,020; 5,591,721; 4,426,330; 4,534,899; 5,013,556; 5,108,921; 5,213,804; 5,227,170; 5,264,221; 5,356,633; 5,395,619; 5,416,016; 5,417,978; 5,462,854; 5,469,854; 5,512,295; 5,527,528; 5,534,259; 5,543,152; 5,556,948; 5,580,575; and 5,595,756, each of which is herein incorporated by reference.

Other examples of sense or antisense oligonucleotides include those oligonucleotides which are covalently linked to organic moieties, such as those described in WO 90/10048, and other moieties that increases affinity of the oligonucleotide for a target nucleic acid sequence, such as poly-(L-lysine). Further still, intercalating agents, such as ellipticine, and alkylating agents or metal complexes may be attached to sense or antisense oligonucleotides to modify binding specificities of the antisense or sense oligonucleotide for the target nucleotide sequence.

Antisense or sense oligonucleotides may be introduced into a cell containing the target nucleic acid sequence by any gene transfer method, including, for example, $CaPO_4$-mediated DNA transfection, electroporation, or by using gene transfer vectors such as Epstein-Barr virus. In a preferred procedure, an antisense or sense oligonucleotide is inserted into a suitable retroviral vector. A cell containing the target nucleic acid sequence is contacted with the recombinant retroviral vector, either *in vivo* or *ex vivo*. Suitable retroviral vectors include, but are not limited to, those derived from the murine retrovirus M-MuLV, N2 (a retrovirus derived from M-MuLV), or the double copy vectors designated DCT5A, DCT5B and DCT5C (see WO 90/13641).

Sense or antisense oligonucleotides also may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell.

Alternatively, a sense or an antisense oligonucleotide may be introduced into a cell containing the target nucleic acid sequence by formation of an oligonucleotide-lipid complex, as described in WO 90/10448. The sense or antisense oligonucleotide-lipid complex is preferably dissociated within the cell by an endogenous lipase.

Antisense or sense RNA or DNA molecules are generally at least about 5 nucleotides in length, alternatively at least about 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750,

760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, or 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length.

Alternatively, a double stranded RNA can be generated. Double stranded RNA that are under 30 nucleotides in length will inhibit the expression of specific genes when introduced into a cell. This mechanism is known as RNA mediated interference (RNAi) and small (under 30 nuclotide) RNAs used as a reagent are known as siRNAs. TAT interfering RNAs may be identified and synthesized using known methods (Shi Y *Trends in Genetics* 19(1):9-12 (2003), WO/2003056012 and WO2003064621). siRNAs are useful to reduce the amount of gene expression in conditions where a reduction in the expression of the target gene would alleviate the condition or disorder.

The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related TAT coding sequences.

Nucleotide sequences encoding a TAT can also be used to construct hybridization probes for mapping the gene which encodes that TAT and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

When the coding sequences for TAT encode a protein which binds to another protein (example, where the TAT is a receptor), the TAT can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor TAT can be used to isolate correlative ligand(s). Screening assays can be designed to find lead compounds that mimic the biological activity of a native TAT or a receptor for TAT. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

Nucleic acids which encode TAT or its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding TAT can be used to clone genomic DNA encoding TAT in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding TAT. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009.

Typically, particular cells would be targeted for TAT transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding TAT introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding TAT. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of TAT can be used to construct a TAT "knock out" animal which has a defective or altered gene encoding TAT as a result of homologous recombination between the endogenous gene encoding TAT and altered genomic DNA encoding TAT introduced into an embryonic stem cell of the animal. For example, cDNA encoding TAT can be used to clone genomic DNA encoding TAT in accordance with established techniques. A portion of the genomic DNA encoding TAT can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, Cell, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., Cell, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the TAT polypeptide.

Nucleic acid encoding the TAT polypeptides may also be used in gene therapy. In gene therapy applications, genes are introduced into cells in order to achieve *in vivo* synthesis of a therapeutically effective genetic product, for example for replacement of a defective gene. "Gene therapy" includes both conventional gene therapy where a lasting effect is achieved by a single treatment, and the administration of gene therapeutic agents, which involves the one time or repeated administration of a therapeutically effective DNA or mRNA. Antisense RNAs and DNAs can be used as therapeutic agents for blocking the expression of certain genes *in vivo*. It has already been shown that short antisense oligonucleotides can be imported into cells where they act as inhibitors, despite their low intracellular concentrations caused by their restricted uptake by the cell membrane. (Zamecnik *et al.*, Proc. Natl. Acad. Sci. USA 83:4143-4146 [1986]). The oligonucleotides can be modified to enhance their uptake, e.g. by substituting their negatively charged phosphodiester groups by uncharged groups.

There are a variety of techniques available for introducing nucleic acids into viable cells. The techniques vary depending upon whether the nucleic acid is transferred into cultured cells *in vitro*, or *in vivo* in the cells of

the intended host. Techniques suitable for the transfer of nucleic acid into mammalian cells *in vitro* include the use of liposomes, electroporation, microinjection, cell fusion, DEAE-dextran, the calcium phosphate precipitation method, etc. The currently preferred *in vivo* gene transfer techniques include transfection with viral (typically retroviral) vectors and viral coat protein-liposome mediated transfection (Dzau et al., Trends in Biotechnology 11, 205-210 [1993]). In some situations it is desirable to provide the nucleic acid source with an agent that targets the target cells, such as an antibody specific for a cell surface membrane protein or the target cell, a ligand for a receptor on the target cell, etc. Where liposomes are employed, proteins which bind to a cell surface membrane protein associated with endocytosis may be used for targeting and/or to facilitate uptake, e.g. capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half-life. The technique of receptor-mediated endocytosis is described, for example, by Wu et al., J. Biol. Chem. 262, 4429-4432 (1987); and Wagner et al., Proc. Natl. Acad. Sci. USA 87, 3410-3414 (1990). For review of gene marking and gene therapy protocols see Anderson et al., Science 256, 808-813 (1992).

The nucleic acid molecules encoding the TAT polypeptides or fragments thereof described herein are useful for chromosome identification. In this regard, there exists an ongoing need to identify new chromosome markers, since relatively few chromosome marking reagents, based upon actual sequence data are presently available. Each TAT nucleic acid molecule of the present invention can be used as a chromosome marker.

The TAT polypeptides and nucleic acid molecules of the present invention may also be used diagnostically for tissue typing, wherein the TAT polypeptides of the present invention may be differentially expressed in one tissue as compared to another, preferably in a diseased tissue as compared to a normal tissue of the same tissue type. TAT nucleic acid molecules will find use for generating probes for PCR, Northern analysis, Southern analysis and Western analysis.

This invention encompasses methods of screening compounds to identify those that mimic the TAT polypeptide (agonists) or prevent the effect of the TAT polypeptide (antagonists). Screening assays for antagonist drug candidates are designed to identify compounds that bind or complex with the TAT polypeptides encoded by the genes identified herein, or otherwise interfere with the interaction of the encoded polypeptides with other cellular proteins, including e.g., inhibiting the expression of TAT polypeptide from cells. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates.

The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays, and cell-based assays, which are well characterized in the art.

All assays for antagonists are common in that they call for contacting the drug candidate with a TAT polypeptide encoded by a nucleic acid identified herein under conditions and for a time sufficient to allow these two components to interact.

In binding assays, the interaction is binding and the complex formed can be isolated or detected in the reaction mixture. In a particular embodiment, the TAT polypeptide encoded by the gene identified herein or the drug candidate is immobilized on a solid phase, e.g., on a microtiter plate, by covalent or non-covalent attachments.

Non-covalent attachment generally is accomplished by coating the solid surface with a solution of the TAT polypeptide and drying. Alternatively, an immobilized antibody, e.g., a monoclonal antibody, specific for the TAT polypeptide to be immobilized can be used to anchor it to a solid surface. The assay is performed by adding the non-immobilized component, which may be labeled by a detectable label, to the immobilized component, e.g., the coated surface containing the anchored component. When the reaction is complete, the non-reacted components are removed, e.g., by washing, and complexes anchored on the solid surface are detected. When the originally non-immobilized component carries a detectable label, the detection of label immobilized on the surface indicates that complexing occurred. Where the originally non-immobilized component does not carry a label, complexing can be detected, for example, by using a labeled antibody specifically binding the immobilized complex.

If the candidate compound interacts with but does not bind to a particular TAT polypeptide encoded by a gene identified herein, its interaction with that polypeptide can be assayed by methods well known for detecting protein-protein interactions. Such assays include traditional approaches, such as, e.g., cross-linking, co-immunoprecipitation, and co-purification through gradients or chromatographic columns. In addition, protein-protein interactions can be monitored by using a yeast-based genetic system described by Fields and co-workers (Fields and Song, Nature (London), 340:245-246 (1989); Chien et al., Proc. Natl. Acad. Sci. USA, 88:9578-9582 (1991)) as disclosed by Chevray and Nathans, Proc. Natl. Acad. Sci. USA, 89: 5789-5793 (1991). Many transcriptional activators, such as yeast GAL4, consist of two physically discrete modular domains, one acting as the DNA-binding domain, the other one functioning as the transcription-activation domain. The yeast expression system described in the foregoing publications (generally referred to as the "two-hybrid system") takes advantage of this property, and employs two hybrid proteins, one in which the target protein is fused to the DNA-binding domain of GAL4, and another, in which candidate activating proteins are fused to the activation domain. The expression of a GAL1-*lacZ* reporter gene under control of a GAL4-activated promoter depends on reconstitution of GAL4 activity via protein-protein interaction. Colonies containing interacting polypeptides are detected with a chromogenic substrate for β-galactosidase. A complete kit (MATCHMAKER™) for identifying protein-protein interactions between two specific proteins using the two-hybrid technique is commercially available from Clontech. This system can also be extended to map protein domains involved in specific protein interactions as well as to pinpoint amino acid residues that are crucial for these interactions.

Compounds that interfere with the interaction of a gene encoding a TAT polypeptide identified herein and other intra- or extracellular components can be tested as follows: usually a reaction mixture is prepared containing the product of the gene and the intra- or extracellular component under conditions and for a time allowing for the interaction and binding of the two products. To test the ability of a candidate compound to inhibit binding, the reaction is run in the absence and in the presence of the test compound. In addition, a placebo may be added to a third reaction mixture, to serve as positive control. The binding (complex formation) between the test compound and the intra- or extracellular component present in the mixture is monitored as described hereinabove. The formation of a complex in the control reaction(s) but not in the reaction mixture containing the test compound indicates that the test compound interferes with the interaction of the test compound and its reaction partner.

To assay for antagonists, the TAT polypeptide may be added to a cell along with the compound to be screened for a particular activity and the ability of the compound to inhibit the activity of interest in the presence of the TAT polypeptide indicates that the compound is an antagonist to the TAT polypeptide. Alternatively, antagonists may be detected by combining the TAT polypeptide and a potential antagonist with membrane-bound TAT polypeptide receptors or recombinant receptors under appropriate conditions for a competitive inhibition assay. The TAT polypeptide can be labeled, such as by radioactivity, such that the number of TAT polypeptide molecules bound to the receptor can be used to determine the effectiveness of the potential antagonist. The gene encoding the receptor can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Coligan et al., Current Protocols in Immun., 1(2): Chapter 5 (1991). Preferably, expression cloning is employed wherein polyadenylated RNA is prepared from a cell responsive to the TAT polypeptide and a cDNA library created from this RNA is divided into pools and used to transfect COS cells or other cells that are not responsive to the TAT polypeptide. Transfected cells that are grown on glass slides are exposed to labeled TAT polypeptide. The TAT polypeptide can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase. Following fixation and incubation, the slides are subjected to autoradiographic analysis. Positive pools are identified and sub-pools are prepared and re-transfected using an interactive sub-pooling and re-screening process, eventually yielding a single clone that encodes the putative receptor.

As an alternative approach for receptor identification, labeled TAT polypeptide can be photoaffinity-linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is resolved by PAGE and exposed to X-ray film. The labeled complex containing the receptor can be excised, resolved into peptide fragments, and subjected to protein micro-sequencing. The amino acid sequence obtained from micro- sequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the gene encoding the putative receptor.

In another assay for antagonists, mammalian cells or a membrane preparation expressing the receptor would be incubated with labeled TAT polypeptide in the presence of the candidate compound. The ability of the compound to enhance or block this interaction could then be measured.

More specific examples of potential antagonists include an oligonucleotide that binds to the fusions of immunoglobulin with TAT polypeptide, and, in particular, antibodies including, without limitation, poly- and monoclonal antibodies and antibody fragments, single-chain antibodies, anti-idiotypic antibodies, and chimeric or humanized versions of such antibodies or fragments, as well as human antibodies and antibody fragments. Alternatively, a potential antagonist may be a closely related protein, for example, a mutated form of the TAT polypeptide that recognizes the receptor but imparts no effect, thereby competitively inhibiting the action of the TAT polypeptide.

Another potential TAT polypeptide antagonist is an antisense RNA or DNA construct prepared using antisense technology, where, e.g., an antisense RNA or DNA molecule acts to block directly the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are

based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes the mature TAT polypeptides herein, is used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res., 6:3073 (1979); Cooney et al., Science, 241: 456 (1988); Dervan et al., Science, 251:1360 (1991)), thereby preventing transcription and the production of the TAT polypeptide. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into the TAT polypeptide (antisense - Okano, Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression (CRC Press: Boca Raton, FL, 1988). The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of the TAT polypeptide. When antisense DNA is used, oligodeoxyribonucleotides derived from the translation-initiation site, e.g., between about -10 and +10 positions of the target gene nucleotide sequence, are preferred.

Potential antagonists include small molecules that bind to the active site, the receptor binding site, or growth factor or other relevant binding site of the TAT polypeptide, thereby blocking the normal biological activity of the TAT polypeptide. Examples of small molecules include, but are not limited to, small peptides or peptide-like molecules, preferably soluble peptides, and synthetic non-peptidyl organic or inorganic compounds.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. Ribozymes act by sequence-specific hybridization to the complementary target RNA, followed by endonucleolytic cleavage. Specific ribozyme cleavage sites within a potential RNA target can be identified by known techniques. For further details see, e.g., Rossi, Current Biology, 4:469-471 (1994), and PCT publication No. WO 97/33551 (published September 18, 1997).

Nucleic acid molecules in triple-helix formation used to inhibit transcription should be single-stranded and composed of deoxynucleotides. The base composition of these oligonucleotides is designed such that it promotes triple-helix formation via Hoogsteen base-pairing rules, which generally require sizeable stretches of purines or pyrimidines on one strand of a duplex. For further details see, e.g., PCT publication No. WO 97/33551, *supra*.

These small molecules can be identified by any one or more of the screening assays discussed hereinabove and/or by any other screening techniques well known for those skilled in the art.

Isolated TAT polypeptide-encoding nucleic acid can be used herein for recombinantly producing TAT polypeptide using techniques well known in the art and as described herein. In turn, the produced TAT polypeptides can be employed for generating anti-TAT antibodies using techniques well known in the art and as described herein.

Antibodies specifically binding a TAT polypeptide identified herein, as well as other molecules identified by the screening assays disclosed hereinbefore, can be administered for the treatment of various disorders, including cancer, in the form of pharmaceutical compositions.

If the TAT polypeptide is intracellular and whole antibodies are used as inhibitors, internalizing antibodies are preferred. However, lipofections or liposomes can also be used to deliver the antibody, or an antibody

fragment, into cells. Where antibody fragments are used, the smallest inhibitory fragment that specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable-region sequences of an antibody, peptide molecules can be designed that retain the ability to bind the target protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology. See, *e.g.*, Marasco *et al.*, Proc. Natl. Acad. Sci. USA, 90: 7889-7893 (1993).

The formulation herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition may comprise an agent that enhances its function, such as, for example, a cytotoxic agent, cytokine, chemotherapeutic agent, or growth-inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Manassas, VA.

EXAMPLE 1: Tissue Expression Profiling Using GeneExpress®

A proprietary database containing gene expression information (GeneExpress®, Gene Logic Inc., Gaithersburg, MD) was analyzed in an attempt to identify polypeptides (and their encoding nucleic acids) whose expression is significantly upregulated in a particular tumor tissue(s) of interest as compared to other tumor(s) and/or normal tissues. Specifically, analysis of the GeneExpress® database was conducted using either software available through Gene Logic Inc., Gaithersburg, MD, for use with the GeneExpress® database or with proprietary software written and developed at Genentech, Inc. for use with the GeneExpress® database. The rating of positive hits in the analysis is based upon several criteria including, for example, tissue specificity, tumor specificity and expression level in normal essential and/or normal proliferating tissues. The following is a list of molecules whose tissue expression profile as determined from an analysis of the GeneExpress® database evidences high tissue expression and significant upregulation of expression in a specific tumor or tumors as compared to other tumor(s) and/or normal tissues and optionally relatively low expression in normal essential and/or normal proliferating tissues. As such, the molecules listed below are excellent polypeptide targets for the diagnosis and therapy of cancer in mammals.

Molecule	upregulation of expression in:	as compared to:

DNA77507 (TAT161)	breast tumor	normal breast tissue
DNA77507 (TAT161)	colon tumor	normal colon tissue
DNA77507 (TAT161)	lung tumor	normal lung tissue
DNA77507 (TAT161)	kidney tumor	normal kidney tissue
DNA77507 (TAT161)	liver tumor	normal liver tissue
DNA77507 (TAT161)	ovarian tumor	normal ovarian tissue
DNA77507 (TAT161)	pancreatic tumor	normal pancreatic tissue
DNA77507 (TAT161)	rectum tumor	normal rectum tissue
DNA77507 (TAT161)	skin tumor	normal skin tissue
DNA77507 (TAT161)	uterine tumor	normal uterine tissue
DNA77507 (TAT161)	brain tumor	normal brain tissue
DNA77507 (TAT161)	soft tissue tumor	normal soft tissue
DNA77507 (TAT161)	bone tumor	normal bone tissue
DNA80894 (TAT101)	breast tumor	normal breast tissue
DNA82343 (TAT157)	colon tumor	normal colon tissue
DNA82343 (TAT157)	ovarian tumor	normal ovarian tissue
DNA82343 (TAT157)	stomach tumor	normal stomach tissue
DNA82343 (TAT157)	liver tumor	normal liver tissue
DNA82343 (TAT157)	rectum tumor	normal rectum tissue
DNA82343 (TAT157)	small intestine tumor	normal small intestine tissue
DNA82343 (TAT157)	esophagus tumor	normal esophagus tissue
DNA82343 (TAT157)	testis tumor	normal testis tissue
DNA82343 (TAT157)	thymus tumor	normal thymus tissue
DNA87994 (TAT160)	breast tumor	normal breast tissue
DNA87994 (TAT160)	pancreatic tumor	normal pancreatic tissue
DNA87994 (TAT160)	rectum tumor	normal rectum tissue
DNA87994 (TAT160)	colon tumor	normal colon tissue
DNA87994 (TAT160)	esophagus tumor	normal esophagus tissue
DNA87994 (TAT160)	ovarian tumor	normal ovarian tissue
DNA87994 (TAT160)	lung tumor	normal lung tissue
DNA87994 (TAT160)	uterine tumor	normal uterine tissue
DNA88131 (TAT158)	bone tumor	normal bone tissue
DNA88131 (TAT158)	breast tumor	normal breast tissue
DNA88131 (TAT158)	colon tumor	normal colon tissue
DNA88131 (TAT158)	uterine tumor	normal uterine tissue
DNA88131 (TAT158)	esophagus tumor	normal esophagus tissue
DNA88131 (TAT158)	lung tumor	normal lung tissue

DNA88131 (TAT158)	ovarian tumor	normal ovarian tissue
DNA88131 (TAT158)	pancreatic tumor	normal pancreatic tissue
DNA88131 (TAT158)	prostate tumor	normal prostate tissue
DNA88131 (TAT158)	skin tumor	normal skin tissue
DNA88131 (TAT158)	soft tissue tumor	normal soft tissue
DNA88131 (TAT158)	stomach tumor	normal stomach tissue

Molecule	upregulation of expression in:	as compared to:
DNA88131 (TAT158)	rectum tumor	normal rectum tissue
DNA88131 (TAT158)	neuroendocrine tumor.	normal neuroendocrine tissue
DNA88131 (TAT158)	brain tumor	normal brain tissue
DNA95930 (TAT110)	colon tumor	normal colon tissue
DNA95930 (TAT110)	uterine tumor	normal uterine tissue
DNA95930 (TAT110)	endometrial tumor	normal endometrial tissue
DNA95930 (TAT110)	rectum tumor	normal rectum tissue
DNA95930 (TAT110)	ovarian tumor	normal ovarian tissue
DNA95930 (TAT110)	breast tumor	normal breast tissue
DNA95930 (TAT110)	lung tumor	normal lung tissue
DNA95930 (TAT110)	prostate tumor	normal prostate tissue
DNA95930-1 (TAT210)	colon tumor	normal colon tissue
DNA95930-1 (TAT210)	uterine tumor	normal uterine tissue
DNA95930-1 (TAT210)	endometrial tumor	normal endometrial tissue
DNA95930-1 (TAT210)	rectum tumor	normal rectum tissue
DNA95930-1 (TAT210)	ovarian tumor	normal ovarian tissue
DNA95930-1 (TAT210)	breast tumor	normal breast tissue
DNA95930-1 (TAT210)	lung tumor	normal lung tissue
DNA95930-1 (TAT210)	prostate tumor	normal prostate tissue
DNA96917 (TAT159)	pancreatic tumor	normal pancreatic tissue
DNA96917 (TAT159)	lung tumor	normal lung tissue
DNA96917 (TAT159)	liver tumor	normal liver tissue
DNA96930 (TAT112)	breast tumor	normal breast tissue
DNA96930 (TAT112)	colon tumor	normal colon tissue
DNA96930 (TAT112)	rectum tumor	normal rectum tissue
DNA96930 (TAT112)	uterine tumor	normal uterine tissue
DNA96930 (TAT112)	lung tumor	normal lung tissue
DNA96930 (TAT112)	ovarian tumor	normal ovarian tissue
DNA96930 (TAT112)	pancreatic tumor	normal pancreatic tissue

DNA96930 (TAT112)	stomach tumor	normal stomach tissue
DNA96936 (TAT147)	breast tumor	normal breast tissue
DNA96936 (TAT147)	colon tumor	normal colon tissue
DNA96936 (TAT147)	testis tumor	normal testis tissue
DNA96936 (TAT147)	ovarian tumor	normal ovarian tissue
DNA98565 (TAT145)	brain tumor	normal brain tissue
DNA98565 (TAT145)	glioma	normal glial tissue
DNA246435 (TAT152)	brain tumor	normal brain tissue
DNA246435 (TAT152)	glioma	normal glial tissue
DNA98591 (TAT162)	colon tumor	normal colon tissue
DNA98591 (TAT162)	rectum tumor	normal rectum tissue
DNA98591 (TAT162)	ovarian tumor	normal ovarian tissue
DNA98591 (TAT162)	pancreatic tumor	normal pancreatic tissue
DNA98591 (TAT162)	stomach tumor	normal stomach tissue
DNA108809 (TAT114)	colon tumor	normal colon tissue
DNA108809 (TAT114)	kidney tumor	normal kidney tissue
DNA119488 (TAT119)	colon tumor	normal colon tissue
DNA119488 (TAT119)	lung tumor	normal lung tissue
DNA119488 (TAT119)	rectum tumor	normal rectum tissue
DNA143493 (TAT103)	breast tumor	normal breast tissue
DNA167234 (TAT130)	prostate tumor	normal prostate tissue
DNA235621 (TAT166)	prostate tumor	normal prostate tissue
DNA235621 (TAT166)	liver tumor	normal liver tissue
DNA176766 (TAT132)	kidney tumor	normal kidney tissue
DNA176766 (TAT132)	ovarian tumor	normal ovarian tissue
DNA176766 (TAT132)	uterine tumor	normal uterine tissue

Molecule	upregulation of expression in:	as compared to:
DNA236463 (TAT150)	kidney tumor	normal kidney tissue
DNA236463 (TAT150)	ovarian tumor	normal ovarian tissue
DNA236463 (TAT150)	uterine tumor	normal uterine tissue
DNA181162 (TAT129)	prostate tumor	normal prostate tissue
DNA188221 (TAT111)	colon tumor	normal colon tissue
DNA188221 (TAT111)	endometrial tumor	normal endometrial tissue
DNA188221 (TAT111)	stomach tumor	normal stomach tissue
DNA233876 (TAT146)	colon tumor	normal colon tissue
DNA233876 (TAT146)	endometrial tumor	normal endometrial tissue

DNA233876 (TAT146)	stomach tumor	normal stomach tissue
DNA193891 (TAT148)	colon tumor	normal colon tissue
DNA248170 (TAT187)	colon tumor	normal colon tissue
DNA248170 (TAT187)	breast tumor	normal breast tissue
DNA194628 (TAT118)	kidney tumor	normal kidney tissue
DNA246415 (TAT167)	kidney tumor	normal kidney tissue
DNA215609 (TAT113)	colon tumor	normal colon tissue
DNA215609 (TAT113)	rectum tumor	normal rectum tissue
DNA220432 (TAT128)	prostate tumor	normal prostate tissue
DNA226094 (TAT164)	breast tumor	normal breast tissue
DNA226094 (TAT164)	brain tumor	normal brain tissue
DNA226094 (TAT164)	lung tumor	normal lung tissue
DNA226094 (TAT164)	skin tumor	normal skin tissue
DNA226165 (TAT122)	breast tumor	normal breast tissue
DNA226165 (TAT122)	endometrial tumor	normal endometrial tissue
DNA226165 (TAT122)	kidney tumor	normal kidney tissue
DNA226165 (TAT122)	lung tumor	normal lung tissue
DNA226165 (TAT122)	ovarian tumor	normal ovarian tissue
DNA226165 (TAT122)	colon tumor	normal colon tissue
DNA226165 (TAT122)	rectum tumor	normal rectum tissue
DNA226165 (TAT122)	skin tumor	normal skin tissue
DNA226165 (TAT122)	soft tissue tumor	normal soft tissue tissue
DNA226165 (TAT122)	bladder tumor	normal bladder tissue
DNA226237 (TAT117)	kidney tumor	normal kidney tissue
DNA246450 (TAT168)	kidney tumor	normal kidney tissue
DNA226456 (TAT144)	breast tumor	normal breast tissue
DNA226456 (TAT144)	colon tumor	normal colon tissue
DNA226456 (TAT144)	rectum tumor	normal rectum tissue
DNA226456 (TAT144)	endometrial tumor	normal endometrial tissue
DNA226456 (TAT144)	kidney tumor	normal kidney tissue
DNA226456 (TAT144)	lung tumor	normal lung tissue
DNA226456 (TAT144)	ovarian tumor	normal ovarian tissue
DNA226456 (TAT144)	skin tumor	normal skin tissue
DNA237637 (TAT188)	breast tumor	normal breast tissue
DNA237637 (TAT188)	colon tumor	normal colon tissue
DNA237637 (TAT188)	rectum tumor	normal rectum tissue
DNA237637 (TAT188)	endometrial tumor	normal endometrial tissue

DNA237637 (TAT188)	kidney tumor	normal kidney tissue
DNA237637 (TAT188)	lung tumor	normal lung tissue
DNA237637 (TAT188)	ovarian tumor	normal ovarian tissue
DNA237637 (TAT188)	skin tumor	normal skin tissue
DNA237637 (TAT188)	liver tumor	normal liver tissue
DNA237637 (TAT188)	lung tumor	normal lung tissue
DNA226539 (TAT126)	breast tumor	normal breast tissue
DNA226539 (TAT126)	colon tumor	normal colon tissue
DNA226539 (TAT126)	rectum tumor	normal rectum tissue

Molecule	upregulation of expression in:	as compared to:
DNA226539 (TAT126)	endometrial tumor	normal endometrial tissue
DNA226539 (TAT126)	lung tumor	normal lung tissue
DNA226539 (TAT126)	ovarian tumor	normal ovarian tissue
DNA226539 (TAT126)	pancreatic tumor	normal pancreatic tissue
DNA236511 (TAT151)	breast tumor	normal breast tissue
DNA236511 (TAT151)	colon tumor	normal colon tissue
DNA236511 (TAT151)	rectum tumor	normal rectum tissue
DNA236511 (TAT151)	endometrial tumor	normal endometrial tissue
DNA236511 (TAT151)	lung tumor	normal lung tissue
DNA236511 (TAT151)	ovarian tumor	normal ovarian tissue
DNA236511 (TAT151)	pancreatic tumor	normal pancreatic tissue
DNA226771 (TAT115)	breast tumor	normal breast tissue
DNA226771 (TAT115)	colon tumor	normal colon tissue
DNA227087 (TAT163)	breast tumor	normal breast tissue
DNA227087 (TAT163)	colon tumor	normal colon tissue
DNA227087 (TAT163)	rectum tumor	normal rectum tissue
DNA227087 (TAT163)	lung tumor	normal lung tissue
DNA227087 (TAT163)	ovarian tumor	normal ovarian tissue
DNA227087 (TAT163)	prostate tumor	normal prostate tissue
DNA227087 (TAT163)	endocrine tumor	normal endocrine tissue
DNA227087 (TAT163)	kidney tumor	normal kidney tissue
DNA227087 (TAT163)	liver tumor	normal liver tissue
DNA227087 (TAT163)	nervous system tumor	normal nervous system tissue
DNA227087 (TAT163)	pancreatic tumor	normal pancreatic tissue
DNA227087 (TAT163)	uterine tumor	normal uterine tissue
DNA227087 (TAT163)	small intestine tumor	normal small intestine tissue

DNA227087 (TAT163)	lymphoid tumor	normal lymphoid tissue
DNA266307 (TAT227)	breast tumor	normal breast tissue
DNA266307 (TAT227)	colon tumor	normal colon tissue
DNA266307 (TAT227)	rectum tumor	normal rectum tissue
DNA266307 (TAT227)	lung tumor	normal lung tissue
DNA266307 (TAT227)	ovarian tumor	normal ovarian tissue
DNA266307 (TAT227)	prostate tumor	normal prostate tissue
DNA266307 (TAT227)	endocrine tumor	normal endocrine tissue
DNA266307 (TAT227)	kidney tumor	normal kidney tissue
DNA266307 (TAT227)	liver tumor	normal liver tissue
DNA266307 (TAT227)	nervous system tumor	normal nervous system tissue
DNA266307 (TAT227)	pancreatic tumor	normal pancreatic tissue
DNA266307 (TAT227)	uterine tumor	normal uterine tissue
DNA266307 (TAT227)	small intestine tumor	normal small intestine tissue
DNA266307 (TAT227)	lymphoid tumor	normal lymphoid tissue
DNA266311 (TAT228)	breast tumor	normal breast tissue
DNA266311 (TAT228)	colon tumor	normal colon tissue
DNA266311 (TAT228)	rectum tumor	normal rectum tissue
DNA266311 (TAT228)	lung tumor	normal lung tissue
DNA266311 (TAT228)	ovarian tumor	normal ovarian tissue
DNA266311 (TAT228)	prostate tumor	normal prostate tissue
DNA266311 (TAT228)	endocrine tumor	normal endocrine tissue
DNA266311 (TAT228)	kidney tumor	normal kidney tissue
DNA266311 (TAT228)	liver tumor	normal liver tissue
DNA266311 (TAT228)	nervous system tumor	normal nervous system tissue
DNA266311 (TAT228)	pancreatic tumor	normal pancreatic tissue
DNA266311 (TAT228)	uterine tumor	normal uterine tissue
DNA266311 (TAT228)	small intestine tumor	normal small intestine tissue
DNA266311 (TAT228)	lymphoid tumor	normal lymphoid tissue

Molecule	upregulation of expression in:	as compared to:
DNA266312 (TAT229)	breast tumor	normal breast tissue
DNA266312 (TAT229)	colon tumor	normal colon tissue
DNA266312 (TAT229)	rectum tumor	normal rectum tissue
DNA266312 (TAT229)	lung tumor	normal lung tissue
DNA266312 (TAT229)	ovarian tumor	normal ovarian tissue
DNA266312 (TAT229)	prostate tumor	normal prostate tissue

DNA266312 (TAT229)	endocrine tumor	normal endocrine tissue
DNA266312 (TAT229)	kidney tumor	normal kidney tissue
DNA266312 (TAT229)	liver tumor	normal liver tissue
DNA266312 (TAT229)	nervous system tumor	normal nervous system tissue
DNA266312 (TAT229)	pancreatic tumor	normal pancreatic tissue
DNA266312 (TAT229)	uterine tumor	normal uterine tissue
DNA266312 (TAT229)	small intestine tumor	normal small intestine tissue
DNA266312 (TAT229)	lymphoid tumor	normal lymphoid tissue
DNA266313 (TAT230)	breast tumor	normal breast tissue
DNA266313 (TAT230)	colon tumor	normal colon tissue
DNA266313 (TAT230)	rectum tumor	normal rectum tissue
DNA266313 (TAT230)	lung tumor	normal lung tissue
DNA266313 (TAT230)	ovarian tumor	normal ovarian tissue
DNA266313 (TAT230)	prostate tumor	normal prostate tissue
DNA266313 (TAT230)	endocrine tumor	normal endocrine tissue
DNA266313 (TAT230)	kidney tumor	normal kidney tissue
DNA266313 (TAT230)	liver tumor	normal liver tissue
DNA266313 (TAT230)	nervous system tumor	normal nervous system tissue
DNA266313 (TAT230)	pancreatic tumor	normal pancreatic tissue
DNA266313 (TAT230)	uterine tumor	normal uterine tissue
DNA266313 (TAT230)	small intestine tumor	normal small intestine tissue
DNA266313 (TAT230)	lymphoid tumor	normal lymphoid tissue
DNA227224 (TAT121)	breast tumor	normal breast tissue
DNA227224 (TAT121)	colon tumor	normal colon tissue
DNA227224 (TAT121)	rectum tumor	normal rectum tissue
DNA227224 (TAT121)	endometrial tumor	normal endometrial tissue
DNA227224 (TAT121)	kidney tumor	normal kidney tissue
DNA227224 (TAT121)	lung tumor	normal lung tissue
DNA227224 (TAT121)	ovarian tumor	normal ovarian tissue
DNA227224 (TAT121)	skin tumor	normal skin tissue
DNA227224 (TAT121)	testis tumor	normal testis tissue
DNA227224 (TAT121)	bladder tumor	normal bladder tissue
DNA247486 (TAT183)	breast tumor	normal breast tissue
DNA247486 (TAT183)	colon tumor	normal colon tissue
DNA247486 (TAT183)	rectum tumor	normal rectum tissue
DNA247486 (TAT183)	endometrial tumor	normal endometrial tissue
DNA247486 (TAT183)	kidney tumor	normal kidney tissue

DNA247486 (TAT183)	lung tumor	normal lung tissue
DNA247486 (TAT183)	ovarian tumor	normal ovarian tissue
DNA247486 (TAT183)	skin tumor	normal skin tissue
DNA247486 (TAT183)	testis tumor	normal testis tissue
DNA247486 (TAT183)	bladder tumor	normal bladder tissue
DNA227800 (TAT131)	prostate tumor	normal prostate tissue
DNA228199 (TAT127)	breast tumor	normal breast tissue
DNA228199 (TAT127)	endometrial tumor	normal endometrial tissue
DNA228199 (TAT127)	ovarian tumor	normal ovarian tissue
DNA228199 (TAT127)	pancreatic tumor	normal pancreatic tissue
DNA228199 (TAT127)	lung tumor	normal lung tissue
DNA228201 (TAT116)	colon tumor	normal colon tissue

Molecule	upregulation of expression in:	as compared to:
DNA228201 (TAT116)	rectum tumor	normal rectum tissue
DNA247488 (TAT189)	colon tumor	normal colon tissue
DNA247488 (TAT189)	rectum tumor	normal rectum tissue
DNA236538 (TAT190)	colon tumor	normal colon tissue
DNA236538 (TAT190)	rectum tumor	normal rectum tissue
DNA247489 (TAT191)	colon tumor	normal colon tissue
DNA247489 (TAT191)	rectum tumor	normal rectum tissue
DNA228211 (TAT133)	uterine tumor	normal uterine tissue
DNA233937 (TAT186)	uterine tumor	normal uterine tissue
DNA233937 (TAT186)	ovarian tumor	normal ovarian tissue
DNA228994 (TAT124)	lung tumor	normal lung tissue
DNA228994 (TAT124)	ovarian tumor	normal ovarian tissue
DNA228994 (TAT124)	skin tumor	normal skin tissue
DNA228994 (TAT124)	breast tumor	normal breast tissue
DNA229410 (TAT105)	breast tumor	normal breast tissue
DNA229411 (TAT107)	breast tumor	normal breast tissue
DNA229413 (TAT108)	breast tumor	normal breast tissue
DNA229700 (TAT139)	breast tumor	normal breast tissue
DNA231312 (TAT143)	breast tumor	normal breast tissue
DNA231312 (TAT143)	colon tumor	normal colon tissue
DNA231542 (TAT100)	brain tumor	normal brain tissue
DNA231542 (TAT100)	glioma	normal glial tissue
DNA231542-1 (TAT284)	brain tumor	normal brain tissue

DNA231542-1 (TAT284)	glioma	normal glial tissue
DNA231542-2 (TAT285)	brain tumor	normal brain tissue
DNA231542-2 (TAT285)	glioma	normal glial tissue
DNA297393 (TAT285-1)	brain tumor	normal brain tissue
DNA297393 (TAT285-1)	glioma	normal glial tissue
DNA234833 (TAT149)	colon tumor	normal colon tissue
DNA268022 (TAT231)	colon tumor	normal colon tissue
DNA268022 (TAT231)	breast tumor	normal breast tissue
DNA268022 (TAT231)	ovarian tumor	normal ovarian tissue
DNA236246 (TAT153)	breast tumor	normal breast tissue
DNA236343 (TAT104)	breast tumor	normal breast tissue
DNA236493 (TAT141)	breast tumor	normal breast tissue
DNA236493 (TAT141)	glioblastoma tumor	normal glial tissue
DNA236534 (TAT102)	breast tumor	normal breast tissue
DNA236534 (TAT102)	colon tumor	normal colon tissue
DNA236534 (TAT102)	rectum tumor	normal rectum tissue
DNA236534 (TAT102)	cervical tumor	normal cervical tissue
DNA236534 (TAT102)	endometrial tumor	normal endometrial tissue
DNA236534 (TAT102)	lung tumor	normal lung tissue
DNA236534 (TAT102)	ovarian tumor	normal ovarian tissue
DNA236534 (TAT102)	pancreatic tumor	normal pancreatic tissue
DNA236534 (TAT102)	prostate tumor	normal prostate tissue
DNA236534 (TAT102)	stomach tumor	normal stomach tissue
DNA236534 (TAT102)	bladder tumor	normal bladder tissue
DNA246430 (TAT109)	breast tumor	normal breast tissue
DNA246430 (TAT109)	prostate tumor	normal prostate tissue
DNA247480 (TAT142)	breast tumor	normal breast tissue
DNA247480 (TAT142)	lung tumor	normal lung tissue
DNA264454 (TAT106)	breast tumor	normal breast tissue

EXAMPLE 2: Microarray Analysis to Detect Upregulation of TAT Polypeptides in Cancerous Tumors

Nucleic acid microarrays, often containing thousands of gene sequences, are useful for identifying differentially expressed genes in diseased tissues as compared to their normal counterparts. Using nucleic acid microarrays, test and control mRNA samples from test and control tissue samples are reverse transcribed and labeled to generate cDNA probes. The cDNA probes are then hybridized to an array of nucleic acids immobilized on a solid support. The array is configured such that the sequence and position of each member of the array is

known. For example, a selection of genes known to be expressed in certain disease states may be arrayed on a solid support. Hybridization of a labeled probe with a particular array member indicates that the sample from which the probe was derived expresses that gene. If the hybridization signal of a probe from a test (disease tissue) sample is greater than hybridization signal of a probe from a control (normal tissue) sample, the gene or genes overexpressed in the disease tissue are identified. The implication of this result is that an overexpressed protein in a diseased tissue is useful not only as a diagnostic marker for the presence of the disease condition, but also as a therapeutic target for treatment of the disease condition.

The methodology of hybridization of nucleic acids and microarray technology is well known in the art. In one example, the specific preparation of nucleic acids for hybridization and probes, slides, and hybridization conditions are all detailed in PCT Patent Application Serial No. PCT/US01/10482, filed on March 30, 2001 and which is herein incorporated by reference.

In the present example, cancerous tumors derived from various human tissues were studied for upregulated gene expression relative to cancerous tumors from different tissue types and/or non-cancerous human tissues in an attempt to identify those polypeptides which are overexpressed in a particular cancerous tumor(s). In certain experiments, cancerous human tumor tissue and non-cancerous human tumor tissue of the same tissue type (often from the same patient) were obtained and analyzed for TAT polypeptide expression. Additionally, cancerous human tumor tissue from any of a variety of different human tumors was obtained and compared to a "universal" epithelial control sample which was prepared by pooling non-cancerous human tissues of epithelial origin, including liver, kidney, and lung. mRNA isolated from the pooled tissues represents a mixture of expressed gene products from these different tissues. Microarray hybridization experiments using the pooled control samples generated a linear plot in a 2-color analysis. The slope of the line generated in a 2-color analysis was then used to normalize the ratios of (test:control detection) within each experiment. The normalized ratios from various experiments were then compared and used to identify clustering of gene expression. Thus, the pooled "universal control" sample not only allowed effective relative gene expression determinations in a simple 2-sample comparison, it also allowed multi-sample comparisons across several experiments.

In the present experiments, nucleic acid probes derived from the herein described TAT polypeptide-encoding nucleic acid sequences were used in the creation of the microarray and RNA from various tumor tissues were used for the hybridization thereto. Below is shown the results of these experiments, demonstrating that various TAT polypeptides of the present invention are significantly overexpressed in various human tumor tissues as compared to their normal counterpart tissue(s). Moreover, all of the molecules shown below are significantly overexpressed in their specific tumor tissue(s) as compared to in the "universal" epithelial control. As described above, these data demonstrate that the TAT polypeptides of the present invention are useful not only as diagnostic markers for the presence of one or more cancerous tumors, but also serve as therapeutic targets for the treatment of those tumors.

Molecule	upregulation of expression in:	as compared to:
DNA95930 (TAT110)	colon tumor	normal colon tissue
DNA95930 (TAT110)	lung tumor	normal lung tissue

DNA95930 (TAT110)	prostate tumor	normal prostate tissue
DNA95930 (TAT110)	endometrial tumor	normal endometrial tissue
DNA95930 (TAT110)	ovarian tumor	normal ovarian tissue
DNA95930-1 (TAT210)	colon tumor	normal colon tissue
DNA95930-1 (TAT210)	lung tumor	normal lung tissue
DNA95930-1 (TAT210)	prostate tumor	normal prostate tissue
DNA95930-1 (TAT210)	endometrial tumor	normal endometrial tissue
DNA95930-1 (TAT210)	ovarian tumor	normal ovarian tissue
DNA96930 (TAT112)	colon tumor	normal colon tissue
DNA96930 (TAT112)	breast tumor	normal breast tissue
DNA96930 (TAT112)	lung tumor	normal lung tissue
DNA96936 (TAT147)	breast tumor	normal breast tissue
DNA96936 (TAT147)	colon tumor	normal colon tissue
DNA96936 (TAT147)	ovarian tumor	normal ovarian tissue
DNA96936 (TAT147)	prostate tumor	normal prostate tissue
DNA108809 (TAT114)	colon tumor	normal colon tissue
DNA119488 (TAT119)	colon tumor	normal colon tissue
DNA119488 (TAT119)	lung tumor	normal lung tissue
DNA143493 (TAT103)	breast tumor	normal breast tissue
DNA181162 (TAT129)	prostate tumor	normal prostate tissue
DNA188221 (TAT111)	colon tumor	normal colon tissue
DNA188221 (TAT111)	lung tumor	normal lung tissue
DNA188221 (TAT111)	ovarian tumor	normal ovarian tissue
DNA233876 (TAT146)	colon tumor	normal colon tissue
DNA233876 (TAT146)	lung tumor	normal lung tissue
DNA233876 (TAT146)	ovarian tumor	normal ovarian tissue
DNA210499 (TAT123)	ovarian tumor	normal ovarian tissue
DNA210499 (TAT123)	lung tumor	normal lung tissue
DNA219894 (TAT211)	ovarian tumor	normal ovarian tissue
DNA219894 (TAT211)	lung tumor	normal lung tissue
DNA215609 (TAT113)	colon tumor	normal colon tissue
DNA220432 (TAT128)	prostate tumor	normal prostate tissue
DNA226165 (TAT122)	breast tumor	normal breast tissue
DNA226165 (TAT122)	colon tumor	normal colon tissue
DNA226165 (TAT122)	rectum tumor	normal rectum tissue
DNA226165 (TAT122)	lung tumor	normal lung tissue
DNA226165 (TAT122)	ovarian tumor	normal ovarian tissue

DNA226165 (TAT122)	prostate tumor	normal prostate tissue
DNA226456 (TAT144)	breast tumor	normal breast tissue
DNA226456 (TAT144)	colon tumor	normal colon tissue
DNA237637 (TAT188)	breast tumor	normal breast tissue
DNA237637 (TAT188)	colon tumor	normal colon tissue
DNA226539 (TAT126)	rectum tumor	normal rectum tissue
DNA226539 (TAT126)	colon tumor	normal colon tissue
DNA226539 (TAT126)	lung tumor	normal lung tissue
DNA226539 (TAT126)	ovarian tumor	normal ovarian tissue
DNA236511 (TAT151)	rectum tumor	normal rectum tissue
DNA236511 (TAT151)	colon tumor	normal colon tissue
DNA236511 (TAT151)	lung tumor	normal lung tissue

Molecule	upregulation of expression in:	as compared to:
DNA236511 (TAT151)	ovarian tumor	normal ovarian tissue
DNA226771 (TAT115)	colon tumor	normal colon tissue
DNA227224 (TAT121)	ovarian tumor	normal ovarian tissue
DNA227224 (TAT121)	rectum tumor	normal rectum tissue
DNA227224 (TAT121)	colon tumor	normal colon tissue
DNA227224 (TAT121)	lung tumor	normal lung tissue
DNA227224 (TAT121)	breast tumor	normal breast tissue
DNA227224 (TAT121)	prostate tumor	normal prostate tissue
DNA247486 (TAT183)	ovarian tumor	normal ovarian tissue
DNA247486 (TAT183)	rectum tumor	normal rectum tissue
DNA247486 (TAT183)	colon tumor	normal colon tissue
DNA247486 (TAT183)	lung tumor	normal lung tissue
DNA247486 (TAT183)	breast tumor	normal breast tissue
DNA247486 (TAT183)	prostate tumor	normal prostate tissue
DNA228199 (TAT127)	ovarian tumor	normal ovarian tissue
DNA228199 (TAT127)	lung tumor	normal lung tissue
DNA228201 (TAT116)	colon tumor	normal colon tissue
DNA247488 (TAT189)	colon tumor	normal colon tissue
DNA236538 (TAT190)	colon tumor	normal colon tissue
DNA247489 (TAT191)	colon tumor	normal colon tissue
DNA228994 (TAT124)	lung tumor	normal lung tissue
DNA228994 (TAT124)	breast tumor	normal breast tissue
DNA228994 (TAT124)	ovarian tumor	normal ovarian tissue

DNA231312 (TAT143)	colon tumor	normal colon tissue
DNA231542 (TAT100)	brain tumor	normal brain tissue
DNA231542 (TAT100)	glioma	normal glial tissue
DNA231542-1 (TAT284)	brain tumor	normal brain tissue
DNA231542-1 (TAT284)	glioma	normal glial tissue
DNA231542-2 (TAT285)	brain tumor	normal brain tissue
DNA231542-2 (TAT285)	glioma	normal glial tissue
DNA297393 (TAT285-1)	brain tumor	normal brain tissue
DNA297393 (TAT285-1)	glioma	normal glial tissue
DNA236246 (TAT153)	breast tumor	normal breast tissue
DNA236343 (TAT104)	breast tumor	normal breast tissue
DNA236534 (TAT102)	breast tumor	normal breast tissue
DNA236534 (TAT102)	colon tumor	normal colon tissue
DNA246430 (TAT109)	prostate tumor	normal prostate tissue
DNA264454 (TAT106)	breast tumor	normal breast tissue
DNA98565 (TAT145)	glioma	normal brain tissue
DNA246435 (TAT152)	glioma	normal brain tissue
DNA226094 (TAT164)	glioma	normal brain tissue

EXAMPLE 3: Quantitative Analysis of TAT mRNA Expression

In this assay, a 5' nuclease assay (for example, TaqMan®) and real-time quantitative PCR (for example, ABI Prizm 7700 Sequence Detection System® (Perkin Elmer, Applied Biosystems Division, Foster City, CA)), were used to find genes that are significantly overexpressed in a cancerous tumor or tumors as compared to other cancerous tumors or normal non-cancerous tissue. The 5' nuclease assay reaction is a fluorescent PCR-based technique which makes use of the 5' exonuclease activity of Taq DNA polymerase enzyme to monitor gene expression in real time. Two oligonucleotide primers (whose sequences are based upon the gene or EST sequence of interest) are used to generate an amplicon typical of a PCR reaction. A third oligonucleotide, or probe, is designed to detect nucleotide sequence located between the two PCR primers. The probe is non-extendible by Taq DNA polymerase enzyme, and is labeled with a reporter fluorescent dye and a quencher fluorescent dye. Any laser-induced emission from the reporter dye is quenched by the quenching dye when the two dyes are located close together as they are on the probe. During the PCR amplification reaction, the Taq DNA polymerase enzyme cleaves the probe in a template-dependent manner. The resultant probe fragments disassociate in solution, and signal from the released reporter dye is free from the quenching effect of the second fluorophore. One molecule of reporter dye is liberated for each new molecule synthesized, and detection of the unquenched reporter dye provides the basis for quantitative interpretation of the data.

The 5' nuclease procedure is run on a real-time quantitative PCR device such as the ABI Prism 7700TM Sequence Detection. The system consists of a thermocycler, laser, charge-coupled device (CCD) camera and

computer. The system amplifies samples in a 96-well format on a thermocycler. During amplification, laser-induced fluorescent signal is collected in real-time through fiber optics cables for all 96 wells, and detected at the CCD. The system includes software for running the instrument and for analyzing the data.

The starting material for the screen was mRNA isolated from a variety of different cancerous tissues. The mRNA is quantitated precisely, e.g., fluorometrically. As a negative control, RNA was isolated from various normal tissues of the same tissue type as the cancerous tissues being tested.

5' nuclease assay data are initially expressed as Ct, or the threshold cycle. This is defined as the cycle at which the reporter signal accumulates above the background level of fluorescence. The ΔCt values are used as quantitative measurement of the relative number of starting copies of a particular target sequence in a nucleic acid sample when comparing cancer mRNA results to normal human mRNA results. As one Ct unit corresponds to 1 PCR cycle or approximately a 2-fold relative increase relative to normal, two units corresponds to a 4-fold relative increase, 3 units corresponds to an 8-fold relative increase and so on, one can quantitatively measure the relative fold increase in mRNA expression between two or more different tissues. Using this technique, the molecules listed below have been identified as being significantly overexpressed in a particular tumor(s) as compared to their normal non-cancerous counterpart tissue(s) (from both the same and different tissue donors) and thus, represent excellent polypeptide targets for the diagnosis and therapy of cancer in mammals.

Molecule	upregulation of expression in:	as compared to:
DNA77507 (TAT161)	breast tumor	normal breast tissue
DNA82343 (TAT157)	colon tumor	normal colon tissue
DNA88131 (TAT158)	breast tumor	normal breast tissue
DNA88131 (TAT158)	colon tumor	normal colon tissue
DNA95930 (TAT110)	colon tumor	normal colon tissue
DNA95930 (TAT110)	lung tumor	normal lung tissue
DNA95930 (TAT110)	prostate tumor	normal prostate tissue
DNA95930 (TAT110)	endometrial tumor	normal endometrial tissue
DNA95930 (TAT110)	ovarian tumor	normal ovarian tissue
DNA95930-1 (TAT210)	colon tumor	normal colon tissue
DNA95930-1 (TAT210)	lung tumor	normal lung tissue
DNA95930-1 (TAT210)	prostate tumor	normal prostate tissue
DNA95930-1 (TAT210)	endometrial tumor	normal endometrial tissue

Molecule	upregulation of expression in:	as compared to:
DNA95930-1 (TAT210)	ovarian tumor	normal ovarian tissue
DNA96930 (TAT112)	colon tumor	normal colon tissue
DNA96936 (TAT147)	colon tumor	normal colon tissue
DNA98591 (TAT162)	colon tumor	normal colon tissue
DNA108809 (TAT114)	kidney tumor	normal kidney tissue

DNA119488 (TAT119)	lung tumor	normal lung tissue
DNA188221 (TAT111)	colon tumor	normal colon tissue
DNA233876 (TAT146)	colon tumor	normal colon tissue
DNA193891 (TAT148)	colon tumor	normal colon tissue
DNA248170 (TAT187)	colon tumor	normal colon tissue
DNA194628 (TAT118)	kidney tumor	normal kidney tissue
DNA246415 (TAT167)	kidney tumor	normal kidney tissue
DNA210499 (TAT123)	lung tumor	normal lung tissue
DNA219894 (TAT211)	lung tumor	normal lung tissue
DNA215609 (TAT113)	colon tumor	normal colon tissue
DNA220432 (TAT128)	prostate tumor	normal prostate tissue
DNA226165 (TAT122)	lung tumor	normal lung tissue
DNA226237 (TAT117)	kidney tumor	normal kidney tissue
DNA246450 (TAT168)	kidney tumor	normal kidney tissue
DNA226456 (TAT144)	breast tumor	normal breast tissue
DNA237637 (TAT188)	breast tumor	normal breast tissue
DNA226539 (TAT126)	ovarian tumor	normal ovarian tissue
DNA236511 (TAT151)	ovarian tumor	normal ovarian tissue
DNA227224 (TAT121)	lung tumor	normal lung tissue
DNA247486 (TAT183)	lung tumor	normal lung tissue
DNA227800 (TAT131)	prostate tumor	normal prostate tissue
DNA228199 (TAT127)	ovarian tumor	normal ovarian tissue
DNA228199 (TAT127)	lung tumor	normal lung tissue
DNA228201 (TAT116)	colon tumor	normal colon tissue
DNA247488 (TAT189)	colon tumor	normal colon tissue
DNA236538 (TAT190)	colon tumor	normal colon tissue
DNA247489 (TAT191)	colon tumor	normal colon tissue
DNA228993 (TAT120)	lung tumor	normal lung tissue
DNA228994 (TAT124)	lung tumor	normal lung tissue
DNA236343 (TAT104)	breast tumor	normal breast tissue
DNA236534 (TAT102)	ovarian tumor	normal ovarian tissue
DNA246430 (TAT109)	breast tumor	normal breast tissue
DNA247480 (TAT142)	lung tumor	normal lung tissue
DNA98565 (TAT145)	glioma	normal brain tissue
DNA246435 (TAT152)	glioma	normal brain tissue
DNA226094 (TAT164)	glioma	normal brain tissue
DNA227578 (TAT165)	glioma	normal brain tissue

DNA231542 (TAT100)	glioma	normal brain tissue
DNA231542-1 (TAT284)	glioma	normal brain tissue
DNA231542-2 (TAT285)	glioma	normal brain tissue
DNA297393 (TAT285-1)	glioma	normal brain tissue

EXAMPLE 4: *In situ* Hybridization

In situ hybridization is a powerful and versatile technique for the detection and localization of nucleic acid sequences within cell or tissue preparations. It may be useful, for example, to identify sites of gene expression, analyze the tissue distribution of transcription, identify and localize viral infection, follow changes in specific mRNA synthesis and aid in chromosome mapping.

In situ hybridization was performed following an optimized version of the protocol by Lu and Gillett, Cell Vision 1:169-176 (1994), using PCR-generated ^{33}P-labeled riboprobes having homology to the target sequence to be detected. Briefly, formalin-fixed, paraffin-embedded human tissues were sectioned, deparaffinized, deproteinated in proteinase K (20 g/ml) for 15 minutes at 37°C, and further processed for *in situ* hybridization as described by Lu and Gillett, *supra*. A [^{33}-P] UTP-labeled antisense riboprobe was generated from a PCR product and hybridized at 55°C overnight. The slides were dipped in Kodak NTB2 nuclear track emulsion and exposed for 4 weeks.

^{33}P-Riboprobe synthesis

6.0 µl (125 mCi) of ^{33}P-UTP (Amersham BF 1002, SA<2000 Ci/mmol) were speed vac dried. To each tube containing dried ^{33}P-UTP, the following ingredients were added:

2.0 µl 5x transcription buffer

1.0 µl DTT (100 mM)

2.0 µl NTP mix (2.5 mM : 10 µ; each of 10 mM GTP, CTP & ATP + 10 µl H_2O)

1.0 µl UTP (50 µM)

1.0 µl Rnasin

1.0 µl DNA template (1µg)

1.0 µl H_2O

1.0 µl RNA polymerase (for PCR products T3 = AS, T7 = S, usually)

The tubes were incubated at 37°C for one hour. 1.0 µl RQ1 DNase were added, followed by incubation at 37°C for 15 minutes. 90 µl TE (10 mM Tris pH 7.6/1mM EDTA pH 8.0) were added, and the mixture was pipetted onto DE81 paper. The remaining solution was loaded in a Microcon-50 ultrafiltration unit, and spun using program 10 (6 minutes). The filtration unit was inverted over a second tube and spun using program 2 (3 minutes). After the final recovery spin, 100 µl TE were added. 1 µl of the final product was pipetted on DE81 paper and counted in 6 ml of Biofluor II.

The probe was run on a TBE/urea gel. 1-3 µl of the probe or 5 µl of RNA Mrk III were added to 3 µl of loading buffer. After heating on a 95°C heat block for three minutes, the probe was immediately placed on ice.

The wells of gel were flushed, the sample loaded, and run at 180-250 volts for 45 minutes. The gel was wrapped in saran wrap and exposed to XAR film with an intensifying screen in -70°C freezer one hour to overnight.

^{33}P-Hybridization

A. Pretreatment of frozen sections

The slides were removed from the freezer, placed on aluminium trays and thawed at room temperature for 5 minutes. The trays were placed in 55°C incubator for five minutes to reduce condensation. The slides were fixed for 10 minutes in 4% paraformaldehyde on ice in the fume hood, and washed in 0.5 x SSC for 5 minutes, at room temperature (25 ml 20 x SSC + 975 ml SQ H $_2$O). After deproteination in 0.5 µg/ml proteinase K for 10 minutes at 37°C (12.5 µl of 10 mg/ml stock in 250 ml prewarmed RNase-free RNAse buffer), the sections were washed in 0.5 x SSC for 10 minutes at room temperature. The sections were dehydrated in 70%, 95%, 100% ethanol, 2 minutes each.

B. Pretreatment of paraffin-embedded sections

The slides were deparaffinized, placed in SQ H_2O, and rinsed twice in 2 x SSC at room temperature, for 5 minutes each time. The sections were deproteinated in 20 µg/ml proteinase K (500µl of 10 mg/ml in 250 ml RNase-free RNase buffer; 37°C, 15 minutes) - human embryo, or 8 x proteinase K (100 µl in 250 ml Rnase buffer, 37°C, 30 minutes) - formalin tissues. Subsequent rinsing in 0.5 x SSC and dehydration were performed as described above.

C. Prehybridization

The slides were laid out in a plastic box lined with Box buffer (4 x SSC, 50% formamide) - saturated filter paper.

D. Hybridization

1.0 x 10^6 cpm probe and 1.0 µl tRNA (50 mg/ml stock) per slide were heated at 95°C for 3 minutes. The slides were cooled on ice, and 48 µl hybridization buffer were added per slide. After vortexing, 50µl ^{33}P mix were added to 50 µl prehybridization on slide. The slides were incubated overnight at 55°C.

E. Washes

Washing was done 2 x 10 minutes with 2xSSC, EDTA at room temperature (400 ml 20 x SSC + 16 ml 0.25M EDTA, V_f=4L), followed by RNaseA treatment at 37°C for 30 minutes (500 µl of 10 mg/ml in 250 ml Rnase buffer = 20 µg/ml), The slides were washed 2 x 10 minutes with 2 x SSC, EDTA at room temperature. The stringency wash conditions were as follows: 2 hours at 55°C, 0.1 x SSC, EDTA (20 ml 20 x SSC + 16 ml EDTA, V_f=4L).

F. Oligonucleotides

In situ analysis was performed on a variety of DNA sequences disclosed herein. The oligonucleotides employed for these analyses were obtained so as to be complementary to the nucleic acids (or the complements thereof) as shown in the accompanying figures.

G. Results

In situ analysis was performed on a variety of DNA sequences disclosed herein. The results from these analyses are as follows.

(1) DNA95930 (TAT110)

In one analysis, significant expression is observed in 3/3 lung tumors, 3/3 colorectal adenocarcinomas, 1/1 prostate cancers, 3/3 transitional cell carcinomas and 3/3 endometrial adenocarcinomas, wherein the level of expression in the counterpart normal tissues is significantly less.

In a second independent analysis, significant expression is observed in 7/7 endometrial and 12/15 ovarian adenocarcinomas, wherein the level of expression in the counterpart normal tissues is significantly less.

In a third independent analysis, significant expression is observed in 24/26 colorectal tumor samples, wherein the level of expression in the counterpart normal tissue is significantly less.

Finally, in a fourth independent analysis, expression is observed in 8/26 samples of non-malignant prostate tissue, 55/82 samples of primary prostate cancer and in 5/23 samples of metastatic prostate cancer.

(2) DNA95930-1 (TAT210)

In one analysis, significant expression is observed in 3/3 lung tumors, 3/3 colorectal adenocarcinomas, 1/1 prostate cancers, 3/3 transitional cell carcinomas and 3/3 endometrial adenocarcinomas, wherein the level of expression in the counterpart normal tissues is significantly less.

In a second independent analysis, significant expression is observed in 7/7 endometrial and 12/15 ovarian adenocarcinomas, wherein the level of expression in the counterpart normal tissues is significantly less.

In a third independent analysis, significant expression is observed in 24/26 colorectal tumor samples, wherein the level of expression in the counterpart normal tissue is significantly less.

Finally, in a fourth independent analysis, expression is observed in 8/26 samples of non-malignant prostate tissue, 55/82 samples of primary prostate cancer and in 5/23 samples of metastatic prostate cancer.

(3) DNA96930 (TAT112)

Strong expression in colorectal cancers. Expression in the malignant epithelium appears significantly stronger than in adjacent benign epithelium. Additionally, strong expression is observed in all 23 of 23 samples of pancreatic adenocarcinoma tested, wherein expression in normal pancreatic tissue is not detectable.

(4) DNA96936 (TAT147)

In one analysis, a strongly positive signal was observed in 6/6 breast tumors. In another independent analysis, a positive signal was observed in 4/4 non small cell lung carcinomas, wherein the tumors appear to have stronger expression compared with normal lung. 1/1 endometrial adenocarcinomas shows strong expression and 3/3 colorectal adenocarcinomas show variable expression.

(5) DNA108809 (TAT114)

Positive signal in all renal cell carcinomas tested (n=3) while no expression observed in normal kidney tissue. Additionally, positive expression is observed in 5/12 stomach tumors, 5/24 colorectal tumors, 3/8 pancreatic tumors and 1/3 lung tumors. Normal non-cancerous tissue expression is limited to stomach and small intestine.

(6) DNA176766 (TAT132)

Positive signal in all endometrial adenocarcinomas tested (n=3) while no expression observed in normal endometrial tissue.

(7) DNA236463 (TAT150)

Positive signal in all endometrial adenocarcinomas tested (n=3) while no expression observed in normal endometrial tissue.

(8) DNA181162 (TAT129)

Neoplastic prostate epithelia are generally positive, with signal intensities varying from weak to strong between cases. Non-prostatic tissues are negative.

(9) DNA188221 (TAT111)

Strong signal seen in colonic multi-tumor array over malignant epithelium. In normal tissues, a certain probe gave specific signal over epithelial cells lining the lower 2/3 of the colonic crypts, the intensity of signal appeared significantly lower than in the colonic carcinomas. Positive expression is observed in 12/18 colorectal adenocarcinomas, 6/8 metastatic adenocarcinomas and 2/9 gastric adenocarcinomas.

(10) DNA233876 (TAT146)

Strong signal seen in colonic multi-tumor array over malignant epithelium. In normal tissues, a certain probe gave specific signal over epithelial cells lining the lower 2/3 of the colonic crypts, the intensity of signal appeared significantly lower than in the colonic carcinomas. Positive expression is observed in 12/18 colorectal adenocarcinomas, 6/8 metastatic adenocarcinomas and 2/9 gastric adenocarcinomas.

(11) DNA210499 (TAT123)

In one analysis, 12/14 ovarian adenocarcinomas are positive and 8/9 endometrial adenocarcinomas are positive. Normal ovarian stroma is negative as is uterine myometrium. Other normal ovarian and uterine tissues are negative.

In an independent analysis, 16/27 non small cell lung carcinomas are positive, wherein the signal is moderate or strong.

(12) DNA219894 (TAT211)

In one analysis, 12/14 ovarian adenocarcinomas are positive and 8/9 endometrial adenocarcinomas are positive. Normal ovarian stroma is negative as is uterine myometrium. Other normal ovarian and uterine tissues are negative.

In an independent analysis, 16/27 non small cell lung carcinomas are positive, wherein the signal is moderate or strong.

(13) DNA215609 (TAT113)

Strong signal seen in colonic carcinomas, with only very low level signal in normal colon. Lung and breast carcinomas were negative.

(14) DNA220432 (TAT128)

The only normal adult tissue expressing this gene is prostatic epithelium. The expression is of moderate to strong intensity and focal, it is more prevalent in hyperplastic epithelium.

In one analysis where 50 cases of primary prostate cancer are available for review, 29 cases (58%) are positive, 18 cases (36%) are negative and 3 cases (6%) are equivocal. In another analysis where 37 cases of primary prostate cancer are available for review, 33 cases (89%) are positive, 4 cases (11%) are negative. Finally, in another independent analysis where 27 cases of metastatic prostate cancer are available for review, 14 cases (52%) are positive, 11 cases (41%) are negative and 2 cases (7%) are equivocal.

(15) DNA226237 (TAT117)

In one analysis, two of 3 renal cell carcinomas are positive, wherein normal kidney expression is negative.

(16) DNA246450 (TAT168)

In one analysis, two of 3 renal cell carcinomas are positive, wherein normal kidney expression is negative.

(17) DNA227087 (TAT163)

A probe for this molecule showed a positive signal in a subpopulation of tumor-associated stromal cells in all tested cases of lung, breast, colon, pancreatic and endometrial carcinomas. The intensity of the labeling was often quite strong. In a case of colon adenocarcinoma with adjacent benign colon, labeling was restricted to the tumor-associated stroma and the normal benign tissue was negative. A breast fibroadenoma also showed labeling of subepithelial stromal cells.

(18) DNA266307 (TAT227)

A probe for this molecule showed a positive signal in a subpopulation of tumor-associated stromal cells in all tested cases of lung, breast, colon, pancreatic and endometrial carcinomas. The intensity of the labeling was often quite strong. In a case of colon adenocarcinoma with adjacent benign colon, labeling was restricted to the tumor-associated stroma and the normal benign tissue was negative. A breast fibroadenoma also showed labeling of subepithelial stromal cells.

(19) DNA266311 (TAT228)

A probe for this molecule showed a positive signal in a subpopulation of tumor-associated stromal cells in all tested cases of lung, breast, colon, pancreatic and endometrial carcinomas. The intensity of the labeling was often quite strong. In a case of colon adenocarcinoma with adjacent benign colon, labeling was restricted to the tumor-associated stroma and the normal benign tissue was negative. A breast fibroadenoma also showed labeling of subepithelial stromal cells.

(20) DNA266312 (TAT229)

A probe for this molecule showed a positive signal in a subpopulation of tumor-associated stromal cells in all tested cases of lung, breast, colon, pancreatic and endometrial carcinomas. The intensity of the labeling was often quite strong. In a case of colon adenocarcinoma with adjacent benign colon, labeling was restricted to the tumor-associated stroma and the normal benign tissue was negative. A breast fibroadenoma also showed labeling of subepithelial stromal cells.

(21) DNA266313 (TAT230)

A probe for this molecule showed a positive signal in a subpopulation of tumor-associated stromal cells in all tested cases of lung, breast, colon, pancreatic and endometrial carcinomas. The intensity of the labeling was often quite strong. In a case of colon adenocarcinoma with adjacent benign colon, labeling was restricted to the

tumor-associated stroma and the normal benign tissue was negative. A breast fibroadenoma also showed labeling of subepithelial stromal cells.

(22) DNA227224 (TAT121)

Expression is seen in 2 of 3 endometrial adenocarcinomas.

(23) DNA247486 (TAT183)

Expression is seen in 2 of 3 endometrial adenocarcinomas.

(24) DNA227800 (TAT131)

In one analysis, 46/64 primary prostate cancers are positive and 6/14 metastatic prostate cancers are positive. Weak to moderate expression is seen in prostate epithelium

(25) DNA228199 (TAT127)

Expression is observed in 13 of 15 ovarian tumors (adenocarcinoma and surface epithelial tumors). Benign ovarian surface epithelium is also positive. The expression level in most positive tumors is strong or moderate and fairly uniform. Expression is also observed in 8 of 9 uterine adenocarcinomas. Seven of 23 non small cell lung carcinomas are positive.

(26) DNA228201 (TAT116)

The malignant cells of 13/16 colorectal adenocarcinomas are positive for TAT116 expression. Additionally, 9/10 metastatic adenocarcinomas are positive for expression. Expression is also observed in the basal portions of normal colonic crypts.

(27) DNA247488 (TAT189)

The malignant cells of 13/16 colorectal adenocarcinomas are positive for TAT189 expression. Additionally, 9/10 metastatic adenocarcinomas are positive for expression. Expression is also observed in the basal portions of normal colonic crypts.

(28) DNA236538 (TAT190)

The malignant cells of 13/16 colorectal adenocarcinomas are positive for TAT190 expression. Additionally, 9/10 metastatic adenocarcinomas are positive for expression. Expression is also observed in the basal portions of normal colonic crypts.

(29) DNA247489 (TAT191)

The malignant cells of 13/16 colorectal adenocarcinomas are positive for TAT191 expression. Additionally, 9/10 metastatic adenocarcinomas are positive for expression. Expression is also observed in the basal portions of normal colonic crypts.

(30) DNA228994 (TAT124)

Thirteen of 61 cass of non small cell lung carcinoma are positive for expression of TAT124. Expression level in these positive tumor samples is significantly higher than in normal adult tissues.

(31) DNA231542 (TAT100)

In situ analysis performed as described above evidences significantly upregulated expression in human glioma and glioblastoma tissues as compared to normal brain (and other) tissue.

(32) DNA231542-1 (TAT284)

In situ analysis performed as described above evidences significantly upregulated expression in human glioma and glioblastoma tissues as compared to normal brain (and other) tissue.

(33) DNA231542-2 (TAT285)

In situ analysis performed as described above evidences significantly upregulated expression in human glioma and glioblastoma tissues as compared to normal brain (and other) tissue.

(34) DNA297393 (TAT285-1)

In situ analysis performed as described above evidences significantly upregulated expression in human glioma and glioblastoma tissues as compared to normal brain (and other) tissue.

(35) DNA236534 (TAT102)

Expression of TAT102 is seen in 14 of 15 ovarian epithelial malignancies (adenocarcinoma, epithelial surface tumors, endometrioid Ca). Also, 8 of 9 endometrial adenocarcinomas of the uterus express TAT102. Moreover, expression of TAT102 is seen in 24 of 27 non-small cell lung cancers, positive cases include squamous and adenocarcinomas. Expression in these tumor tissues is significantly higher than in their normal tissue counterparts.

(36) DNA246430 (TAT109)

Fourteen of 92 breast tumor samples are positive for TAT109 expression. Expression in all normal tissues is undetectable.

(37) DNA264454 (TAT106)

Expression of TAT106 is observed in 38/88 breast tumors. Expression in normal breast tissue is weak or undetectable.

(38) DNA98565 (TAT145)

Positive signal for TAT145 was observed in most gliomas, glioblastomas, some melanomas, and normal brain (primarily localized to astrocytes). The signal intensity in the glioblastomas appeared to be greater than that in normal astrocytes. While the majority of glioma and glioblastoma samples tested were positive for TAT145 expression, the majority of normal brain samples tested were negative for such expression.

(39) DNA246435 (TAT152)

Positive signal for TAT152 was observed in most glioblastomas, some melanomas, and normal brain (primarily localized to astrocytes). The signal intensity in the glioblastomas appeared to be greater than that in normal astrocytes. While the majority of glioma and glioblastoma samples tested were positive for TAT152 expression, the majority of normal brain samples tested were negative for such expression.

(40) DNA167234 (TAT130)

Seventy cases of primary adenocarcinoma of the prostate were available for review. Of these 70 cases, 56 cases (80%) are positive for TAT130 expression. TAT130 expression in non-prostatic tissues is weak or undetectable.

(41) DNA235621 (TAT166)

Seventy cases of primary adenocarcinoma of the prostate were available for review. Of these 70 cases, 56 cases (80%) are positive for TAT166 expression. TAT166 expression in non-prostatic tissues is weak or undetectable.

(42) DNA236493 (TAT141)

Positive expression is observed in 70/148 breast carcinomas, 2/63 colorectal adenocarcinomas, 4/42 ovarian tumors, 9/69 non small cell lung carcinomas, 9/67 prostate adenocarcinomas and 5/25 gliomas. Expression in normal non-cancerous tissues appears restricted to prostate and breast epithelium.

(43) DNA226094 (TAT164)

Twenty one of 37 glioblastoma samples and 8 or 8 glioma samples were positive for TAT164 expression while all other tumor and normal tissues examined (including normal brain tissue) were negative.

(44) DNA227578 (TAT165)

Fifteen of 25 glioblastoma samples teste4d were positive for expression while significantly weaker expression was observed in the normal brain samples tested.

EXAMPLE 5: Immunohistochemistry Analysis

Antibodies against certain TAT polypeptides disclosed herein were prepared and immunohistochemistry analysis was performed as follows. Tissue sections were first fixed for 5 minutes in acetone/ethanol (frozen or paraffin-embedded). The sections were then washed in PBS and then blocked with avidin and biotin (Vector kit) for 10 minutes each followed by a wash in PBS. The sections were then blocked with 10% serum for 20 minutes and then blotted to remove the excess. A primary antibody was then added to the sections at a concentration of 10μg/ml for 1 hour and then the sections were washed in PBS. A biotinylated secondary antibody (anti-primary antibody) was then added to the sections for 30 minutes and then the sections were washed with PBS. The sections were then exposed to the reagents of the Vector ABC kit for 30 minutes and then the sections were washed in PBS. The sections were then exposed to Diaminobenzidine (Pierce) for 5 minutes and then washed in PBS. The sections were then counterstained with Mayers hematoxylin, covered with a coverslip and visualized. Immunohistochemistry analysis can also be performed as described in Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989 and Ausubel et al., Current Protocols of Molecular Biology, Unit 3.16, John Wiley and Sons (1997). The results from these analyses are show below.

(1) DNA96930 (TAT112)

Significantly higher expression was detected in the apical surface of the colonic crypts of colon tumors than on the apical surface of the normal colonic crypts. Additionally, TAT112 was found to be significantly overexpressed in pancreatic adenocarcinoma cells as compared to normal pancreatic cells. Finally, IHC analysis performed as described above evidenced that TAT112 is significantly overexpressed in lung carcinoma as compared to normal lung tissue, non small cell lung carcinoma as compared to normal lung tissue and stomach carcinoma as compared to normal stomach tissue.

(2) DNA226539 (TAT126)

Positive expression is observed in 2/10 uterine adenocarcinomas, 9/17 ovarian adenocarcinomas and 2/20 non small cell lung carcinomas. Using this procedure, expression of TAT126 was not detectable in any normal tissue.

(3) DNA236511 (TAT151)

Positive expression is observed in 2/10 uterine adenocarcinomas, 9/17 ovarian adenocarcinomas and 2/20 non small cell lung carcinomas. Using this procedure, expression of TAT151 was not detectable in any normal tissue.

EXAMPLE 6: Verification and Analysis of Differential TAT Polypeptide Expression by GEPIS

TAT polypeptides which may have been identified as a tumor antigen as described in one or more of the above Examples were analyzed and verified as follows. An expressed sequence tag (EST) DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) was searched and interesting EST sequences were identified by GEPIS. Gene expression profiling *in silico* (GEPIS) is a bioinformatics tool developed at Genentech, Inc. that characterizes genes of interest for new cancer therapeutic targets. GEPIS takes advantage of large amounts of EST sequence and library information to determine gene expression profiles. GEPIS is capable of determining the expression profile of a gene based upon its proportional correlation with the number of its occurrences in EST databases, and it works by integrating the LIFESEQ® EST relational database and Genentech proprietary information in a stringent and statistically meaningful way. In this example, GEPIS is used to identify and cross-validate novel tumor antigens, although GEPIS can be configured to perform either very specific analyses or broad screening tasks. For the initial screen, GEPIS is used to identify EST sequences from the LIFESEQ® database that correlate to expression in a particular tissue or tissues of interest (often a tumor tissue of interest). The EST sequences identified in this initial screen (or consensus sequences obtained from aligning multiple related and overlapping EST sequences obtained from the initial screen) were then subjected to a screen intended to identify the presence of at least one transmembrane domain in the encoded protein. Finally, GEPIS was employed to generate a complete tissue expression profile for the various sequences of interest. Using this type of screening bioinformatics, various TAT polypeptides (and their encoding nucleic acid molecules) were identified as being significantly overexpressed in a particular type of cancer or certain cancers as compared to other cancers and/or normal non-cancerous tissues. The rating of GEPIS hits is based upon several criteria including, for example, tissue specificity, tumor specificity and expression level in normal essential and/or normal proliferating tissues. The following is a list of molecules whose tissue expression profile as determined by GEPIS evidences high tissue expression and significant upregulation of expression in a specific tumor or tumors as compared to other tumor(s) and/or normal tissues and optionally relatively low expression in normal essential and/or normal proliferating tissues. As such, the molecules listed below are excellent polypeptide targets for the diagnosis and therapy of cancer in mammals.

Molecule	upregulation of expression in:	as compared to:
DNA77507 (TAT161)	breast tumor	normal breast tissue
DNA77507 (TAT161)	colon tumor	normal colon tissue

DNA77507 (TAT161)	lung tumor	normal lung tissue
DNA77507 (TAT161)	kidney tumor	normal kidney tissue
DNA77507 (TAT161)	liver tumor	normal liver tissue
DNA77507 (TAT161)	ovarian tumor	normal ovarian tissue
DNA77507 (TAT161)	pancreatic tumor	normal pancreatic tissue
DNA77507 (TAT161)	rectum tumor	normal rectum tissue
DNA77507 (TAT161)	skin tumor	normal skin tissue
DNA77507 (TAT161)	uterine tumor	normal uterine tissue
DNA77507 (TAT161)	brain tumor	normal brain tissue
DNA77507 (TAT161)	soft tissue tumor	normal soft tissue
DNA77507 (TAT161)	bone tumor	normal bone tissue
DNA82343 (TAT157)	colon tumor	normal colon tissue
DNA82343 (TAT157)	ovarian tumor	normal ovarian tissue
DNA82343 (TAT157)	stomach tumor	normal stomach tissue
DNA82343 (TAT157)	thymus tumor	normal thymus tissue
DNA82343 (TAT157)	small intestine tumor	normal small intestine tissue
DNA87994 (TAT160)	breast tumor	normal breast tissue
DNA87994 (TAT160)	pancreatic tumor	normal pancreatic tissue
DNA87994 (TAT160)	colon tumor	normal colon tissue
DNA87994 (TAT160)	esophagus tumor	normal esophagus tissue
DNA87994 (TAT160)	ovarian tumor	normal ovarian tissue
DNA87994 (TAT160)	prostate tumor	normal prostate tissue
DNA88131 (TAT158)	breast tumor	normal breast tissue
DNA88131 (TAT158)	colon tumor	normal colon tissue
DNA88131 (TAT158)	lung tumor	normal lung tissue

Molecule	upregulation of expression in:	as compared to:
DNA88131 (TAT158)	pancreatic tumor	normal pancreatic tissue
DNA88131 (TAT158)	prostate tumor	normal prostate tissue
DNA88131 (TAT158)	stomach tumor	normal stomach tissue
DNA88131 (TAT158)	bladder tumor	normal bladder tissue
DNA88131 (TAT158)	brain tumor	normal brain tissue
DNA95930 (TAT110)	colon tumor	normal colon tissue
DNA95930 (TAT110)	lung tumor	normal lung tissue
DNA95930 (TAT110)	prostate tumor	normal prostate tissue
DNA95930 (TAT110)	endometrial tumor	normal endometrial tissue
DNA95930 (TAT110)	ovarian tumor	normal ovarian tissue

DNA95930 (TAT110)	breast tumor	normal breast tissue
DNA95930-1 (TAT210)	colon tumor	normal colon tissue
DNA95930-1 (TAT210)	lung tumor	normal lung tissue
DNA95930-1 (TAT210)	prostate tumor	normal prostate tissue
DNA95930-1 (TAT210)	endometrial tumor	normal endometrial tissue
DNA95930-1 (TAT210)	ovarian tumor	normal ovarian tissue
DNA95930-1 (TAT210)	breast tumor	normal breast tissue
DNA96917 (TAT159)	pancreatic tumor	normal pancreatic tissue
DNA96917 (TAT159)	lung tumor	normal lung tissue
DNA96917 (TAT159)	liver tumor	normal liver tissue
DNA96917 (TAT159)	prostate tumor	normal prostate tissue
DNA96930 (TAT112)	breast tumor	normal breast tissue
DNA96930 (TAT112)	colon tumor	normal colon tissue
DNA96930 (TAT112)	lung tumor	normal lung tissue
DNA96930 (TAT112)	ovarian tumor	normal ovarian tissue
DNA96930 (TAT112)	pancreatic tumor	normal pancreatic tissue
DNA96930 (TAT112)	stomach tumor	normal stomach tissue
DNA96936 (TAT147)	breast tumor	normal breast tissue
DNA96936 (TAT147)	colon tumor	normal colon tissue
DNA96936 (TAT147)	prostate tumor	normal prostate tissue
DNA96936 (TAT147)	uterine tumor	normal uterine tissue
DNA98565 (TAT145)	brain tumor	normal brain tissue
DNA98565 (TAT145)	colon tumor	normal colon tissue
DNA246435 (TAT152)	brain tumor	normal brain tissue
DNA246435 (TAT152)	colon tumor	normal colon tissue
DNA98591 (TAT162)	colon tumor	normal colon tissue
DNA98591 (TAT162)	small intestine tumor	normal small intestine tissue
DNA98591 (TAT162)	ovarian tumor	normal ovarian tissue
DNA98591 (TAT162)	esophagus tumor	normal esophagus tissue
DNA108809 (TAT114)	colon tumor	normal colon tissue
DNA108809 (TAT114)	lung tumor	normal lung tissue
DNA108809 (TAT114)	ovarian tumor	normal ovarian tissue
DNA108809 (TAT114)	brain tumor	normal brain tissue
DNA143493 (TAT103)	breast tumor	normal breast tissue
DNA167234 (TAT130)	prostate tumor	normal prostate tissue
DNA235621 (TAT166)	prostate tumor	normal prostate tissue
DNA176766 (TAT132)	kidney tumor	normal kidney tissue

DNA176766 (TAT132)	uterine tumor	normal uterine tissue
DNA236463 (TAT150)	kidney tumor	normal kidney tissue
DNA236463 (TAT150)	uterine tumor	normal uterine tissue
DNA181162 (TAT129)	prostate tumor	normal prostate tissue
DNA188221 (TAT111)	colon tumor	normal colon tissue
DNA188221 (TAT111)	liver tumor	normal liver tissue
DNA188221 (TAT111)	lung tumor	normal lung tissue
DNA233876 (TAT146)	colon tumor	normal colon tissue

Molecule	upregulation of expression in:	as compared to:
DNA233876 (TAT146)	liver tumor	normal liver tissue
DNA233876 (TAT146)	lung tumor	normal lung tissue
DNA193891 (TAT148)	prostate tumor	normal prostate tissue
DNA193891 (TAT148)	breast tumor	normal breast tissue
DNA248170 (TAT187)	breast tumor	normal breast tissue
DNA248170 (TAT187)	prostate tumor	normal prostate tissue
DNA194628 (TAT118)	kidney tumor	normal kidney tissue
DNA246415 (TAT167)	kidney tumor	normal kidney tissue
DNA215609 (TAT113)	colon tumor	normal colon tissue
DNA220432 (TAT128)	prostate tumor	normal prostate tissue
DNA226094 (TAT164)	breast tumor	normal breast tissue
DNA226094 (TAT164)	brain tumor	normal brain tissue
DNA226094 (TAT164)	ovarian tumor	normal ovarian tissue
DNA226094 (TAT164)	lung tumor	normal lung tissue
DNA226165 (TAT122)	breast tumor	normal breast tissue
DNA226165 (TAT122)	endometrial tumor	normal endometrial tissue
DNA226165 (TAT122)	lung tumor	normal lung tissue
DNA226165 (TAT122)	colon tumor	normal colon tissue
DNA226237 (TAT117)	kidney tumor	normal kidney tissue
DNA246450 (TAT168)	kidney tumor	normal kidney tissue
DNA246450 (TAT168)	brain tumor	normal brain tissue
DNA226456 (TAT144)	breast tumor	normal breast tissue
DNA226456 (TAT144)	brain tumor	normal brain tissue
DNA226456 (TAT144)	endometrial tumor	normal endometrial tissue
DNA226456 (TAT144)	kidney tumor	normal kidney tissue
DNA226456 (TAT144)	lung tumor	normal lung tissue
DNA237637 (TAT188)	breast tumor	normal breast tissue

DNA237637 (TAT188)	brain tumor	normal brain tissue
DNA237637 (TAT188)	endometrial tumor	normal endometrial tissue
DNA237637 (TAT188)	kidney tumor	normal kidney tissue
DNA237637 (TAT188)	lung tumor	normal lung tissue
DNA226539 (TAT126)	colon tumor	normal colon tissue
DNA226539 (TAT126)	endometrial tumor	normal endometrial tissue
DNA226539 (TAT126)	ovarian tumor	normal ovarian tissue
DNA226539 (TAT126)	pancreatic tumor	normal pancreatic tissue
DNA236511 (TAT151)	colon tumor	normal colon tissue
DNA236511 (TAT151)	endometrial tumor	normal endometrial tissue
DNA236511 (TAT151)	ovarian tumor	normal ovarian tissue
DNA236511 (TAT151)	pancreatic tumor	normal pancreatic tissue
DNA226771 (TAT115)	colon tumor	normal colon tissue
DNA227087 (TAT163)	breast tumor	normal breast tissue
DNA227087 (TAT163)	colon tumor	normal colon tissue
DNA227087 (TAT163)	endocrine tumor	normal endocrine tissue
DNA227087 (TAT163)	kidney tumor	normal kidney tissue
DNA227087 (TAT163)	liver tumor	normal liver tissue
DNA227087 (TAT163)	lung tumor	normal lung tissue
DNA227087 (TAT163)	pancreatic tumor	normal pancreatic tissue
DNA227087 (TAT163)	uterine tumor	normal uterine tissue
DNA227087 (TAT163)	prostate tumor	normal prostate tissue
DNA227087 (TAT163)	bladder tumor	normal bladder tissue
DNA266307 (TAT227)	breast tumor	normal breast tissue
DNA266307 (TAT227)	colon tumor	normal colon tissue
DNA266307 (TAT227)	endocrine tumor	normal endocrine tissue
DNA266307 (TAT227)	kidney tumor	normal kidney tissue
DNA266307 (TAT227)	liver tumor	normal liver tissue

Molecule	upregulation of expression in:	as compared to:
DNA266307 (TAT227)	lung tumor	normal lung tissue
DNA266307 (TAT227)	pancreatic tumor	normal pancreatic tissue
DNA266307 (TAT227)	uterine tumor	normal uterine tissue
DNA266307 (TAT227)	prostate tumor	normal prostate tissue
DNA266307 (TAT227)	bladder tumor	normal bladder tissue
DNA266311 (TAT228)	breast tumor	normal breast tissue
DNA266311 (TAT228)	colon tumor	normal colon tissue

DNA266311 (TAT228)	endocrine tumor	normal endocrine tissue
DNA266311 (TAT228)	kidney tumor	normal kidney tissue
DNA266311 (TAT228)	liver tumor	normal liver tissue
DNA266311 (TAT228)	lung tumor	normal lung tissue
DNA266311 (TAT228)	pancreatic tumor	normal pancreatic tissue
DNA266311 (TAT228)	uterine tumor	normal uterine tissue
DNA266311 (TAT228)	prostate tumor	normal prostate tissue
DNA266311 (TAT228)	bladder tumor	normal bladder tissue
DNA266312 (TAT229)	breast tumor	normal breast tissue
DNA266312 (TAT229)	colon tumor	normal colon tissue
DNA266312 (TAT229)	endocrine tumor	normal endocrine tissue
DNA266312 (TAT229)	kidney tumor	normal kidney tissue
DNA266312 (TAT229)	liver tumor	normal liver tissue
DNA266312 (TAT229)	lung tumor	normal lung tissue
DNA266312 (TAT229)	pancreatic tumor	normal pancreatic tissue
DNA266312 (TAT229)	uterine tumor	normal uterine tissue
DNA266312 (TAT229)	prostate tumor	normal prostate tissue
DNA266312 (TAT229)	bladder tumor	normal bladder tissue
DNA266313 (TAT230)	breast tumor	normal breast tissue
DNA266313 (TAT230)	colon tumor	normal colon tissue
DNA266313 (TAT230)	endocrine tumor	normal endocrine tissue
DNA266313 (TAT230)	kidney tumor	normal kidney tissue
DNA266313 (TAT230)	liver tumor	normal liver tissue
DNA266313 (TAT230)	lung tumor	normal lung tissue
DNA266313 (TAT230)	pancreatic tumor	normal pancreatic tissue
DNA266313 (TAT230)	uterine tumor	normal uterine tissue
DNA266313 (TAT230)	prostate tumor	normal prostate tissue
DNA266313 (TAT230)	bladder tumor	normal bladder tissue
DNA227224 (TAT121)	breast tumor	normal breast tissue
DNA227224 (TAT121)	endometrial tumor	normal endometrial tissue
DNA227224 (TAT121)	lung tumor	normal lung tissue
DNA227224 (TAT121)	skin tumor	normal skin tissue
DNA247486 (TAT183)	breast tumor	normal breast tissue
DNA247486 (TAT183)	endometrial tumor	normal endometrial tissue
DNA247486 (TAT183)	lung tumor	normal lung tissue
DNA247486 (TAT183)	skin tumor	normal skin tissue
DNA227578 (TAT165)	brain tumor	normal brain tissue

DNA227800 (TAT131)	prostate tumor	normal prostate tissue
DNA227800 (TAT131)	kidney tumor	normal kidney tissue
DNA227904 (TAT140)	breast tumor	normal breast tissue
DNA228199 (TAT127)	uterine tumor	normal uterine tissue
DNA228199 (TAT127)	fallopian tube tumor	normal fallopian tube tissue
DNA228199 (TAT127)	ovarian tumor	normal ovarian tissue
DNA228199 (TAT127)	lung tumor	normal lung tissue
DNA228201 (TAT116)	colon tumor	normal colon tissue
DNA247488 (TAT189)	colon tumor	normal colon tissue
DNA236538 (TAT190)	colon tumor	normal colon tissue
DNA247489 (TAT191)	colon tumor	normal colon tissue

Molecule	upregulation of expression in:	as compared to:
DNA231312 (TAT143)	colon tumor	normal colon tissue
DNA231542 (TAT100)	brain tumor	normal brain tissue
DNA231542 (TAT100)	glioma	normal glial tissue
DNA231542-1 (TAT284)	brain tumor	normal brain tissue
DNA231542-1 (TAT284)	glioma	normal glial tissue
DNA231542-2 (TAT285)	brain tumor	normal brain tissue
DNA231542-2 (TAT285)	glioma	normal glial tissue
DNA297393 (TAT285-1)	brain tumor	normal brain tissue
DNA297393 (TAT285-1)	glioma	normal glial tissue
DNA232754 (TAT125)	lung tumor	normal lung tissue
DNA236246 (TAT153)	breast tumor	normal breast tissue
DNA236343 (TAT104)	breast tumor	normal breast tissue
DNA236493 (TAT141)	breast tumor	normal breast tissue
DNA236493 (TAT141)	glioblastoma tumor	normal glial tissue
DNA236534 (TAT102)	breast tumor	normal breast tissue
DNA236534 (TAT102)	lung tumor	normal lung tissue
DNA236534 (TAT102)	pancreatic tumor	normal pancreatic tissue
DNA236534 (TAT102)	prostate tumor	normal prostate tissue
DNA236534 (TAT102)	bladder tumor	normal bladder tissue
DNA247480 (TAT142)	lung tumor	normal lung tissue
DNA264454 (TAT106)	breast tumor	normal breast tissue
DNA264454 (TAT106)	prostate tumor	normal prostate tissue
DNA264454 (TAT106)	ovarian tumor	normal ovarian tissue

EXAMPLE 7: Use of TAT as a hybridization probe

The following method describes use of a nucleotide sequence encoding TAT as a hybridization probe for, i.e., diagnosis of the presence of a tumor in a mammal.

DNA comprising the coding sequence of full-length or mature TAT as disclosed herein can also be employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of TAT) in human tissue cDNA libraries or human tissue genomic libraries.

Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled TAT-derived probe to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42 °C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

DNAs having a desired sequence identity with the DNA encoding full-length native sequence TAT can then be identified using standard techniques known in the art.

EXAMPLE 8: Expression of TAT in *E. coli*

This example illustrates preparation of an unglycosylated form of TAT by recombinant expression in *E. coli*.

The DNA sequence encoding TAT is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the TAT coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., supra. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized TAT protein can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

TAT may be expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding TAT is initially amplified using selected PCR primers. The primers will contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences are then ligated into an expression vector, which is used to transform an *E. coli* host based on strain 52 (W3110 fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq). Transformants are first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an O.D.600 of 3-5 is reached. Cultures are then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g $(NH_4)_2SO_4$, 0.71 g sodium citrate•2H2O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 mL water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM $MgSO_4$) and grown for approximately 20-30 hours at 30°C with shaking. Samples are removed to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets are frozen until purification and refolding.

E. coli paste from 0.5 to 1 L fermentations (6-10 g pellets) is resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution is stirred overnight at 4°C. This step results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution is centrifuged at 40,000 rpm in a Beckman Ultracentifuge for 30 min. The supernatant is diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. The clarified extract is loaded onto a 5 ml Qiagen Ni-NTA metal chelate column equilibrated in the metal chelate column buffer. The column is washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein is eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein are pooled and stored at 4°C. Protein concentration is estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

The proteins are refolded by diluting the sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM EDTA. Refolding volumes are chosen so that the final protein concentration is between 50 to 100 micrograms/ml. The refolding solution is stirred gently at 4°C for 12-36 hours. The refolding reaction is quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the solution is filtered through a 0.22 micron filter and acetonitrile is added to 2-10% final concentration. The refolded protein is chromatographed on a Poros R1/H reversed phase column using a mobile buffer of 0.1% TFA with elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A280 absorbance are analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein are pooled. Generally, the properly refolded species of most proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded forms of proteins from the desired form, the reversed phase step also removes endotoxin from the samples.

Fractions containing the desired folded TAT polypeptide are pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins are formulated into 20 mM Hepes, pH 6.8 with 0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

Certain of the TAT polypeptides disclosed herein have been successfully expressed and purified using this technique(s).

EXAMPLE 9: Expression of TAT in mammalian cells

This example illustrates preparation of a potentially glycosylated form of TAT by recombinant expression in mammalian cells.

The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the TAT DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the TAT DNA using ligation methods such as described in Sambrook et al., supra. The resulting vector is called pRK5-TAT.

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10μg pRK5-TAT DNA is mixed with about 1 μg DNA encoding the VA RNA gene [Thimmappaya et al., Cell, 31:543 (1982)] and dissolved in 500 μl of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M $CaCl_2$. To this mixture is added, dropwise, 500μl of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM $NaPO_4$, and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 μCi/ml ^{35}S-cysteine and 200 μCi/ml ^{35}S-methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of TAT polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, TAT may be introduced into 293 cells transiently using the dextran sulfate method described by Somparyrac et al., Proc. Natl. Acad. Sci., 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μg pRK5-TAT DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 μg/ml bovine insulin and 0.1 μg/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris.

The sample containing expressed TAT can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, TAT can be expressed in CHO cells. The pRK5-TAT can be transfected into CHO cells using known reagents such as $CaPO_4$ or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ^{35}S-methionine. After determining the presence of TAT polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed TAT can then be concentrated and purified by any selected method.

Epitope-tagged TAT may also be expressed in host CHO cells. The TAT may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged TAT insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged TAT can then be concentrated and purified by any selected method, such as by Ni^{2+}-chelate affinity chromatography.

TAT may also be expressed in CHO and/or COS cells by a transient expression procedure or in CHO cells by another stable expression procedure.

Stable expression in CHO cells is performed using the following procedure. The proteins are expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g. extracellular domains) of the respective proteins are fused to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or is a poly-His tagged form.

Following PCR amplification, the respective DNAs are subcloned in a CHO expression vector using standard techniques as described in Ausubel et al., Current Protocols of Molecular Biology, Unit 3.16, John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used expression in CHO cells is as described in Lucas et al., Nucl. Acids Res. 24:9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

Twelve micrograms of the desired plasmid DNA is introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect® (Quiagen), Dosper® or Fugene® (Boehringer Mannheim). The cells are grown as described in Lucas et al. supra. Approximately 3 x 10^7 cells are frozen in an ampule for further growth and production as described below.

The ampules containing the plasmid DNA are thawed by placement into water bath and mixed by vortexing. The contents are pipetted into a centrifuge tube containing 10 mLs of media and centrifuged at 1000 rpm for 5 minutes. The supernatant is aspirated and the cells are resuspended in 10 mL of selective media (0.2 μm filtered PS20 with 5% 0.2 μm diafiltered fetal bovine serum). The cells are then aliquoted into a 100 mL spinner

containing 90 mL of selective media. After 1-2 days, the cells are transferred into a 250 mL spinner filled with 150 mL selective growth medium and incubated at 37°C. After another 2-3 days, 250 mL, 500 mL and 2000 mL spinners are seeded with 3 x 10^5 cells/mL. The cell media is exchanged with fresh media by centrifugation and resuspension in production medium. Although any suitable CHO media may be employed, a production medium described in U.S. Patent No. 5,122,469, issued June 16, 1992 may actually be used. A 3L production spinner is seeded at 1.2 x 10^6 cells/mL. On day 0, the cell number pH is determined. On day 1, the spinner is sampled and sparging with filtered air is commenced. On day 2, the spinner is sampled, the temperature shifted to 33°C, and 30 mL of 500 g/L glucose and 0.6 mL of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning 365 Medical Grade Emulsion) taken. Throughout the production, the pH is adjusted as necessary to keep it at around 7.2. After 10 days, or until the viability dropped below 70%, the cell culture is harvested by centrifugation and filtering through a 0.22 μm filter. The filtrate was either stored at 4°C or immediately loaded onto columns for purification.

For the poly-His tagged constructs, the proteins are purified using a Ni-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4 °C. After loading, the column is washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein is subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc-containing) constructs are purified from the conditioned media as follows. The conditioned medium is pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 ml fractions into tubes containing 275 μL of 1 M Tris buffer, pH 9. The highly purified protein is subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity is assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

Certain of the TAT polypeptides disclosed herein have been successfully expressed and purified using this technique(s).

EXAMPLE 10: Expression of TAT in Yeast

The following method describes recombinant expression of TAT in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of TAT from the ADH2/GAPDH promoter. DNA encoding TAT and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of TAT. For secretion, DNA encoding TAT can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, a native TAT signal peptide or other mammalian signal peptide, or, for example, a yeast alpha-factor or invertase secretory signal/leader sequence, and linker sequences (if needed) for expression of TAT.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant TAT can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing TAT may further be purified using selected column chromatography resins.

Certain of the TAT polypeptides disclosed herein have been successfully expressed and purified using this technique(s).

EXAMPLE 11: Expression of TAT in Baculovirus-Infected Insect Cells

The following method describes recombinant expression of TAT in Baculovirus-infected insect cells.

The sequence coding for TAT is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the sequence encoding TAT or the desired portion of the coding sequence of TAT such as the sequence encoding an extracellular domain of a transmembrane protein or the sequence encoding the mature protein if the protein is extracellular is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 - 5 days of incubation at 28C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression are performed as described by O'Reilley et al., Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

Expressed poly-his tagged TAT can then be purified, for example, by $^{2+}$Ni-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al Nature, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM $MgCl_2$; 0.1 mM EDTA; 10% glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% glycerol, pH 7.8) and filtered through a 0.45 μm filter. A Ni^{2+}-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A_{280} with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A_{280} baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by

SDS-PAGE and silver staining or Western blot with Ni^{2+}-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His_{10}-tagged TAT are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) TAT can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

Certain of the TAT polypeptides disclosed herein have been successfully expressed and purified using this technique(s).

EXAMPLE 12: Preparation of Antibodies that Bind TAT

This example illustrates preparation of monoclonal antibodies which can specifically bind TAT. This example further illustrates preparation of monoclonal antibodies which can specifically bind TAT188(E16) polypeptide.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, supra. Immunogens that may be employed include purified TAT, fusion proteins containing TAT, and cells expressing recombinant TAT on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the TAT immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-TAT antibodies.

The TAT188 polypeptide, E16, is a twelve-transmembrane protein with both the C- and N-terminus located intracellularly. TAT(E16) is a subunit of the Na2+ independent large neutral amino acid transporter, heterodimerized with a common heavy chain 4F2hc (CD98hc) to form a functional unit (see diagram in Fig. 155). To prepare anti-TAT188 (anti-E16) antibodies targeting extracellular domain(s), PC3 cells, which endogenously express E16, were used as the immunogen. Briefly, PC3 cells (22×10^6 cells/ml) were injected into Balb-c mice. Antibodies titers were tested against control 293 cells and 293-E16 cells (293 cells transfected with and producing E16). Six hybridomas were prepared which expressed anti-E16 antibodies. Of these, three antibodies were shown to have good specificity for E16 by FACS analysis (see Fig. 156) in which cell surface binding paralleled E16 surface expression. Specifically E16 siRNA transfection inhibited E16 expression (see Example 17 herein) and also decreased the amount of anti-E16 antibody bound to the surface of PC3 cells. These antibodies are referred to herein as 3B5.1 (or 3B5), 12G12.1 (or 12G12), and 12B9.1 (or 12B9) and were used for further experiments disclosed herein.

Hybridomas expressing the antibodies of the invention were prepared as follows. After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of TAT. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen

cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells were screened in an ELISA for reactivity against TAT. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against TAT is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-TAT monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

Antibodies directed against certain of the TAT polypeptides disclosed herein have been successfully produced using this technique(s). More specifically, functional monoclonal antibodies that are capable of recognizing and binding to TAT protein (as measured by standard ELISA, FACS sorting analysis and/or immunohistochemistry analysis) have been successfully generated against the following TAT proteins as disclosed herein: TAT110 (DNA95930), TAT210 (DNA95930-1), TAT113 (DNA215609), TAT126 (DNA226539), TAT151 (DNA236511), TAT111 (DNA188221), TAT146 (DNA233876), TAT112 (DNA96930), TAT145 (DNA98565), TAT152 (DNA246435), TAT141 (DNA236493), TAT114 (DNA108809), TAT104 (DNA236343), TAT100 (DNA231542), TAT284 (DNA231542-1), TAT285 (DNA231542-2), TAT285-1 (DNA297393), TAT144 (DNA226456), TAT188 (DNA237637), TAT123 (DNA210499), TAT211 (DNA219894), TAT102 (DNA236534), TAT127 (DNA228199) and TAT128 (DNA220432). Interestingly, Applicants have identified that the monoclonal antibodies prepared against TAT111 (DNA188221) and TAT146 (DNA233876) are capable of blocking activation of the EphB2R receptor encoded by the DNA188221 and DNA233876 molecules by its associated ligand polypeptide. As such, antibodies and methods for using those antibodies for blocking activation of the EphB2R receptor (i.e., TAT111 and TAT146 polypeptides) by its associated ligand are encompassed within the presently described invention. Moreover, Applicants have identified that monoclonal antibodies directed against the TAT110 (DNA95930) and TAT210 (DNA95930-1) polypeptides (i.e., IL-20 receptor alpha polypeptides) are capable of inhibiting activation of the IL20 receptor alpha by IL-19 protein.As such, antibodies and methods for using those antibodies for inhibiting activation of the IL-20 receptor alpha (i.e., TAT110 and TAT210 polypeptides) by IL-19 are encompassed within the presently described invention.

In addition to the successful preparation of monoclonal antibodies directed against the TAT polypeptides as described herein, many of those monoclonal antibodies have been successfully conjugated to a cell toxin for use in directing the cellular toxin to a cell (or tissue) that expresses a TAT polypeptide of interested (both *in vitro* and *in vivo*). For example, toxin (e.g., DM1) derivitized monoclonal antibodies have been successfully generated to the following TAT polypeptides as described herein: TAT110 (DNA95930), TAT210 (DNA95930-1), TAT112 (DNA96930), TAT113 (DNA215609), TAT111 (DNA188221) and TAT146 (DNA233876). As shown herein below, other useful conjugate toxins include the auristatins MMAE and MMAF.

EXAMPLE 13: Purification of TAT Polypeptides Using Specific Antibodies

Native or recombinant TAT polypeptides may be purified by a variety of standard techniques in the art of protein purification. For example, pro-TAT polypeptide, mature TAT polypeptide, or pre-TAT polypeptide is purified by immunoaffinity chromatography using antibodies specific for the TAT polypeptide of interest. In general, an immunoaffinity column is constructed by covalently coupling the anti-TAT polypeptide antibody to an activated chromatographic resin.

Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSE™ (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

Such an immunoaffinity column is utilized in the purification of TAT polypeptide by preparing a fraction from cells containing TAT polypeptide in a soluble form. This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by other methods well known in the art. Alternatively, soluble TAT polypeptide containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

A soluble TAT polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of TAT polypeptide (*e.g.*, high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/TAT polypeptide binding (*e.g.*, a low pH buffer such as approximately pH 2-3, or a high concentration of a chaotrope such as urea or thiocyanate ion), and TAT polypeptide is collected.

EXAMPLE 14: *In Vitro* Tumor Cell Killing Assay

Mammalian cells expressing the TAT polypeptide of interest may be obtained using standard expression vector and cloning techniques. Alternatively, many tumor cell lines expressing TAT polypeptides of interest are publicly available, for example, through the ATCC and can be routinely identified using standard ELISA or FACS analysis. Anti-TAT polypeptide monoclonal antibodies (and toxin conjugated derivatives thereof) may then be employed in assays to determine the ability of the antibody to kill TAT polypeptide expressing cells *in vitro*.

For example, cells expressing the TAT polypeptide of interest are obtained as described above and plated into 96 well dishes. In one analysis, the antibody/toxin conjugate (or naked antibody) is included throughout the cell incubation for a period of 4 days. In a second independent analysis, the cells are incubated for 1 hour with the antibody/toxin conjugate (or naked antibody) and then washed and incubated in the absence of antibody/toxin conjugate for a period of 4 days. Cell viability is then measured using the CellTiter-Glo Luminescent Cell Viability Assay from Promega (Cat# G7571). Untreated cells serve as a negative control.

In one specific analysis, the ability of monoclonal antibodies directed against TAT112 (DNA96930) were analyzed for the ability to kill cells expressing that polypeptide. In one analysis, an expression vector called gD.NCA was prepared. The TAT112 polypeptide encoding sequences inserted into that vector are driven by an SV40 promoter and the vector also contains the SV40 early poly A signal. The gD.NCA vector was co-transfected into PC3 cells along with an SV40 vector that expresses Neo resistance in PC3 cells, and positive transformants were selected in 800 μg/ml G418. Positive clones were isolated in 96 well plates and analyzed by flow cytometry using an anti-TAT112 monoclonal antibody prepared as described above and called 3E6. Clone 3 was selected for the analysis as it was found to express a high level of TAT112 polypeptide on its surface. In a second independent analysis, the pancreatic cancer cell line, Hpaf II, was obtained from the ATCC and employed in the assay.

In another specific analysis, the ability of monoclonal antibodies directed against TAT188 (DNA237637) were analyzed for the ability to kill cells expressing that polypeptide.

First, the ability of a naked anti-TAT188(E16) antibody to affect cellular activity was demonstrated.

A. Anti-TAT188(E16) monoclonal antibodies stimulate TAT188(E16) internalization and reduce amino acid transport activity.

The binding of anti-E16 antibody to cells expressing the E16 protein elicited antibody internalization which paralleled a reduction in amino acid transport. PC3 cells were incubated with primary antibodies for indicated time with proteasome inhibitors (Sigma) at 37°C. Then antibodies are removed and cells are washed with PBS several times. Cells are fixed with 4% PFA, then permeabilized with PBS/0.1%TritonX-100. After blocking with 10% goat serum, cells were stained with secondary antibody (goat anti-mouse IgG-Cy3 conjugated (Jackson immunolabs)) and analyzed by fluorescent microscope. The results show that an anti-E16 monoclonal antibody binds and is internalized by PC3 cells expressing TAT188(E16) on their surface (see Fig. 157). The demonstration of anti-TAT188 antibody internalization highlights a useful feature of the antibody-E16 interaction because the internalization of a bound antibody can result in the uptake of toxins conjugated to the anti-TAT188 antibodies (see sections B and C of this Example 14).

The function of TAT188(E16) as an amino acid transport protein was inhibited by the binding of anti-TAT188(E16) antibody. The leucine transport functional assay was performed as follows. Cells were grown in 24 well plate and treated with anti-E16 antibody for over night. Cells were then rinsed with warm Na++ free Hepes-Ringer, and incubated for 10min at 37°C. Buffer was changed to 200 μl/well of assay buffer containing radiolabeled amino acid (L-[2, 3-^{3}H]-Arginine or L-[3, 4, 5-^{3}H(N)]-Leucine) 5 μCi/ml, 50 μl insodium-free Hepes-Ringer). After incubating the radioactive mixture for30 sec at 37°C, buffer was removed by aspiration and cells were wash with ice-cold sodium-free Hepes-Ringer 1ml/well for 4 times. Thep lates were dried and to them was added 0.2ml/well of 0.2%SDS/0.2N NaOH. 0.1ml aliquots of lysates from each sample were neutralized with 0.1ml of 0.2N Hcl, after which liquid scintillation cocktail was added. (^{3}H) radioactivity of the lysates was tested using scintillation counter. The results in Figure 158 also demonstrate that anti-E16 antibody binding to E16 protein on the cell surface causes overall reduction in amino acid transport after approximately 24 hours. The reduction is due to concomitant E16 polypeptide internalization and loss of

available E16 for transport function.

B. The binding of toxin-conjugated anti-TAT188(E16) antibodies to TAT188(E16)-expressing cells results in cell killing.

Preparation of anti-TAT188-MC-vc-PAB-MMAE by conjugation of anti-TAT188 and MC-vc-PAB-MMAE

Antibodies provide a convenient and highly specific method for delivering cytotoxins to cells where a relatively unique feature of the cell, such as a tissue specific (or disease specific or both) surface protein, allows the cell to be targeted for killing while avoiding the killing of healthy cells. The auristatin cytotoxins coupled to an antibody via a MC-vc-PAB linker as disclosed hereinabove, are useful to demonstrate cell killing resulting from E16 mediated internalization of the antibody-toxin conjugate. The procedure was performed as follows. Using a MC-vc-PAB linker, the MMAE toxin was covalently linked to cysteine residues on anti-TAT188 antibodies. Anti-TAT188, dissolved in 500mM sodium borate and 500 mM sodium chloride at pH 8.0 is treated with an excess of 100mM dithiothreitol (DTT). After incubation at 37 °C for about 30 minutes, the buffer is exchanged by elution over Sephadex G25 resin and eluted with PBS with 1mM DTPA. The thiol/Ab value is checked by determining the reduced antibody concentration from the absorbance at 280 nm of the solution and the thiol concentration by reaction with DTNB (Aldrich, Milwaukee, WI) and determination of the absorbance at 412 nm. The reduced antibody dissolved in PBS is chilled on ice.

The drug linker reagent, maleimidocaproyl-valine-citrulline-PAB-monomethyl auristatin E (MMAE), i.e. MC-vc-PAB-MMAE, dissolved in DMSO, is diluted in acetonitrile and water at known concentration, and added to the chilled reduced anti-TAT188(E16) antibody in PBS. After about one hour, an excess of maleimide is added to quench the reaction and cap any unreacted antibody thiol groups. The reaction mixture is concentrated by centrifugal ultrafiltration and anti-TAT188(E16)-MC-vc-PAB-MMAE was purified and desalted by elution through G25 resin in PBS, filtered through 0.2 μm filters under sterile conditions, and frozen for storage.

Preparation of anti-TAT188-MC-vc-PAB-MMAF by conjugation of anti-TAT188 and MC-vc-PAB-MMAF

Anti-TAT188-MC-val-cit-PAB-MMAF was prepared by conjugation of anti-TAT188(E16) and MC-val-cit-PAB-MMAF following the procedure disclosed above for conjugation of MMAE. MMAF is a derivative of MMAE with a phenylalanine at the C-terminus of the drug.

Determination of in vitro Cytotoxicity of toxin-conjugated anti-TAT188

The toxin-conjugated 3B5, 12B9, and 12G12 antibodies for this experiment were prepared as 3B5-MC-vc-PAB-MMAE, 12B9-MC-vc-PAB-MMAE, and 3B5-MC-vc-PAB-MMAE conjugates, respectively. Antibody-toxin conjugate, alpha-IL8-MC-vc-PAB-MMAE, was used as a negative control. PC3 cells (a prostate cancer cell line, available from ATCC and which endogenously expressed E16 on its surface) and

Colo205 cells (a human colon cancer cell line, available from ATCC and which endogenously expresses E16 on its surface) were contacted with increasing concentrations of toxin-conjugated antibodies and monitored for cell killing by a standard cell viability bioluminescence assay (CellTiterGlo™ Kit, Promega) as described above. The results shown in Fig. 159A (PC3 cells) and Fig. 159B (Colo205 cells) demonstrate that the three toxin-conjugated anti-E16 antibodies promote killing of cells that express E16.

Another toxin, MMAF, was linked to antibody and tested for cytotoxicity. The toxin-linked 3B5 antibodies for this experiment were designated 3B5-MC-vc-PAB-MMAE and 3B5-MC-vc-PAB-MMAF. Antibody-toxin conjugate, alpha-IL8-vc-MMAE, was used as a negative control. The average number of toxin molecules attached to each antibody was approximately four toxin molecules per antibody as shown in Fig. 160. Human colon cancer cell line, Colo205, cells were seeded in 96well plates at 4000cells/well in 50 µl of culture medium. The next day cells were incubated with 3B5-MC-vc-PAB-MMAE, 3B5-MC-vc-PAB-MMAF or alpha-IL8-vc-MMAE antibodies in graded concentrations diluted in 50 µl of culture medium for 72hrs. Cell viability was measured using a CellTiter-Glo™ luminescent cell viability assay kit (Promega). The results shown in Fig. 160 demonstrate that MC-vc-PAB-MMAE and MC-vc-PAB-MMAF-conjugated anti-E16 antibodies have the ability to kill cells expressing E16, with MMAF toxin (EC50 approximately 0.06 µg/ml) providing greater cell killing than MMAF (EC50 approximately 0.007 µg/ml).

Development of the antibodies of the invention as therapeutic agent in humans requires early toxicity studies in monkeys and/or primates. The E16 amino acid sequence varies across species such as human versus monkey, rat and mouse E16 amino acid sequences is 96.26%, 91.05% and 90.66, respectively. To determine whether the anti-human TAT188(E16) antibodies were also capable of binding and killing monkey cells, the following in vitro experiment was performed. A monkey cell line, COS7 (available from ATCC), that endogenously expresses monkey E16 on its surface was used. The anti-E16 antibodies 3B5, 12B9, and 12G12 were contacted with human, monkey, and mouse cells endogenously expressing human, monkey and mouse E16, respectively, and analyzed by standard FACS analysis. The FACS showed that each of the three anti-human E16 antibodies 3B5, 12G12, and 12B9 bind monkey COS7 cells expressing human E16 (see Fig. 161A) as well as human cells expressing human E16 (see Fig. 161B), but do not bind mouse E16 expressing cells (see Fig. 161B). Thus, the anti-TAT188(E16) antibodies of the invention are useful for clinical activity, efficacy and toxicity studies.

EXAMPLE 15: *In Vivo* Tumor Cell Killing Assay

Anti-TAT112 activity:

To test the efficacy of unconjugated anti-TAT112 monoclonal antibodies, anti-TAT112 antibody was injected intraperitoneally into nude mice 24 hours prior to receiving PC3.gD.NCA clone 3 cells (obtained as described in Example 14 above) subcutaneously in the flank. Antibody injections continued twice per week for the remainder of the study. Tumor volume was measured twice per week.

To test the efficacy of DM1-conjugated anti-TAT112 antibody, PC3.gD.NCA clone 3 cells (obtained as described in Example 14 above) were inoculated into the flank of nude mice. When the tumors

reached a mean volume of approximately 100mm3, mice were treated with DM1-conjugated anti-TAT112 antibody intravenously either once or twice per week.

The results of the above analyses demonstrated that both the unconjugated anti-TAT112 as well as the DM1-conjugated anti-TAT112 antibody were highly efficacious in reducing tumor volume in this *in vivo* model. These analyses demonstrate that anti-TAT polypeptide monoclonal antibodies are efficacious for killing tumor cells that express a TAT polypeptide of interest.

Anti-TAT118 (E16) activity:

To test the efficacy of unconjugated anti-TAT188 monoclonal antibodies, anti-TAT188 antibodies 3B5, 12B9, and 12G12 for cell killing in vitro, the following procedure was performed. PC3 cells, which express E16 endogenously, were injected into athymic nude mice (nu/nu) at the dorsal flank. On the same day as cell inoculation, antibodies were injected at 2 mg/kg intraperitoneously twice per week. The control mice were injected with PBS without antibody. Ten mice in each group were monitored for 4 weeks and tumor volume was measured twice per week. By injecting tumor cells and anti-E16 antibody simultaneously, this procedure tested the ability of the administered antibodies to inhibit tumor establishment in the xenograft mice. The results in Fig. 162 show that in this experiment, mean tumor volume was not significantly lower than for control antibody.

To test the efficacy of MC-vc-PAB-MMAE and MC-vc-PAB-MMAF conjugated anti-TAT188(E16) antibody, Colo205 cells were inoculated into the flank of athymic nude mice (nu/nu), and after approximately 1 week, when the tumor had reached a mean volume of approximately 100-200mm^3, mice were injected with conjugated antibody at 3 mg/kg intravenously once per week. The following antibody treatments were performed on the day tumor-containing mice were divided into groups: (1) Mock Control - 0.2ml or less PBS, intravenous injection once per week for 4 weeks; (2) Antibody control - anti-IL8-MC-vc-PAB-MMAE in 0.2ml or less PBS, intravenous injection once per week for 4 weeks; (3) Antibody - 3B5-MC-vc-PAB-MMAE in 0.2ml or less PBS, intravenous injection once per week for 4 weeks; (4) Antibody - 3B5-MC-vc-PAB-MMAF in 0.2ml or less PBS, intravenous injection once per week for 4 weeks. Mean tumor volume was measured and plotted versus time. The results in Fig. 163 show that MMAE- and MMAF-conjugated anti-E16 antibody significantly reduced tumor volume in this *in vivo* model. These analyses demonstrate that anti-TAT polypeptide monoclonal antibodies, such as anti-188(E16) antibodies are efficacious for killing tumor cells that express a TAT polypeptide of interest, such as the TAT188(E16) polypeptide.

EXAMPLE 16: Northern Blot Analysis

Northern blot analysis was performed essentially as described by Sambrook et al., supra. Northern blot analysis using probes derived from DNA231542, DNA231542-1, DNA231542-2 and DNA297393 evidences significant upregulation of expression in human glioma tissue as compared to normal human brain tissue.

Northern blot analysis using probes derived from DNA237637 (TAT188, E16) were also

performed and the results showed that expression in human breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver evidences significant upregulation of expression in human cancers as compared to normal human tissue of the same tissue type.

EXAMPLE 17: Modulating TAT expression by siRNA

siRNAs have proven useful as a tool in studies of modulating gene expression where traditional antagonists such as small molecules or antibodies have failed. (Shi Y., *Trends in Genetics* 19(1):9-12 (2003)). In vitro synthesized, double stranded RNAs that are 21 to 23 nucleotides in length can act as interfering RNAs (iRNAs) and can specifically inhibit gene expression (Fire A., *Trends in Genetics* 391; 806-810 (1999)). These iRNAs act by mediating degradation of their target RNAs. Since they are under 30 nuclotides in length, however they do not trigger a cell antiviral defense mechanism. Such mechanisms include interferon production, and a general shutdown of host cell protein synthesis. Practically, siRNAs can by synthesized and then cloned into DNA vectors. Such vectors can be transfected and made to express the siRNA at high levels. The high level of siRNA expression is used to "knockdown" or significantly reduce the amount of protein produced in a cell, and thus it is useful in experiments where overexpression of a protein is believed to be linked to a disorder such as cancer. The TATs are cell surface expressed tumor antigens and TAT overexpression was shown by microarray, Taqman™ and in situ hybridization analyses. While antibody binding to TATs such as TAT188 are shown herein to inhibit function of the tumor antigen and, with respect to toxin-conjugated antibodies, cause cell killing), siRNAs are useful antagonists to TAT proteins by limiting cellular production of the antigen.

Experimental Methods and Results: *TAT188 (E16)*

The following experiments demonstrate that siRNA specific for TAT188 (also referred to herein as E16) reduces RNA expression, protein production, cell surface expression, and protein function. The same or similar procedures are useful demonstrating siRNA downregulation of other TAT RNAs of the invention. Selection of a target sequence may be performed by any suitable procedure (see, for example, Amarzguioul, M. And Prydz, H., Biocham Biophys Res Comm 316:1050-1058 (2004).

A. Transient transfection of E16 siRNA reduces the expression of E16-GFP in PC3-E16-GFP cells.

The prostate cancer cell line, PC3, was used in this set of experiments as it overexpresses TAT188 endogenously. PC3 cells were transfected using double-stranded 21mer RNA oligos directed against TAT188 (siTAT188), lamin or mock transfected. These oligos are listed below.

siRNA oligos:

TAT188(E16)

The DNA sequence of the target region within the coding sequence of TAT188(E16) is shown below followed by the sequences of the sense and antisense strands of the double stranded siRNA.

AAGGAAGAGGCGCGGGAGAAG (SEQ ID NO:155), is the nucleic acid sequence within the TAT188 coding region targeted by the siRNA oligos used in the TAT188 siRNA examples disclosed herein. Other sequences may be chosen as targets according to standard practice with the field of siRNA gene silencing.

TAT188(E16) siRNA oligos:

Sense: r(GGAAGAGGCGCGGGAGAAG)TT (SEQ ID NO:156)

Antisense: r(CUUCUCCCGCGCCUCUUCC)TT (SEQ ID NO:157)

lamin siRNA oligos:

Sense: r(CUGGACUUCCAGAAGAACA)TT (SEQ ID NO:158)

Antisense: r(UGUUCUGGAAGUCCAG)TT (SEQ ID NO:159)

(see Elbashir et al., Nature 411:494-498 (2001)

Two days post-transfection and the relative level of TAT188(E16) polypeptide versus the knock-down of E16-EGFP expression was analyzed by fluorescent microscope monitoring of EGFP fluorescence. A decrease in E16 expression parallels a reduction in EGFP production and bioluminescence detected. The results in Fig. 164 demonstrate that transient transfection with siTAT188(E16) siRNA reduced E16-EGFP fusion expression. The fusion protein refered to as "GFP" in Figs. 164 and 164 refer to an E16-EGFP fusion produced by cloning the E16 gene into pEGFP-C1 vector from BDClontech).

B. Transient transfection of E16 siRNA reduces TAT188(E16) polypeptide production -

Reduction of E16 expression was demonstrated by reduction in E16 protein using Western blot analysis. siTAT188(siE16) RNA oligos SEQ ID NO:156 and SEQ ID NO:157 and the lamin siRNA oligos SEQ ID NO:158 and SEQ ID NO:159 used in this example were the same as shown above in part A of this example. PC3 cells transiently transfected with E16-EGFP were harvested 3 days post-transfection and a Western blot was performed on beta-tubulin band to indicate protein loading. Western blot of TAT188(E16)-GFP fusions were visualized using detectably labelled anti-GFP antibody (see Fig. 165, left lane - mock transfection; right lane, E16 siRNA transfection). The results demonstrate that the presence TAT188(E16) siRNA in a cell expressing E16-GFP fusion protein reduced the expression of E16-GFP significantly.

C. Transient transfection of E16 siRNA inhibits endogenous E16 amino acid transporter activity in PC3 cells.

The same procedure for transient transfection of TAT188(E16) siRNA and lamin siRNA (control) was performed. The test and control cells were tested for E16 activity as an amino acid transport

protein as follows using the leucine transport assay described herein above. The results in Fig. 166 show that E16 siRNA reduction in E16 RNA and protein production is consistent with reduction in the amount of E16 available to function in amino acid transport in the cell.

D. Transient transfection of E16 siRNA reduces the surface expression of E16

This example was performed as follows. Forty eight hours after transfection of siRNA, cells were analysed by FACS. Briefly, cells were incubated with primary antibodies for one hour on ice. The cells were then labeled with a secondary antibody conjugated with a fluorescent dye (such as nonlimiting examples Cy-3 or PE) and collecting data with FACScaliber™ or FACScan™. The collected data was analyzed using, as non-limiting examples, Cellquest™ (BD Clontech) or FlowJo ™ (Tree Star, Inc.). As disclosed in Example 12, herein, the three anti-E16 monoclonal antibodies, 3b5, 12B9, and 12G12, were shown to bind less on PC3 cells after transfection of TAT188(E16) siRNA consistent with reduced expression of the E16 protein in the presence of targeted siRNA (see Fig. 156).

E. Transient transfection of TAT188(E16) siRNA inhibits PC3 cell proliferation.

Proliferation of PC3 cells transfected with siRNA targeting TAT188 can be assayed according to the following commercial protocol (Promega Corp. Technical Bulletin TB288; Mendoza et al. (2202) Cancer Res. 62:5485-5488). For example, PC3 cells were plated at 3.5×10^5 cells/well in 6 well plates. The next day, double stranded RNA (dsRNA, SEQ ID NO:155 and SEQ ID NO:156) or lamin control siRNA (using siRNA oligos SEQ ID NO:157 and SEQ ID NO:158) at a final concentration of 100 nM was used to transfect the cells using Lipofectamin2000 (Invitrogen). After 6 hrs of transfection, cells were reseeded into 96 well plates, and cell viability was measured using Cell Counting Kit-8™ (Dojindo) at the time points indicated in Fig. 167, where RLU = relative luminescence units. Other cells that may be assayed by this method if they express TAT188(either endogenous ly or after transfection with nucleic acid encoding TAT188) include without limitation SKBR-3, BT474, MCF7 or MDA-MB-468 demonstrating that reduction in E16 expression in breast cancer cells results in cell killing and that breast cancer, as well as prostate, colorectal and other types of cancer, are desirable targets for treatment using E16 expression knockdown.

These data shows that siRNAs can "knockdown" or reduce the expression of TAT proteins such as TAT188 (E16) protein. The reduction is shown by reduction of RNA expression (in *in situ* hybridization assays), reduced protein production (as shown in Western blot analyses), reduced cell surface expression of TAT188 (E16) (as shown by FACS analyses and internalization assays), and reduced function (as shown by reduction in amino acid transport). siRNA reduction of TAT proteins, such as TAT188 (E16) over a period of time results in reduction of cell proliferation in cancer cell lines. Therefore, siRNA is useful in reducing TAT proteins such as TAT188 (E16). Reduction in the level of TAT proteins, such as TAT188 (E16), causes a reduction in the proliferation of cancer cell lines. Since cancer is a cell proliferative disorder, these results show that antagonists of TAT polypeptides, such as TAT188 (E16) would be useful in the reduction of cancer cell growth. Reduction of cancer cell growth would be useful in the alleviation of cancer in mammals.

Example 18: Deposit of Materials

The following hybridoma cell line has been deposited with the American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209 USA (ATCC):

Hybridoma/Antibody Designation	ATCC No.	Deposit Date
3B5.1	PTA-6193	September 8, 2004
12B9.1	PTA-6194	September 8, 2004
12G12.1	PTA-6195	September 8, 2004

This deposit was made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture for 30 years from the date of deposit. The cell line will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures (a) that access to the culture will be available during pendency of the patent application to one determined by the Commissioner to be entitled thereto under 37 CFR §1.14 and 35 USC §122, and (b) that all restrictions on the availability to the public of the culture so deposited will be irrevocably removed upon the granting of the patent.

The assignee of the present application has agreed that if the culture on deposit should die or be lost or destroyed when cultivated under suitable conditions, it will be promptly replaced on notification with a viable specimen of the same culture. Availability of the deposited cell line is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

WHAT IS CLAIMED IS:

1. An isolated antibody that binds to a polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

2. An isolated antibody that binds to a polypeptide having:

(a) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154);

(b) the amino acid sequence shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

3. The antibody of Claim 1 which is a monoclonal antibody.

4. The antibody of Claim 1 which is an antibody fragment.

5. The antibody of Claim 1 which is a chimeric or a humanized antibody.

6. The antibody of Claim 1 which is conjugated to a growth inhibitory agent.

7. The antibody of Claim 1 which is conjugated to a cytotoxic agent.

8. The antibody of Claim 7, wherein the cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

9. The antibody of Claim 7, wherein the cytotoxic agent is a toxin.

10. The antibody of Claim 9, wherein the toxin is selected from the group consisting of auristatin, maytansinoid and calicheamicin.

11. The antibody of Claim 9, wherein the toxin is a maytansinoid.

12. The antibody of Claim 1 which is produced in bacteria.

13. The antibody of Claim 1 which is produced in CHO cells.

14. The antibody of Claim 1 which induces death of a cell to which it binds.

15. The antibody of Claim 1 which is detectably labeled.

16. The antibody of Claim 9, wherein the toxin is selected from the group consisting of MMAE (mono-methyl auristatin E), MMAF and AEVB (auristatin E valeryl benzylhydrazone), and AFP (Auristatin F phenylene diamine).

17. The antibody of claim 9, wherein the toxin is covalently attached to the antibody by a linker.

18. The antibody of claim 17, wherein the linker is selected from the group consisting of maleimidocaproyl (MC), valine-citrulline (val-cit, vc), citrulline (2-amino-5-ureido pentanoic acid), PAB (p-aminobenzylcarbamoyl), Me (N-methyl-valine citrulline), MC(PEG)6-OH (maleimidocaproyl- polyethylene glycol), SPP (N-Succinimidyl 4-(2-pyridylthio) pentanoate), SMCC (N-Succinimidyl 4-(N-maleimidomethyl) cyclohexane-1 carboxylate), and MC-vc-PAB.

19. The antibody of claim 16, wherein the toxin is MMAE.

20. The antibody of claim 16, wherein the toxin is MMAF.

21. The antibody of claim 19 or 20, wherein the linker is MC-vc-PAB.

22. The antibody of any one of claims 1-21, wherein the antibody binds the TAT188 polypeptide.

23. The antibody of claim 22, wherein the antibody inhibits proliferation or promotes cell death of a cell expressing TAT188.

24. The antibody of claim 23, wherein the cell is a cancer cell.

25. The antibody of claim 24, where the cancer cell is selected from the group of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

26. An isolated antibody that competes with binding to the epitopes of TAT188 polypeptide bound by an antibody produced by the hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195).

27. An isolated antibody having the biological activity of an antibody produced by the hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195), wherein the biological activity is inhibition of cell proliferation or promotion of cell death in a cell expressing TAT188.

28. An isolated antibody comprising in a corresponding complementary determining region (CDR) an amino acid sequence having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid sequence of at least 1, 2, 3, 4, 5, or 6 of the CDR(s) of the antibody produced by a hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195).

29. The antibody of any one of claims 1-21, wherein the antibody comprises in a corresponding complementary determining region (CDR) an amino acid sequence having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid sequence of at least 1, 2, 3, 4, 5, or 6 of the CDR(s) of the antibody produced by a hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195).

30. The antibody of any one of claims 1-21, wherein the antibody exhibits the biological activity of an antibody produced by the hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1

(ATCC Accession No. PTA-6195), wherein the biological activity is inhibition of cell proliferation or promotion of cell death in a cell expressing TAT188.

31. The antibody of claim 30, wherein the cell is a cancer cell.

32. The antibody of claim 31, wherein the cancer cell is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

33. A method of inhibiting growth of a cell expressing TAT188, the method comprising contacting the cell with an antibody of any one of claims 1-21.

34. The method of claim 33, wherein the cell is a cancer cell.

35. The method of claim 34, wherein the cancer cell is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

36. The method of claim 35, wherein the cancer cell is a mammalian cell.

37. The method of claim 36, wherein the mammalian cell is a human cell.

38. A method of inhibiting growth of a cell expressing TAT188, the method comprising contacting the cell with an antibody of any one of claims 1-21, wherein the antibody comprises in a corresponding complementary determining region (CDR) an amino acid sequence having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid sequence of at least 1, 2, 3, 4, 5, or 6 of the CDR(s) of the antibody produced by a hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. PTA-6193), 12B9.1 (ATCC Accession No. PTA-6194) and 12G12.1 (ATCC Accession No. PTA-6195).

39. A method of detecting the level of TAT188 polypeptide expressed in a test cell relative to a control cell, the method comprising:

(a) contacting the test cell and the control cell with an isolated anti-TAT188 antibody of claim 26;

(b) detecting binding of the antibody; and

(c) determining the relative binding of the antibody to the test and control cell.

40. The method of claim 39, wherein the test cell and control cell are lysed .

41. The method of claim 39, wherein the test cell is in a tissue.

42. The method of claim 41, wherein the tissue is a tumor tissue.

43. The method of claim 42, wherein the tumor tissue is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

44. A method of detecting the level of TAT188 polypeptide or a polypeptide having at least 80% sequence identity to the amino acid sequence shown in Figure 115 (SEQ ID NO:115) in a test cell relative to a control cell,, the method comprising:
(a) contacting the test cell and the control cell with an isolated antibody of any one of claims 1-5, 12-15, and 26-28;
(b) detecting binding of the antibody; and
(c) determining the relative binding of the antibody to the test and control cell.

45. The method of claim 44, wherein the level of TAT188 polypeptide in the test cell is greater than that in the control cell.

46. The method of claim 45, wherein the method diagnoses cancer in a tissue containing or having contained the test cell.

47. A TAT binding interfering RNA (siRNA) which binds to a nucleic acid having at least 80% sequence identity to:
(a) a nucleotide sequence shown in Figure 37 (SEQ ID NO:37; and
(b) the complement of (a), wherein the siRNA reduces expression of TAT188.

48. An expression vector comprising the siRNA of Claim 47.

49. The expression vector of Claim 48, wherein said siRNA is operably linked to control sequences recognized by a host cell transfected with the vector.

50. A host cell comprising the expression vector of Claim 49.

51. A composition of matter comprising:
(a) the antibody of Claim 1, or
(b) the siRNA of claim 47,
in combination with a carrier.

52. The composition of matter of Claim 51, wherein said carrier is a pharmaceutically acceptable carrier.

53. An article of manufacture:

(a) a container; and

(b) the composition of matter of Claim 51 contained within said container.

54. The article of manufacture of Claim 53, further comprising a label affixed to said container, or a package insert included with said container, referring to the use of said composition of matter for the therapeutic treatment of or the diagnostic detection of a cancer.

55. A method of inhibiting the growth of a cancer cell that expresses a polypeptide having at least 80% amino acid sequence identity to the amino acid sequence shown in Figure 115 (SEQ ID NO:115), said method comprising contacting said cancer cell with a siRNA that binds to a nucleic acid encoding the amino acid in said cancer cell, thereby inhibiting the growth of said cancer cell.

56. The method of claim 55, wherein the nucleic acid has the sequence shown in Figure 37 (SEQ ID NO:37).

57. The method of Claim 44, wherein the detecting the level of expression of the polypeptide comprises employing an antibody in an immunohistochemistry analysis.

58. A method for treating or preventing a cell proliferative disorder associated with increased expression or activity of a polypeptide having at least 80% amino acid sequence identity to the amino acid sequence shown in Figure 115 (SEQ ID NO:115), said method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a TAT188 polypeptide.

59. The method of claim 58, wherein said antagonist is an isolated anti-TAT188 polypeptide antibody of any one of claims 1-21 and 23-28.

60. The method of claim 59, wherein the cell proliferative disorder is cancer.

61. The method of claim 60, wherein the cancer is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

FIGURE 1

CCAAGAATTCGGCACGAGGGTGACCCGGACAGCTGTCCTCTCTGACACCACCCCGGCCTGC
CTCTTTGTTGCCATGAGAGCTGCCTACCTCTTCCTGCTATTCCTGCCTGCAGGCTTGCTGG
CTCAGGGCCAGTATGACCTGGACCCGCTGCCGCCGTTCCCTGACCACGTCCAGTACACCCA
CTATAGCGACCAGATCGACAACCCAGACTACTATGATTATCAAGAGGTGACTCCTCGGCCC
TCCGAGGAACAGTTCCAGTTCCAGTCCCAGCAGCAAGTCCAACAGGAAGTCATCCCAGCCC
CAACCCCAGAACCAGGAAATGCAGAGCTGGAGCCCACAGAGCCTGGGCCTCTTGACTGCCG
TGAGGAACAGTACCCGTGCACCCGCCTCTACTCCATACACAGGCCTTGCAAACAGTGTCTC
AACGAGGTCTGCTTCTACAGCCTCCGCCGTGTGTACGTCATTAACAAGGAGATCTGTGTTC
GTACAGTGTGTGCCCACGAGGAGCTCCTCCGAGCTGACCTCTGTCGGGACAAGTTCTCCAA
ATGTGGCGTGATGGCCAGCAGCGGCCTGTGCCAATCCGTGGCGGCCTCCTGTGCCAGGAGC
TGTGGGAGCTGCTAGGGTGGTGCTGGCATCCTGAGTCCTGGCCCTCCTGGGATCTGGGGCC
CTCGGGCCCTGCCTGACCTGGTGCTTTTTTCCCCATCCCCATGTTCCTTTTATTCTGTAAA
AAGTTAGTGGACTGCAGCCCTGGGGGTTGCAGGCTGCGGTGCCTCAGGCCCCTCCTTCAGC
CTGTGGCCACCTCTGGGGCACGATGGGGGCTCCCCACTGCCCAGTCTGCCCCTCGGGTTGG
GGGAGTATCCCAGGCCTCTCTGTGGGACCTGGGCCCCTGACGGGCCTTCTCAGCCCGTTTT
GAGGACAGACAGTCCCCCGAGGTAGGCTACATCCCCCCACCCCAGCTGGTCTGCTTGGATT
TCCTACAGCCCCCGTGGGCATGGACCACCTTTATTTTATACAAAATTAAAAACAAGTTTTT
AC

FIGURE 2A

CACTAACGCTCTTCCTAGTCCCCGGGCCAACTCGGACAGTTTGCTCATTTATTGCAACGGT
CAAGGCTGGCTTGTGCCAGAACGGCGCGCGCGCGACGCACGCACACACACGGGGGGAAACT
TTTTTAAAAATGAAAGGCTAGAAGAGCTCAGCGGCGGCGCGGGCCGTGCGCGAGGGCTCCG
GAGCTGACTCGCCGAGGCAGGAAATCCCTCCGGTCGCGACGCCCGGCCCCGCTCGGCGCCC
GCGTGGGATGGTGCAGCGCTCGCCGCCGGGCCCGAGAGCTGCTGCACTGAAGGCCGGCGAC
GATGGCAGCGCGCCCGCTGCCCGTGTCCCCCGCCCGCGCCCTCCTGCTCGCCCTGGCCGGT
GCTCTGCTCGCGCCCTGCGAGGCCCGAGGGGTGAGCTTATGGAACGAAGGAAGAGCTGATG
AAGTTGTCAGTGCCTCTGTTCGGAGTGGGGACCTCTGGATCCCAGTGAAGAGCTTCGACTC
CAAGAATCATCCAGAAGTGCTGAATATTCGACTACAACGGGAAAGCAAAGAACTGATCATA
AATCTGGAAAGAAATGAAGGTCTCATTGCCAGCAGTTTCACGGAAACCCACTATCTGCAAG
ACGGTACTGATGTCTCCCTCGCTCGAAATTACACGGTAATTCTGGGTCACTGTTACTACCA
TGGACATGTACGGGGATATTCTGATTCAGCAGTCAGTCTCAGCACGTGTTCTGGTCTCAGG
GGACTTATTGTGTTTGAAAATGAAAGCTATGTCTTAGAACCAATGAAAAGTGCAACCAACA
GATACAAACTCTTCCCAGCGAAGAAGCTGAAAAGCGTCCGGGGATCATGTGGATCACATCA
CAACACACCAAACCTCGCTGCAAAGAATGTGTTTCCACCACCCTCTCAGACATGGGCAAGA
AGGCATAAAAGAGAGACCCTCAAGGCAACTAAGTATGTGGAGCTGGTGATCGTGGCAGACA
ACCGAGAGTTTCAGAGGCAAGGAAAAGATCTGGAAAAAGTTAAGCAGCGATTAATAGAGAT
TGCTAATCACGTTGACAAGTTTTACAGACCACTGAACATTCGGATCGTGTTGGTAGGCGTG
GAAGTGTGGAATGACATGGACAAATGCTCTGTAAGTCAGGACCCATTCACCAGCCTCCATG
AATTTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAAATCCCATGACAATGCGCAGCT
TGTCAGTGGGGTTTATTTCCAAGGGACCACCATCGGCATGGCCCCAATCATGAGCATGTGC
ACGGCAGACCAGTCTGGGGGAATTGTCATGGACCATTCAGACAATCCCCTTGGTGCAGCCG
TGACCCTGGCACATGAGCTGGGCCACAATTTCGGGATGAATCATGACACACTGGACAGGGG
CTGTAGCTGTCAAATGGCGGTTGAGAAAGGAGGCTGCATCATGAACGCTTCCACCGGGTAC
CCATTTCCCATGGTGTTCAGCAGTTGCAGCAGGAAGGACTTGGAGACCAGCCTGGAGAAAG
GAATGGGGGTGTGCCTGTTTAACCTGCCGGAAGTCAGGGAGTCTTTCGGGGGCCAGAAGTG
TGGGAACAGATTTGTGGAAGAAGGAGAGGAGTGTGACTGTGGGGAGCCAGAGGAATGTATG
AATCGCTGCTGCAATGCCACCACCTGTACCCTGAAGCCGGACGCTGTGTGCGCACATGGGC
TGTGCTGTGAAGACTGCCAGCTGAAGCCTGCAGGAACAGCGTGCAGGGACTCCAGCAACTC
CTGTGACCTCCCAGAGTTCTGCACAGGGGCCAGCCCTCACTGCCCAGCCAACGTGTACCTG
CACGATGGGCACTCATGTCAGGATGTGGACGGCTACTGCTACAATGGCATCTGCCAGACTC
ACGAGCAGCAGTGTGTCACACTCTGGGGACCAGGTGCTAAACCTGCCCCTGGGATCTGCTT
TGAGAGAGTCAATTCTGCAGGTGATCCTTATGGCAACTGTGGCAAAGTCTCGAAGAGTTCC
TTTGCCAAATGCGAGATGAGAGATGCTAAATGTGGAAAAATCCAGTGTCAAGGAGGTGCCA

GCCGGCCAGTCATTGGTACCAATGCCGTTTCCATAGAAACAAACATCCCCCTGCAGCAAGG
AGGCCGGATTCTGTGCCGGGGGACCCACGTGTACTTGGGCGATGACATGCCGGACCCAGGG
CTTGTGCTTGCAGGCACAAAGTGTGCAGATGGAAAAATCTGCCTGAATCGTCAATGTCAAA
ATATTAGTGTCTTTGGGGTTCACGAGTGTGCAATGCAGTGCCACGGCAGAGGGGTGTGCAA
CAACAGGAAGAACTGCCACTGCGAGGCCCACTGGGCACCTCCCTTCTGTGACAAGTTTGGC
TTTGGAGGAAGCACAGACAGCGGCCCCATCCGGCAAGCAGATAACCAAGGTTTAACCATAG
GAATTCTGGTGACCATCCTGTGTCTTCTTGCTGCCGGATTTGTGGTTTATCTCAAAAGGAA
GACCTTGATACGACTGCTGTTTACAAATAAGAAGACCACCATTGAAAAACTAAGGTGTGTG
CGCCCTTCCCGGCCACCCCGTGGCTTCCAACCCTGTCAGGCTCACCTCGGCCACCTTGGAA
AAGGCCTGATGAGGAAGCCGCCAGATTCCTACCCACCGAAGGACAATCCCAGGAGATTGCT
GCAGTGTCAGAATGTTGACATCAGCAGACCCCTCAACGGCCTGAATGTCCCTCAGCCCCAG
TCAACTCAGCGAGTGCTTCCTCCCCTCCACCGGGCCCCACGTGCACCTAGCGTCCCTGCCA
GACCCCTGCCAGCCAAGCCTGCACTTAGGCAGGCCCAGGGGACCTGTAAGCCAAACCCCCC
TCAGAAGCCTCTGCCTGCAGATCCTCTGGCCAGAACAACTCGGCTCACTCATGCCTTGGCC
AGGACCCCAGGACAATGGGAGACTGGGCTCCG

FIGURE 2B.

CCTGGCACCCCTCAGACCTGCTCCACAATATCCACACCAAGTGCCCAGATCCACCCACACC
GCCTATATTAAGTGAGAAGCCGACACCTTTTTTCAACAGTGAAGACAGAAGTTTGCACTAT
CTTTCAGCTCCAGTTGGAGTTTTTTGTACCAACTTTTAGGATTTTTTTTAATGTTTAAAAC
ATCATTACTATAAGAACTTTGAGCTACTGCCGTCAGTGCTGTGCTGTGCTATGGTGCTCTG
TCTACTTGCACAGGTACTTGTAAATTATTAATTTATGCAGAATGTTGATTACAGTGCAGTG
CGCTGTAGTAGGCATTTTTACCATCACTGAGTTTTCCATGGCAGGAAGGCTTGTTGTGCTT
TTAGTATTTTAGTGAACTTGAAATATCCTGCTTGATGGGATTCTGGACAGGATGTGTTTGC
TTTCTGATCAAGGCCTTATTGGAAAGCAGTCCCCCAACTACCCCCAGCTGTGCTTATGGTA
CCAGATGCAGCTCAAGAGATCCCAAGTAGAATCTCAGTTGATTTTCTGGATTCCCCATCTC
AGGCCAGAGCCAAGGGGCTTCAGGTCCAGGCTGTGTTTGGCTTTCAGGGAGGCCCTGTGCC
CCTTGACAACTGGCAGGCAGGCTCCCAGGGACACCTGGGAGAAATCTGGCTTCTGGCCAGG
AAGCTTTGGTGAGAACCTGGGTTGCAGACAGGAATCTTAAGGTGTAGCCACACCAGGATAG
AGACTGGAACACTAGACAAGCCAGAACTTGACCCTGAGCTGACCAGCCGTGAGCATGTTTG
GAAGGGGTCTGTAGTGTCACTCAAGGCGGTGCTTGATAGAAATGCCAAGCACTTCTTTTTC
TCGCTGTCCTTTCTAGAGCACTGCCACCAGTAGGTTATTTAGCTTGGGAAAGGTGGTGTTT
CTGTAAGAAACCTACTGCCCAGGCACTGCAAACCGCCACCTCCCTATACTGCTTGGAGCTG
AGCAAATCACCACAAACTGTAATACAATGATCCTGTATTCAGACAGATGAGGACTTTCCAT
GGGACCACAACTATTTTCAGATGTGAACCATTAACCAGATCTAGTCAATCAAGTCTGTTTA
CTGCAAGGTTCAACTTATTAACAATTAGGCAGACTCTTTATGCTTGCAAAAACTACAACCA
ATGGAATGTGATGTTCATGGGTATAGTTCATGTCTGCTATCATTATTCGTAGATATTGGAC
AAAGAACCTTCTCTATGGGGCATCCTCTTTTTCCAACTTGGCTGCAGGAATCTTTAAAAGA
TGCTTTTAACAGAGTCTGAACCTATTTCTTAAACACTTGCAACCTACCTGTTGAGCATCAC
AGAATGTGATAAGGAAATCAACTTGCTTATCAACTTCCTAAATATTATGAGATGTGGCTTG
GGCAGCATCCCCTTGAACTCTTCACTCTTCAAATGCCTGACTAGGGAGCCATGTTTCACAA
GGTCTTTAAAGTGACTAATGGCATGAGAAATACAAAAATACTCAGATAAGGTAAAATGCCA
TGATGCCTCTGTCTTCTGGACTGGTTTTCACATTAGAAGACAATTGACAACAGTTACATAA
TTCACTCTGAGTGTTTTATGAGAAAGCCTTCTTTTGGGGTCAACAGTTTTCCTATGCTTTG
AAACAGAAAAATATGTACCAAGAATCTTGGTTTGCCTTCCAGAAAACAAAACTGCATTTCA
CTTTCCCGGTGTTCCCCACTGTATCTAGGCAACATAGTATTCATGACTATGGATAAACTAA
ACACGTGACACAAACACACACAAAAGGGAACCCAGCTCTAATACATTCCAACTCGTATAGC
ATGCATCTGTTTATTCTATAGTTATTAAGTTCTTTAAAATGTAAAGCCATGCTGGAAAATA
ATACTGCTGAGATACATACAGAATTACTGTAACTGATTACACTTGGTAATTGTACTAAAGC
CAAACATATATATACTATTAAAAAGGTTTACAGAATTTTATGGTGCATTACGTGGGCATTG
TCTTTTTAGATGCCCAAATCCTTAGATCTGGCATGTTAGCCCTTCCTCCAATTATAAGAGG
ATATGAACCAAAAAAAAAAAAAAAAAAA

FIGURE 3

GGAGTTTCAAAGAAGCGCTAGTAAGGTCTCTGAGATCCTTGCACTAGCTACATCCTCAGGG
TAGGAGGAAGATGGCTTCCAGAAGCATGCGGCTGCTCCTATTGCTGAGCTGCCTGGCCAAA
ACAGGAGTCCTGGGTGATATCATCATGAGACCCAGCTGTGCTCCTGGATGGTTTTACCACA
AGTCCAATTGCTATGGTTACTTCAGGAAGCTGAGGAACTGGTCTGATGCCGAGCTCGAGTG
TCAGTCTTACGGAAACGGAGCCCACCTGGCATCTATCCTGAGTTTAAAGGAAGCCAGCACC
ATAGCAGAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTGGCCTGCACGACC
CACAGAAGAGGCAGCAGTGGCAGTGGATTGATGGGGCCATGTATCTGTACAGATCCTGGTC
TGGCAAGTCCATGGGTGGGAACAAGCACTGTGCTGAGATTAGCTCCAATAACAACTTTTTA
ACTTGGAGCAGCAACGAATGCAACAAGCGCCAACACTTCCTGTGCAAGTACCGACCATAGA
GCAAGAATCAAGATTCTGCTAACTCCTGCACAGCCCCGTCCTCTTCCTTTCTGCTAGCCTG
GCTAAATCTGCTCATTATTTCAGAGGGGAAACCTAGCAAACTAAGAGTGATAAGGGCCCTA
CTACACTGGCTTTTTTAGGCTTAGAGACAGAAACTTTAGCATTGGCCCAGTAGTGGCTTCT
AGCTCTAAATGTTTGCCCCGCCATCCCTTTCCACAGTATCCTTCTTCCCTCCTCCCCTGTC
TCTGGCTGTCTCGAGCAGTCTAGAAGAGTGCATCTCCAGCCTATGAAACAGCTGGGTCTTT
GGCCATAAGAAGTAAAGATTTGAAGACAGAAGGAAGAAACTCAGGAGTAAGCTTCTAGACC
CCTTCAGCTTCTACACCCTTCTGCCCTCTCTCCATTGCCTGCACCCCACCCCAGCCACTCA
ACTCCTGCTTGTTTTTCCTTTGGCCATAGGAAGGTTTACCAGTAGAATCCTTGCTAGGTTG
ATGTGGGCCATACATTCCTTTAATAAACCATTGTGTACATAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 4

GGGGAGCAGAGAGGAGGCAATGGCCACCATGGAGAACAAGGTGATCTGCGCCCTGGTCCTG
GTGTCCATGCTGGCCCTCGGCACCCTGGCCGAGGCCCAGACAGAGACGTGTACAGTGGCCC
CCCGTGAAAGACAGAATTGTGGTTTTCCTGGTGTCACGCCCTCCCAGTGTGCAAATAAGGG
CTGCTGTTTCGACGACACCGTTCGTGGGGTCCCCTGGTGCTTCTATCCTAATACCATCGAC
GTCCCTCCAGAAGAGGAGTGTGAATTTTAGACACTTCTGCAGGGATCTGCCTGCATCCTGA
CGCGGTGCCATCCCCAGCACGGTGATTAGTCCCAGAGCTCGGCTGCCACCTCCACCGGACA
CCTCAGACACGCTTCTGCAGCTGTGCCTCGGCTCACAACACAGATTGACTGCTCTGACTTT
GACTACTCAAAATTGGCCTAAAAATTAAAAGAGCTCGATATTAAAAA

FIGURE 5

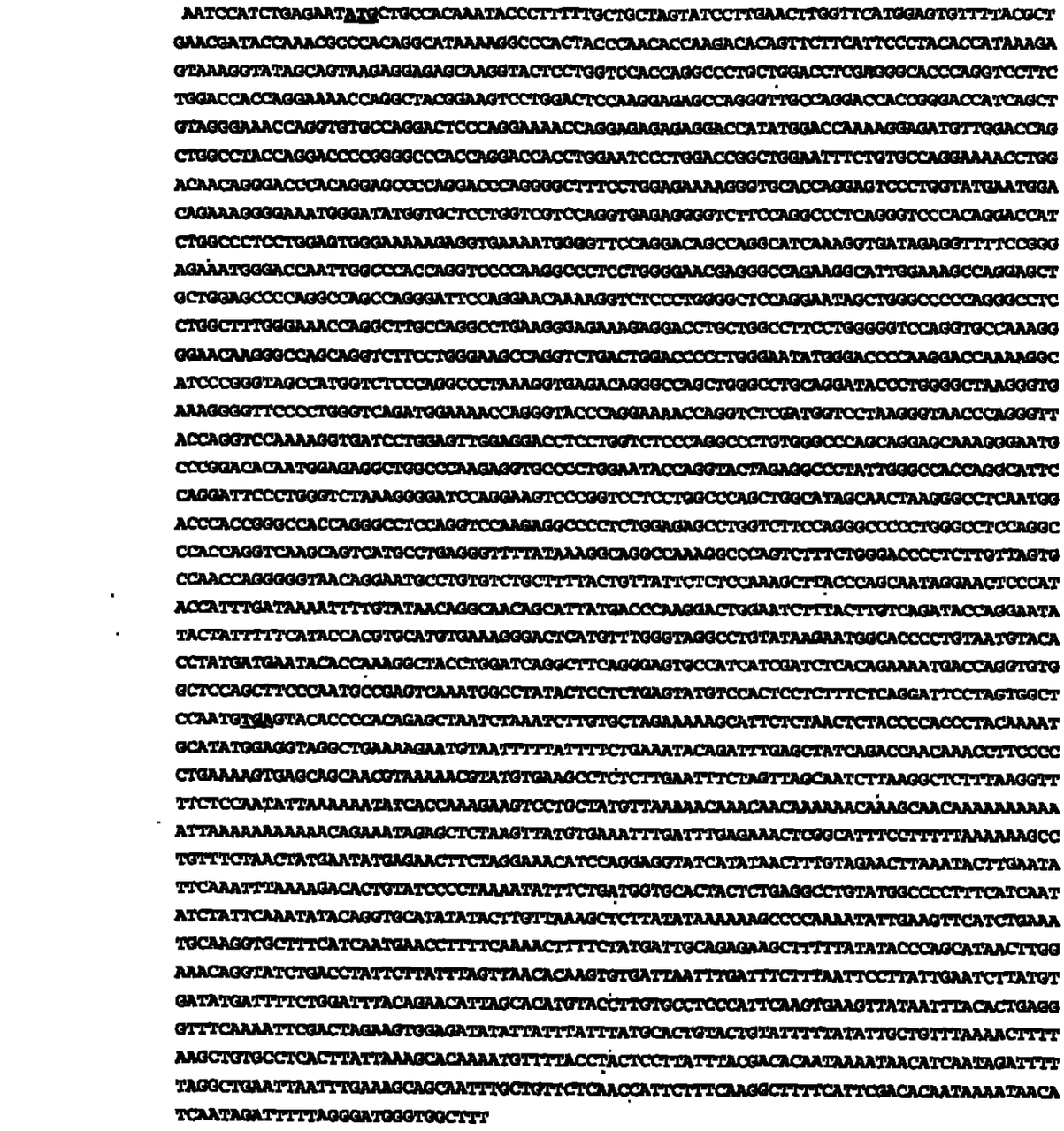


FIGURE 6

CATGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCC
TGTGTCTCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGA
ATGTCCTACAATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCA
GTATTTCATATATGGGCAAAAGAAATGGCTGAATAAATCAGAATGCAGAAATATCAATAGA
ACCTACTGTGATCTTTCTGCTGAAACTTCTGACTACGAACACCAGTATTATGCCAAAGTTA
AGGCCATTTGGGGAACAAAGTGTTCCAAATGGGCTGAAAGTGGACGGTTCTATCCTTTTTT
AGAAACACAAATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTT
GTCCTGACAGCTCCAGAGAAGTGGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAAC
AAATATACTCCAATCTGAAGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTG
GTCCCAGTGTGTGACCAACCACACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTAC
TGCGTACACGTGGAGTCCTTCGTCCCAGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGC
AGTGTGCCAGGACTTTGAAAGATCAATCATCAGAGTTCAAGGCTAAAATCATCTTCTGGTA
TGTTTTGCCCATATCTATTACCGTGTTTCTTTTTTCTGTGATGGGCTATTCCATCTACCGA
TATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTTGATTTTGATTTATGGAAATGAAT
TTGACAAAAGATTCTTTGTGCCTGCTGAAAAAATCGTGATTAACTTTATCACCCTCAATAT
CTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGAAAAAGCAGTGATGTA
TCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGAGGAAGAGGAGG
TGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAGAAAACAC
GGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAGCAGAACAATACCCCCGGATAAAACA
GTCATTGAATATGAATATGATGTCAGAACCACTGACATTTGTGCGGGGCCTGAAGAGCAGG
AGCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTT
GGCAGTCTTGGGCCCGCAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGAC
CCCCTGGCGCAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCC
TGGTCGACTGGGATCCCCAAACTGGCAGGCTGTGTATTCCTTCGCTGTCCAGCTTCGACCA
GGATTCAGAGGGCTGCGAGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCT
AGACTCTATGAGGAGCCGGCTCCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGC
AATTCATGGAGGAATGGGGGTTATATGTGCAGATGGAAAACTGATGCCAACACTTCCTTTT
GCCTTTTGTTTCCTGTGCAAACAAGTGAGTCACCCCTTTGATCCCAGCCATAAAGTACCTG
GGATGAAAGAAGTTTTTTCCAGTTTGTCAGTGTCTGTGAGAA

FIGURE 7

AGCGGGCGTGGGACTGAGCAGTCTGCTGCCCCCCGACATGTGACCCAGCCCCGCCGCCCAT
GCGGGCTCCCGGCCGCCCGGCCCTGCGGCCGCTGCCGCTGCCGCCGCTGCTGCTGTTGCTC
CTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGTGTCTCTGGTGGTTTGCCTAAACCTGCAA
ACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACAATGGACTCCACCAGAGGGTCT
TCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTCATATATGGGCAAAAGAAATGGCTG
AATAAATCAGAATGCAGAAATATCAATAGAACCTACTGTGATCTTTCTGCTGAAACTTCTG
ACTACGAACACCAGTATTATGCCAAAGTTAAGGCCATTTGGGGAACAAAGTGTTCCAAATG
GGCTGAAAGTGGACGGTTCTATCCTTTTTTAGAAACACAAATTGGCCCACCAGAGGTGGCA
CTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCCAGAGAAGTGGAAGAGAA
ATCCAGAAGACCTTCCTGTTTCCATGCAACAAATATACTCCAATCTGAAGTATAACGTGTC
TGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCACACGCTGGTG
CTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCTTCGTCCCAGGGC
CCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCATC
AGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTTCTT
TTTTCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAG
CAAATTTGATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAA
AATCGTGATTAACTTTATCACCCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGAT
ATGAGTTTACTGGGAAAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGA
ACCTGAGGCCCCCTCAGGAGGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGAT
GGAAATTTTTTGTGACTCTGAAGAAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCC
CTCAGCAGAACAATACCCCCGGATAAAACAGTCATTGAATATGAATATGATGTCAGAACCA
CTGACATTTGTGCGGGGCCTGAAGAGCAGGAGCTCAGTTTGCAGGAGGAGGTGTCCACACA
AGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAGTCTTGGGCCCGCAAACGTTACAGTAC
TCATACACCCCTCAGCTCCAAGACTTAGACCCCCTGGCGCAGGAGCACACAGACTCGGAGG
AGGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGATCCCCAAACTGGCAGGCT
GTGTATTCCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCGAGCCTTCTGAGGGG
GATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCTCCAGACAGGC
CACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGTTATATGTGCA
GATGGAAAACTGATGCCAACACTTCCTTTTGCCTTTTGTTTCCTGTGCAAACAAGTGAGTC
ACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTTCCAGTTTGTCAGT
GTCTGTGAGAA

FIGURE 8

GCTTGGTCTAGTGCCCATTTACTCTGGACTCCGGATGGCTGCCGCACGCCTCTGCCTCTCC
CTGCTGCTCCTGTCCACCTGCGTGGCTCTGTTACTACAGCCACTGCTGGGTGCCCAGGGAG
CCCCACTGGAGCCAGTGTACCCAGGGGACAATGCCACACCAGAGCAGATGGCCCAGTATGC
AGCTGATCTCCGTAGATACATCAACATGCTGACCAGGCCTAGGTATGGGAAAAGACACAAA
GAGGACACGCTGGCCTTCTCGGAGTGGGGGTCCCCGCATGCTGCTGTCCCCAGGGAGCTCA
GCCCGCTGGACTTATAATGCCACCTTCTGTCTCCTACGACTCCATGAGCAGCGCCAGCCCA
GCTCTCCCCTCTGCACCCTTGGCTCTGGCCAAAGCTTGCTCCCTGCTCCCACACAGGCTCA
ATAAAGCAAGTCAAAGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 9

GAAGGACAGCAGGGCCAACAGTCACAGCAGCCCTGACCAGAGCATTCCTGGAGCTCAAGCT
CCTCTACAAAGAGGTGGACAGAGAAGACAGCAGAGACCATGGGACCCCCCTCAGCCCCTCC
CTGCAGATTGCATGTCCCCTGGAAGGAGGTCCTGCTCACAGCCTCACTTCTAACCTTCTGG
AACCCACCCACCACTGCCAAGCTCACTATTGAATCCACGCCGTTCAATGTCGCAGAGGGGA
AGGAGGTTCTTCTACTCGCCCACAACCTGCCCCAGAATCGTATTGGTTACAGCTGGTACAA
AGGCGAAAGAGTGGATGGCAACAGTCTAATTGTAGGATATGTAATAGGAACTCAACAAGCT
ACCCCAGGGCCCGCATACAGTGGTCGAGAGACAATATACCCCAATGCATCCCTGCTGATCC
AGAACGTCACCCAGAATGACACAGGATTCTATACCCTACAAGTCATAAAGTCAGATCTTGT
GAATGAAGAAGCAACCGGACAGTTCCATGTATACCCGGAGCTGCCCAAGCCCTCCATCTCC
AGCAACAACTCCAACCCCGTGGAGGACAAGGATGCTGTGGCCTTCACCTGTGAACCTGAGG
TTCAGAACACAACCTACCTGTGGTGGGTAAATGGTCAGAGCCTCCCGGTCAGTCCCAGGCT
GCAGCTGTCCAATGGCAACATGACCCTCACTCTACTCAGCGTCAAAAGGAACGATGCAGGA
TCCTATGAATGTGAAATACAGAACCCAGCGAGTGCCAACCGCAGTGACCCAGTCACCCTGA
ATGTCCTCTATGGCCCAGATGGCCCCACCATTTCCCCCTCAAAGGCCAATTACCGTCCAGG
GGAAAATCTGAACCTCTCCTGCCACGCAGCCTCTAACCCACCTGCACAGTACTCTTGGTTT
ATCAATGGGACGTTCCAGCAATCCACACAAGAGCTCTTTATCCCCAACATCACTGTGAATA
ATAGCGGATCCTATATGTGCCAAGCCCATAACTCAGCCACTGGCCTCAATAGGACCACAGT
CACGATGATCACAGTCTCTGGAAGTGCTCCTGTCCTCTCAGCTGTGGCCACCGTCGGCATC
ACGATTGGAGTGCTGGCCAGGGTGGCTCTGATATAGCAGCCCTGGTGTATTTTCGATATTT
CAGGAAGACTGGCAGATTGGACCAGACCCTGAATTCTTCTAGCTCCTCCAATCCCATTTTA
TCCCATGGAACCACTAAAAACAAGGTCTGCTCTGCTCCTGAAGCCCTATATGCTGGAGATG
GACAACTCAATGAAAATTTAAAGGGAAAACCCTCAGGCCTGAGGTGTGTGCCACTCAGAGA
CTTCACCTAACTAGAGACAGGCAAACTGCAAACCATGGTGAGAAATTGACGACTTCACACT
ATGGACAGCTTTTCCCAAGATGTCAAAACAAGACTCCTCATCATGATAAGGCTCTTACCCC
CTTTTAATTTGTCCTTGCTTATGCCTGCCTCTTTCGCTTGGCAGGATGATGCTGTCATTAG
TATTTCACAAGAAGTAGCTTCAGAGGGTAACTTAACAGAGTATCAGATCTATCTTGTCAAT
CCCAACGTTTTACATAAAATAAGAGATCCTTTAGTGCACCCAGTGACTGACATTAGCAGCA
TCTTTAACACAGCCGTGTGTTCAAATGTACAGTGGTCCTTTTCAGAGTTGGACTTCTAGAC
TCACCTGTTCTCACTCCCTGTTTTAATTCAACCCAGCCATGCAATGCCAAATAATAGAATT
GCTCCCTACCAGCTGAACAGGGAGGAGTCTGTGCAGTTTCTGACACTTGTTGTTGAACATG
GCTAAATACAATGGGTATCGCTGAGACTAAGTTGTAGAAATTAACAAATGTGCTGCTTGGT
TAAAATGGCTACACTCATCTGACTCATTCTTTATTCTATTTTAGTTGGNTTGTATCTTGCC
TAAGGTGCGTAGTCCAACTCTTGGTATTACCCTCCTAATAGTCATACTAGTAGTCATACTC
CCTGGTGTAGTGTATTCTCTAAAAGCTTTAAATGTCTGCATGCAGCCAGCCATCAAATAGT
GAATGGGTCTTCTTTTGGCTGGAATTACAAAACTCAGAGAAATGTGTCATTAGGAGAACATC

ATAACCCATGAAGGATAAAAGCCCCAAATGGTGGTAACTGATAATAGCACTAATGCTTTAA
GATTTGGTCACACTCTCACCTAGGTGAGCGCATTGAGCCAGTGGTGCATAAATGCTACATA
CTCCAACTGAAATGTTAAGGAAGAAGATAGATCCAATTAAAAAAAATTAAAACCAATTTAA
AAAAAAAAAGAACACAGGAGATTCCAGTCTACTTGAGTTAGCATAATACAGAAGTCCCCTC
TACTTTAACTTTTACAAAAAAGTAACCTGAACTAATCTGATGTTAACCAATGTATTTATTT
CTGTGGTTCTGTTTCCTTGTTCCAATTTGACAAAACCCACTGTTCTTGTATTGTATTGCCC
AGGGGGAGCTATCACTGTACTTGTAGAGTGGTGCTGCTTTAATTCATAAATCACAAATAAA
AGCCAATTAGCTCTAAAAAAAAAAAA

FIGURE 10

GACCTGGGGGGGCGAGGGCCCCCCAAACTCAGTTCGGATCCTACCCGAGTGAGGCGGCGCC
ATGGAGCTCCGGGTGCTGCTCTGCTGGGCTTCGTTGGCCGCAGCTTTGGAAGAGACCCTGC
TGAACACAAAATTGGAAACTGCTGATCTGAAGTGGGTGACATTCCCTCAGGTGGACGGGCA
GTGGGAGGAACTGAGCGGCCTGGATGAGGAACAGCACAGCGTGCGCACCTACGAAGTGTGT
GACGTGCAGCGTGCCCCGGGCCAGGCCCACTGGCTTCGCACAGGTTGGGTCCCACGGCGGG
GCGCCGTCCACGTGTACGCCACGCTGCGCTTCACCATGCTCGAGTGCCTGTCCCTGCCTCG
GGCTGGGCGCTCCTGCAAGGAGACCTTCACCGTCTTCTACTATGAGAGCGATGCGGACACG
GCCACGGCCCTCACGCCAGCCTGGATGGAGAACCCCTACATCAAGGTGGACACGGTGGCCG
CGGAGCATCTCACCCGGAAGCGCCCTGGGGCCGAGGCCACCGGGAAGGTGAATGTCAAGAC
GCTGCGTCTGGGACCGCTCAGCAAGGCTGGCTTCTACCTGGCCTTCCAGGACCAGGGTGCC
TGCATGGCCCTGCTATCCCTGCACCTCTTCTACAAAAAGTGCGCCCAGCTGACTGTGAACC
TGACTCGATTCCCGGAGACTGTGCCTCGGGAGCTGGTTGTGCCCGTGGCCGGTAGCTGCGT
GGTGGATGCCGTCCCCGCCCCTGGCCCCAGCCCCAGCCTCTACTGCCGTGAGGATGGCCAG
TGGGCCGAACAGCCGGTCACGGGCTGCAGCTGTGCTCCGGGGTTCGAGGCAGCTGAGGGGA
ACACCAAGTGCCGAGCCTGTGCCCAGGGCACCTTCAAGCCCCTGTCAGGAGAAGGGTCCTG
CCAGCCATGCCCAGCCAATAGCCACTCTAACACCATTGGATCAGCCGTCTGCCAGTGCCGC
GTCGGGTACTTCCGGGCACGCACAGACCCCCGGGGTGCACCCTGCACCACCCCTCCTTCGG
CTCCGCGGAGCGTGGTTTCCCGCCTGAACGGCTCCTCCCTGCACCTGGAATGGAGTGCCCC
CCTGGAGTCTGGTGGCCGAGAGGACCTCACCTACGCCCTCCGCTGCCGGGAGTGCCGACCC
GGAGGCTCCTGTGCGCCCTGCGGGGGAGACCTGACTTTTGACCCCGGCCCCCGGGACCTGG
TGGAGCCCTGGGTGGTGGTTCGAGGGCTACGTCCTGACTTCACCTATACCTTTGAGGTCAC
TGCATTGAACGGGGTATCCTCCTTAGCCACGGGGCCCGTCCCATTTGACCTGTCAATGTC
ACCACTGACCGAGAGGTACCTCCTGCAGTGTCTGACATCCGGGTGACGCGGTCCTCACCCA
GCAGCTTGAGCCTGGCCTGGGCTGTTCCCCGGGCACCCAGTGGGGCTGTGCTGGACTACGA
GGTCAAATACCATGAGAAGGGCGCCGAGGGTCCCAGCAGCGTGCGGTTCCTGAAGACGTCA
GAAAACCGGGCAGAGCTGCGGGGGCTGAAGCGGGGAGCCAGCTACCTGGTGCAGGTACGGG
CGCGCTCTGAGGCCGGCTACGGGCCCTTCGGCCAGGAACATCACAGCCAGACCCAACTGGA
TGAGAGCGAGGGCTGGCGGGAGCAGCTGGCCCTGATTGCGGGCACGGCAGTCGTGGGTGTG
GTCCTGGTCCTGGTGGTCATTGTGGTCGCAGTTCTCTGCCTCAGGAAGCAGAGCAATGGGA
GAGAAGCAGAATATTCGGACAAACACGGACAGTATCTCATCGGACATGGTACTAAGGTCTA
CATCGACCCCTTCACTTATGAAGACCCTAATGAGGCTGTGAGGGAATTTGCAAAAGAGATC
GATGTCTCCTACGTCAAGATTGAAGAGGTGATTGGTGCAGGTGAGTTTGGCGAGGTGTGTC
GGGGGCGGCTCAAGGCCCCAGGGAAGAAGGAGAGCTGTGTGGCAATCAAGACCCTGAAGGG
TGGCTACACGGAGCGGCAGCGGCGTGAGTTTCTGAGCGAGGCCTCCATCATGGGCCAGTTC
GAGCACCCCAATATCATCCGCCTGGAGGGCGTGGTCACCAACAGCATGCCCGTCATGATTC

TCACAGAGTTCATGGAGAACGGCGCCCTGGACTCCTTCCTGCGGCTAAACGACGGACAGTT
CACAGTCATCCAGCTCGTGGGCATGCTGCGGGGCATCGCCTCGGGCATGCGGTACCTTGCC
GAGATGAGCTACGTCCACCGAGACCTGGCTGCTCGCAACATCCTAGTCAACAGCAACCTCG
TCTGCAAAGTGTCTGACTTTGGCCTTTCCCGATTCCTGGAGGAGAACTCTTCCGATCCCAC
CTACACGAGCTCCCTGGGAGGAAAGATTCCCATCCGATGGACTGCCCCGGAGGCCATTGCC
TTCCGGAAGTTCACTTCCGCCAGTGATGCCTGGAGTTACGGGATTGTGATGTGGGAGGTGA
TGTCATTTGGGGAGAGGCCGTACTGGGACATGAGCAATCAGGACGTGATCAATGCCATTGA
ACAGGACTACCGGCTGCCCCCGCCCCCAGACTGTCCCACCTCCCTCCACCAGCTCATGCTG
GACTGTTGGCAGAAAGACCGGAATGCCCGGCCCCGCTTCCCCCAGGTGGTCAGCGCCCTGG
ACAAGATGATCCGGAACCCCGCCAGCCTCAAAATCGTGGCCCGGGAGAATGGCGGGGCCTC
ACACCCTCTCCTGGACCAGCGGCAGCCTCACTACTCAGCTTTTGGCTCTGTGGGCGAGTGG
CTTCGGGCCATCAAAATGGGAAGATACGAAGAAAGTTTCGCAGCCGCTGGCTTTGGCTCCT
TCGAGCTGGTCAGCCAGATCTCTGCTGAGGACCTGCTCCGAATCGGAGTCACTCTGGCGGG
ACACCAGAAGAAAATCTTGGCCAGTGTCCAGCACATGAAGTCCCAGGCCAAGCCGGGAACC
CCGGGTGGGACAGGAGGACCGGCCCCGCAGTACTGACCTGCAGGAACTCCCCACCCCAGGG
ACACCGCCTCCCCATTTTCCGGGGCAGAGTGGGGACTCACAGAGGCCCCCAGCCCTGTGCC
CCGCTGGATTGCACTTTGAGCCCGTGGGGTGAGGAGTTGGCAATTTGGAGAGACAGGATTT
GGGGGTTCTGCCATAATAGGAGGGGAAAATCACCCCCCAGCCACCTCGGGGAACTCCAGAC
CAAGGGTGAGGGCGCCTTTCCCTCAGGACTGGGTGTGACCAGAGGAAAAGGAAGTGCCCAA
CATCTCCCAGCCTCCCCAGGTGCCCCCCTCACCTTGATGGGTGCGTTCCCGCAGACCAAAG
AGAGTGTGACTCCCTTGCCAGCTCCAGAGTGGGGGGGCTGTCCCAGGGGGCAAGAAGGGGT
GTCAGGGCCCAGTGACAAAATCATTGGGGTTTGTAGTCCCAACTTGCTGCTGTCACCACCA
AACTCAATCATTTTTTTCCCTTGTAAATGCCCCTCCCCCAGCTGCTGCCTTCATATTGAAG
GTTTTTGAGTTTTGTTTTTGGTCTTAATTTTTCTCCCCGTTCCCTTTTTGTTTCTTCGTTT
TGTTTTTCTACCGTCCTTGTCATAACTTTGTGTTGGAGGGAACCTGTTTCACTATGGCCTC
CTTTGCCCAAGTTGAAACAGGGGCCCATCATCATGTCTGTTTCCAGAACAGTGCCTTGGTC
ATCCCACATCCCCGGACCCCGCCTGGGACCCCCAAGCTGTGTCCTATGAAGGGGTGTGGGG
TGAGGTAGTGAAAAGGGCGGTAGTTGGTGGTGGAACCCAGAAACGGACGCCGGTGCTTGGA
GGGGTTCTTAAATTATATTTAAAAAAGTAACTTTTTGTATAAATAAAAGAAAATGGGACGT
GTCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGA

FIGURE 11

GCTCTATGCCGCCTACCTTGCTCTCGCCGCTGCTGCCGGAGCCGAAGCAGAGAAGGCAGCG
GGTCCCGTGACCGTCCCGAGAGCCCCGCGCTCCCGACCAGGGGGCGGGGGCGGCCCCGGGG
AGGGCGGGGCAGGGGCGGGGGGAAGAAAGGGGGTTTTGTGCTGCGCCGGGAGGGCCGGCGC
CCTCTTCCGAATGTCCTGCGGCCCCAGCCTCTCCTCACGCTCGCGCAGTCTCCGCCGCAGT
CTCAGCTGCAGCTGCAGGACTGAGCCGTGCACCCGGAGGAGACCCCCGGAGGAGGCGACAA
ACTTCGCAGTGCCGCGACCCAACCCCAGCCCTGGGTAGCCTGCAGCATGGCCCAGCTGTTC
CTGCCCCTGCTGGCAGCCCTGGTCCTGGCCCAGGCTCCTGCAGCTTTAGCAGATGTTCTGG
AAGGAGACAGCTCAGAGGACCGCGCTTTTCGCGTGCGCATCGCGGGCGACGCGCCACTGCA
GGGCGTGCTCGGCGGCGCCCTCACCATCCCTTGCCACGTCCACTACCTGCGGCCACCGCCG
AGCCGCCGGGCTGTGCTGGGCTCTCCGCGGGTCAAGTGGACTTTCCTGTCCCGGGGCCGGG
AGGCAGAGGTGCTGGTGGCGCGGGGAGTGCGCGTCAAGGTGAACGAGGCCTACCGGTTCCG
CGTGGCACTGCCTGCGTACCCAGCGTCGCTCACCGACGTCTCCCTGGCGCTGAGCGAGCTG
CGCCCCAACGACTCAGGTATCTATCGCTGTGAGGTCCAGCACGGCATCGATGACAGCAGCG
ACGCTGTGGAGGTCAAGGTCAAAGGGGTCGTCTTTCTCTACCGAGAGGGCTCTGCCCGCTA
TGCTTTCTCCTTTTCTGGGGCCCAGGAGGCCTGTGCCCGCATTGGAGCCCACATCGCCACC
CCGGAGCAGCTCTATGCCGCCTACCTTGGGGGCTATGAGCAATGTGATGCTGGCTGGCTGT
CGGATCAGACCGTGAGGTATCCCATCCAGACCCCACGAGAGGCCTGTTACGGAGACATGGA
TGGCTTCCCCGGGGTCCGGAACTATGGTGTGGTGGACCCGGATGACCTCTATGATGTGTAC
TGTTATGCTGAAGACCTAAATGGAGAACTGTTCCTGGGTGACCCTCCAGAGAAGCTGACAT
TGGAGGAAGCACGGGCGTACTGCCAGGAGCGGGGTGCAGAGATTGCCACCACGGGCCAACT
GTATGCAGCCTGGGATGGTGGCCTGGACCACTGCAGCCCAGGGTGGCTAGCTGATGGCAGT
GTGCGCTACCCCATCGTCACACCCAGCCAGCGCTGTGGTGGGGGCTTGCCTGGTGTCAAGA
CTCTCTTCCTCTTCCCCAACCAGACTGGCTTCCCCAATAAGCACAGCCGCTTCAACGTCTA
CTGCTTCCGAGACTCGGCCCAGCCTTCTGCCATCCCTGAGGCCTCCAACCCAGCCTCCAAC
CCAGCCTCTGATGGACTAGAGGCTATCGTCACAGTGACAGAGACCCTGGAGGAACTGCAGC
TGCCTCAGGAAGCCACAGAGAGTGAATCCCGTGGGGCCATCTACTCCATCCCCATCATGGA
GGACGGAGGAGGTGGAAGCTCCACTCCAGAAGACCCAGCAGAGGCCCCTAGGACGCTCCTA
GAATTTGAAACACAATCCATGGTACCGCCCACGGGGTTCTCAGAAGAGGAAGGTAAGGCAT
TGGAGGAAGAAGAGAAATATGAAGATGAAGAAGAGAAAGAGGAGGAAGAAGAAGAGGAGGA
GGTGGAGGATGAGGCTCTGTGGGCATGGCCCAGCGAGCTCAGCAGCCCGGGCCCTGAGGCC
TCTCTCCCCACTGAGCCAGCAGCCCAGGAGAAGTCACTCTCCCAGGCGCCAGCAAGGGCAG
TCCTGCAGCCTGGTGCATCACCACTTCCTGATGGAGAGTCAGAAGCTTCCAGGCCTCCAAG
GGTCCATGGACCACCTACTGAGACTCTGCCCACTCCCAGGGAGAGGAACCTAGCATCCCCA
TCACCTTCCACTCTGGTTGAGGCAAGAGAGGTGGGGGAGGCAACTGGTGGTCCTGAGCTAT
CTGGGGTCCCTCGAGGAGAGAGCGAGGAGACAGGAAGCTCCGAGGGTGCCCCTTCCCTGCT

TCCAGCCACACGGGCCCCTGAGGGTACCAGGGAGCTGGAGGCCCCCTCTGAAGATAATTCT
GGAAGAACTGCCCCAGCAGGGACCTCAGTGCAGGCCCAGCCAGTGCTGCCGACTGACAGCG
CCAGCCGAGGTGGAGTGGCCGTGGTCCCCGCATCAGGTGACTGTGTCCCCAGCCCCTGCCA
CAATGGTGGGACATGCTTGGAGGAGGAGGAAGGGGTCCGCTGCCTATGTCTGCCTGGCTAT
GGGGGGGACCTGTGCGATGTTGGCCTCCGCTTCTGCAACCCCGGCTGGGACGCCTTCCAGG
GCGCCTGCTACAAGCACTTTTCCACACGAAGGAGCTGGGAGGAGGCAGAGACCCAGTGCCG
GATGTACGGCGCGCATCTGGCCAGCATCAGCACACCCGAGGAACAGGACTTCATCAACAAC
CGGTACCGGGAGTACCAGTGGATCGGACTCAACGACAGGACCATCGAAGGCGACTTCTTGT
GGTCGGATGGCGTCCCCCTGCTCTATGAGAACTGGAACCCTGGGCAGCCTGACAGCTACTT
CCTGTCTGGAGAGAACTGCGTGGTCATGGTGTGGCATGATCAGGGACAATGGAGTGACGTG
CCCTGCAACTACCACCTGTCCTACACCTGCAAGATGGGGCTGGTGTCCTGTGGGCCGCCAC
CGGAGCTGCCCCTGGCTCAAGTGTTCGGCCGCCCACGGCTGCGCTATGAGGTGGACACTGT
GCTTCGCTACCGGTGCCGGGAAGGACTGGCCCAGCGCAATCTGCCGCTGATCCGATGCCAA
GAGAACGGTCGTTGGGAGGCCCCCCAGATCTCCTGTGTGCCCAGAAGACCTGCCCGAGCTC
TGCACCCAGAGGAGGACCCAGAAGGACGTCAGGGGAGGCTACTGGGACGCTGGAAGGCGCT
GTTGATCCCCCCTTCCAGCCCCATGCCAGGTCCCTAGGGGGCAAGGCCTTGAACACTGCCG
GCCACAGCACTGCCCTGTCACCCAAATTTTCCCTCACACCTTGCGCTCCCGCCACCACAGG
AAGTGACAACATGACGAGGGGTGGTGCTGGAGTCCAGGTGACAGTTCCTGAAGGGGCTTCT
GGGAAATACCTAGGAGGCTCCAGCCCAGCCCAGGCCCTCTCCCCCTACCCTGGGCACCAGA
TCTTCCATCAGGGCCGGAGTAAATCCCTAAGTGCCTCAACTGCCCTCTCCCTGGCAGCCAT
CTTGTCCCCTCTATTCCTCTAGGGAGCACTGTGCCCACTCTTTCTGGGTTTTCCAAGGGAA
TGGGCTTGCAGGATGGAGTGTCTGTAAAATCAACAGGAAATAAAACTGTGTATGAGCCCA

FIGURE 12

TGTGGCACTGCCTGCGTACCCAACCCCAGCCCTGGGTAGCCTGCAGCATGGCCCAGCTGTT
CCTGCCCCTGCTGGCAGCCCTGGTCCTGGCCCAGGCTCCTGCAGCTTTAGCAGATGTTCTG
GAAGGAGACAGCTCAGAGGACCGCGCTTTTCGCGTGCGCATCGCGGGCGACGCGCCACTGC
AGGGCGTGCTCGGCGGCGCCCTCACCATCCCTTGCCACGTCCACTACCTGCGGCCACCGCC
GAGCCGCCGGGCTGTGCTGGGCTCTCCGCGGGTCAAGTGGACTTTCCTGTCCCGGGGCCGG
GAGGCAGAGGTGCTGGTGGCGCGGGGAGTGCGCGTCAAGGTGAACGAGGCCTACCGGTTCC
GCGTGGCACTGCCTGCGTACCCAGCGTCGCTCACCGACGTCTCCCTGGCGCTGAGCGAGCT
GCGCCCCAACGACTCAGGTATCTATCGCTGTGAGGTCCAGCACGGCATCGATGACAGCAGC
GACGCTGTGGAGGTCAAGGTCAAAGGGGTCGTCTTTCTCTACCGAGAGGGCTCTGCCCGCT
ATGCTTTCTCCTTTTCTGGGGCCCAGGAGGCCTGTGCCCGCATTGGAGCCCACATCGCCAC
CCCGGAGCAGCTCTATGCCGCCTACCTTGGGGGCTATGAGCAATGTGATGCTGGCTGGCTG
TCGGATCAGACCGTGAGGTATCCCATCCAGACCCCACGAGAGGCCTGTTACGGAGACATGG
ATGGCTTCCCCGGGGTCCGGAACTATGGTGTGGTGGACCCGGATGACCTCTATGATGTGTA
CTGTTATGCTGAAGACCTAAATGGAGAATTGTTCCTGGGTGACCCTCCAGAGAAGCTGACA
TTGGAGGAAGCACGGGCGTACTGCCAGGAGCGGGGTGCAGAGATTGCCACCACGGGCCAAC
TGTATGCAGCCTGGGATGGTGGCCTGGACCACTGCAGCCCAGGGTGGCTAGCTGATGGCAG
TGTGCGCTACCCCATCGTCACACCCAGCCAGCGCTGTGGTGGGGGCTTGCCTGGTGTCAAG
ACTCTCTTCCTCTTCCCCAACCAGACTGGCTTCCCCAATAAGCACAGCCGCTTCAACGTCT
ACTGCTTCCGAGACTCGGCCCAGCCTTCTGCCATCCCTGAGGCCTCCAACCCAGCCTCCAA
CCCAGCCTCTGATGGACTAGAGGCTATCGTCACAGTGACAGAGACCCTGGAGGAACTGCAG
CTGCCTCAGGAAGCCACAGAGAGTGAATCCCGTGGGGCCATCTACTCCATCCCCATCATGG
AGGACGGAGGAGGTGGAAGCTCCACTCCAGAAGACCCAGCAGAGGCCCCTAGGACGCTCCT
AGAATTTGAAACACAATCCATGGTACCGCCCACGGGGTTCTCAGAAGAGGAAGGTAAGGCA
TTGGAGGAAGAAGAGAAATATGAAGATGAAGAAGAGAAAGAGGAGGAAGAAGAAGAGGAGG
AGGTGGAGGATGAGGCTCTGTGGGCATGGCCCAGCGAGCTCAGCAGCCCGGGCCCTGAGGC
CTCTCTCCCCACTGAGCCAGCAGCCCAGGAGGAGTCACTCTCCCAGGCGCCAGCAAGGGCA
GTCCTGCAGCCTGGTGCATCACCACTTCCTGATGGAGAGTCAGAAGCTTCCAGGCCTCCAA
GGGTCCATGGACCACCTACTGAGACTCTGCCCACTCCCAGGGAGAGGAACCTAGCATCCCC
ATCACCTTCCACTCTGGTTGAGGCAAGAGAGGTGGGGGAGGCAACTGGTGGTCCTGAGCTA
TCTGGGGTCCCTCGAGGAGAGAGCGAGGAGACAGGAAGCTCCGAGGGTGCCCCTTCCCTGC
TTCCAGCCACACGGGCCCCTGAGGGTACCAGGGAGCTGGAGGCCCCCTCTGAAGATAATTC
TGGAAGAACTGCCCCAGCAGGGACCTCAGTGCAGGCCCAGCCAGTGCTGCCCACTGACAGC
GCCAGCCGAGGTGGAGTGGCCGTGGTCCCCGCATCAGGTAATTCTGCCCAAGGCTCAACTG
CCCTCTCTATCCTACTCCTTTTCTTCCCCCTGCAGCTCTGGGTCACCTGACCTGTAGTCCT

CTATGGGTCTCCAATCTCGGATATCCACCTTGTGGGTATCTCAGCTCTCCGCGTCTTTACC
CTGTGATCCCAGCCCCGCCACTGACCATCTGTGACCCTTCCCTGCCATTGGGCCCTCCACC
TGTGGCTCACATCTCGCCAGCCCCACAGAGCATCCTCAGGCCTCTCCAAGGGTCCTCATCA
CCTATTGCAGCCTTCAGGGCTCGGCCTATTTTCCACTACTCCCTTCATCCGCCTGTGTGCC
GTCCCCTTTAGCTGCCTCCTATTGATCTCAGGGAAGCCTGGGAGTCCCTTCTCACCCCTCA
ACCTCCGGAGTCCAGGAGAACCCGTACCCCCACAGAGCCTTAAGCAACTACTTCTGTGAAG
TATTTTTTGACTGTTTCATGGAAAACAAGCCTTGGAAATAAATCTCTATTAAACCGC

FIGURE 13

AGCCAGCTCAGGCTACACTATCCCAGGATCAGCATGGCCGTCCGCCAGTGGGTAATCGCCC
TGGCCTTGGCTGCCCTCCTTGTTGTGGACAGGGAAGTGCCAGTGGCAGCAGGAAAGCTCCC
TTTCTCAAGAATGCCCATCTGTGAACACATGGTAGAGTCTCCAACCTGTTCCCAGATGTCC
AACCTGGTCTGCGGCACTGATGGGCTCACATATACGAATGAATGCCAGCTCTGCTTGGCCC
GGATAAAAACCAAACAGGACATCCAGATCATGAAAGATGGCAAATGCTGATCCCACAGGAG
CACCTCAAGCCATGAAGTGTCAGCTGGAGAACAGTGGTGGGCATGGAGAGGATATGACATG
AAATAAAAGATCCAGCCCAAAAAAAAAAGAAAAAA

FIGURE 14

CGTTTCCAATGCACGTACAGCCCGTACACACCGTGTGCTGGGACACCCCACAGTCAGCCGC
ATGGCTCCCCTGTGCCCCAGCCCCTGGCTCCCTCTGTTGATCCCGGCCCCTGCTCCAGGCC
TCACTGTGCAACTGCTGCTGTCACTGCTGCTTCTGGTGCCTGTCCATCCCCAGAGGTTGCC
CCGGATGCAGGAGGATTCCCCCCTGGGAGGAGGCTCTTCTGGGGAAGATGACCCACTGGGC
GAGGAGGATCTGCCCAGTGAAGAGGATTCACCCAGAGAGGAGGATCCACCCGGAGAGGAGG
ATCTACCTGGAGAGGAGGATCTACCTGGAGAGGAGGATCTACCTGAAGTTAAGCCTAAATC
AGAAGAAGAGGGCTCCCTGAAGTTAGAGGATCTACCTACTGTTGAGGCTCCTGGAGATCCT
CAAGAACCCCAGAATAATGCCCACAGGGACAAAGAAGGGGATGACCAGAGTCATTGGCGCT
ATGGAGGCGACCCGCCCTGGCCCCGGGTGTCCCCAGCCTGCGCGGGCCGCTTCCAGTCCCC
GGTGGATATCCGCCCCCAGCTCGCCGCCTTCTGCCCGGCCCTGCGCCCCCTGGAACTCCTG
GGCTTCCAGCTCCCGCCGCTCCCAGAACTGCGCCTGCGCAACAATGGCCACAGTGTGCAAC
TGACCCTGCCTCCTGGGCTAGAGATGGCTCTGGGTCCCGGGCGGGAGTACCGGGCTCTGCA
GCTGCATCTGCACTGGGGGGCTGCAGGTCGTCCGGGCTCGGAGCACACTGTGGAAGGCCAC
CGTTTCCCTGCCGAGATCCACGTGGTTCACCTCAGCACCGCCTTTGCCAGAGTTGACGAGG
CCTTGGGGCGCCCGGGAGGCCTGGCCGTGTTGGCCGCCTTTCTGGAGGAGGGCCCGGAAGA
AAACAGTGCCTATGAGCAGTTGCTGTCTCGCTTGGAAGAAATCGCTGAGGAAGGCTCAGAG
ACTCAGGTCCCAGGACTGGACATATCTGCACTCCTGCCCTCTGACTTCAGCCGCTACTTCC
GATATGAGGGGTCTCTGACTACACCGCCCTGTGCCCAGGGTGTCATCTGGACTGTGTTTAA
CCAGACAGTGATGCTGAGTGCTAAGCAGCTCCACACCCTCTCTGACACCCTGTGGGGACCT
GGTGACTCTCGGCTACAGCTGAACTTCCGAGCGACGCAGCCTTTGAATGGGCGAGTGATTG
AGGCCTCCTTCCCTGCTGGAGTGGACAGCAGTCCTCGGGCTGCTGAGCCAGTCCAGCTGAA
TTCCTGCCTGGCTGCTGGTGACATCCTAGCCCTGGTTTTTGGCCTCCTTTTTGCTGTCACC
AGCGTCGCGTTCCTTGTGCAGATGAGAAGGCAGCACAGAAGGGGAACCAAAGGGGGTGTGA
GCTACCGCCCAGCAGAGGTAGCCGAGACTGGAGCCTAGAGGCTGGATCTTGGAGAATGTGA
GAAGCCAGCCAGAGGCATCTGAGGGGGAGCCGGTAACTGTCCTGTCCTGCTCATTATGCCA
CTTCCTTTTAACTGCCAAGAAATTTTTTAAAATAAATATTTATAATA

FIGURE 15

TCCACATCCAGGTCTTTGTGCTCCTCGCTTGCCTGTTCCTTTTCCACGCATTTTCCAGGAT
AACTGTGACTCCAGGCCCGCAATGGATGCCCTGCAACTAGCAAATTCGGCTTTTGCCGTTG
ATCTGTTCAAACAACTATGTGAAAAGGAGCCACTGGGCAATGTCCTCTTCTCTCCAATCTG
TCTCTCCACCTCTCTGTCACTTGCTCAAGTGGGTGCTAAAGGTGACACTGCAAATGAAATT
GGACAGGTTCTTCATTTTGAAAATGTCAAAGATGTACCCTTTGGATTTCAAACAGTAACAT
CGGATGTAAACAAACTTAGTTCCTTTTACTCACTGAAACTAATCAAGCGGCTCTACGTAGA
CAAATCTCTGAATCTTTCTACAGAGTTCATCAGCTCTACGAAGAGACCGTATGCAAAGGAA
TTGGAAACTGTTGACTTCAAAGATAAATTGGAAGAAACGAAAGGTCAGATCAACAACTCAA
TTAAGGATCTCACAGATGGCCACTTTGAGAACATTTTAGCTGACAACAGTGTGAACGACCA
GACCAAAATCCTTGTGGTTAATGCTGCCTACTTTGTTGGCAAGTGGATGAAGAAATTTTCT
GAATCAGAAACAAAAGAATGTCCTTTCAGAGTCAACAAGACAGACACCAAACCAGTGCAGA
TGATGAACATGGAGGCCACGTTCTGTATGGGAAACATTGACAGTATCAATTGTAAGATCAT
AGAGCTTCCTTTTCAAAATAAGCATCTCAGCATGTTCATCCTACTACCCAAGGATGTGGAG
GATGAGTCCACAGGCTTGGAGAAGATTGAAAAACAACTCAACTCAGAGTCACTGTCACAGT
GGACTAATCCCAGCACCATGGCCAATGCCAAGGTCAAACTCTCCATTCCAAAATTTAAGGT
GGAAAAGATGATTGATCCCAAGGCTTGTCTGGAAAATCTAGGGCTGAAACATATCTTCAGT
GAAGACACATCTGATTTCTCTGGAATGTCAGAGACCAAGGGAGTGGCCCTATCAAATGTTA
TCCACAAAGTGTGCTTAGAAATAACTGAAGATGGTGGGGATTCCATAGAGGTGCCAGGAGC
ACGGATCCTGCAGCACAAGGATGAATTGAATGCTGACCATCCCTTTATTTACATCATCAGG
CACAACAAAACTCGAAACATCATTTTCTTTGGCAAATTCTGTTCTCCTTAAGTGGCATAGC
CCATGTTAAGTCCTCCCTGACTTTTCTGTGGATGCCGATTTCTGTAAACTCTGCATCCAGA
GATTCATTTTCTAGATACAATAAATTGCTAATGTTGCTGGATCAGGAAGCCGCCAGTACTT
GTCATATGTAGCCTTCACACAGATAGACCTTTTTTTTTTTTTCCAATTCTATCTTTTGTTTC
CTTTTTTCCCATAAGACAATGACATACGCTTTTAATGAAAAGGAATCACGTTAGAGGAAAA
ATATTTATTCATTATTTGTCAAATTGTCCGGGTTAGTTGGCAGAAATACAGTCTTCCACAA
AGAAAATTCCTATAAGGAAGATTTGGAAGCTCTTCTTCCCAGCACTATGCTTTCCTTCTTT
GGGATAGAGAATGTTCCAGACATTCTCGCTTCCCTGAAAGACTGAAGAAAGTGTAGTGCAT
GGGACCCACGAAACTGCCCTGGCTCCAGTGAAACTTGGGCACATGCTCAGGCTACTATAGG
TCCAGAAGTCCTTATGTTAAGCCCTGGCAGGCAGGTGTTTATTAAAATTCTGAATTTTGGG
GATTTTCAAAAGATAATATTTTACATACACTGTATGTTATAGAACTTCATGGATCAGATCT
GGGGCAGCACCCTATAAATCAACACCTTAATATGCTGCAACAAAATGTAGAATATTCAGAC
AAAATGGATACATAAAGACTAAGTAGCCCATAAGGGGTCAAAATTTGCTGCCAAATGCGTA
TGCCACCAACTTACAAAAACACTTCGTTCGCAGAGCTTTTCAGATTGTGGAATGTTGGATA
AGGAATTATAGACCTCTAGTAGCTGAAATGCAAGACCCCAAGAGGAAGTTCAGATCTTAAT

TGGCAGGGGCTTCTAGCTGACTTGCACAGGGATTCTCACAATAGCCGATATCAGAATTTGT
GTTGAAGGAACTTGTCTCTTCATCTAATATGATAGCGGGAAAAGGAGAGGAAACTACTGCC
TTTAGAAAATATAAGTAAAGTGATTAAAGTGCTCACGTTACCTTGACACATAGTTTTTCAG
TCTATGGGTTTAGTTACTTTAGATGGCAAGCATGTAACTTATATTAATAGTAATTTGTAAA
GTTGGTTGGATAAGCTATCCGTGTTGCAGGTTCATGGATTACTTCTCTATAAAAAATATGT
ATTTACCAAAAATTTTGTGACATTCCTTCTCCCATCTCTTCCTTGACCTGCATTGTAAATA
GGTTCTTCTTGTTCTGAGATTCAATATTGAATTTTTCCTATGCTATTGACAATAAAATATT
ATTGAACTACA

FIGURE 16

GCCGAGCCGGGGCTCCGGAAGCCGGCGGGGGCGCCGCGGCCGTGCGGGGCGTCAATGGATC
GCCACTCCAGCTACATCTTCATCTGGCTGCAGCTGGAGCTCTGCGCCATGGCCGTGCTGCT
CACCAAAGGTGAAATTCGATGCTACTGTGATGCTGCCCACTGTGTAGCCACTGGTTATATG
TGTAAATCTGAGCTCAGCGCCTGCTTCTCTAGACTTCTTGATCCTCAGAACTCAAATTCCC
CACTCACCCATGGCTGCCTGGACTCTCTTGCAAGCACGACAGACATCTGCCAAGCCAAACA
GGCCCGAAACCACTCTGGCACCACCATACCCACATTGGAATGCTGTCATGAAGACATGTGC
AATTACAGAGGGCTGCACGATGTTCTCTCTCCTCCCAGGGGTGAGGCCTCAGGACAAGGAA
ACAGGTATCAGCATGATGGTAGCAGAAACCTTATCACCAAGGTGCAGGAGCTGACTTCTTC
CAAAGAGTTGTGGTTCCGGGCAGCGGTCATTGCCGTGCCCATTGCTGGAGGGCTGATTTTA
GTGTTGCTTATTATGTTGGCCCTGAGGATGCTTCGAAGTGAAAATAAGAGGCTGCAGGATC
AGCGGCAACAGATGCTCTCCCGTTTGCACTACAGCTTTCACGGACACCATTCCAAAAAGGG
GCAGGTTGCAAAGTTAGACTTGGAATGCATGGTGCCGGTCAGTGGGCACGAGAACTGCTGT
CTGACCTGTGATAAAATGAGACAAGCAGACCTCAGCAACGATAAGATCCTCTCGCTTGTTC
ACTGGGGCATGTACAGTGGGCACGGGAAGCTGGAATTCGTATGACGGAGTCTTATCTGAAC
TACACTTACTGAACAGCTTGAAGGCCTTTTGAGTTCTGCTGGACAGGAGCACTTTATCTGA
AGACAAACTCATTTAATCATCTTTGACAGACAAAATGACCTCTGCAAACAGAATCTTGGAT
ATTTCTTCTGAAGGATTATTTGCACAGACTTAAATACAGTTAAATGTGTTATTTGCTTTTA
AAATTATAAAAAGCAAAGAGAAGACTTTGTACACACTGTCACCAGGGTTATTTGCATCCAA
GGGAGCTGGAATTGAGTACCTAAATAAACAAAAATGTGCCCTAAAAA

FIGURE 17A

GTCACTTAGGAAAAGGTGTCCTTTCGGGCAGCCGGGCTCAGCATGAGGAACAGAAGGAATG
ACACTCTGGACAGCACCCGGACCCTGTACTCCAGCGCGTCTCGGAGCACAGACTTGTCTTA
CAGTGAAAGCGACTTGGTGAATTTTATTCAAGCAAATTTTAAGAAACGAGAATGTGTCTTC
TTTACCAAAGATTCCAAGGCCACGGAGAATGTGTGCAAGTGTGGCTATGCCCAGAGCCAGC
ACATGGAAGGCACCCAGATCAACCAAAGTGAGAAATGGAACTACAAGAAACACACCAAGGA
ATTTCCTACCGACGCCTTTGGGGATATTCAGTTTGAGACACTGGGGAAGAAAGGGAAGTAT
ATACGTCTGTCCTGCGACACGGACGCGGAAATCCTTTACGAGCTGCTGACCCAGCACTGGC
ACCTGAAAACACCCAACCTGGTCATTTCTGTGACCGGGGGCGCCAAGAACTTCGCCCTGAA
GCCGCGCATGCGCAAGATCTTCAGCCGGCTCATCTACATCGCGCAGTCCAAAGGTGCTTGG
ATTCTCACGGGAGGCACCCATTATGGCCTGACGAAGTACATCGGGGAGGTGGTGAGAGATA
ACACCATCAGCAGGAGTTCAGAGGAGAATATTGTGGCCATTGGCATAGCAGCTTGGGGCAT
GGTCTCCAACCGGGACACCCTCATCAGGAATTGCGATGCTGAGGGCTATTTTTTAGCCCAG
TACCTTATGGATGACTTCACAAGGGATCCACTGTATATCCTGGACAACAACCACACACATT
TGCTGCTCGTGGACAATGGCTGTCATGGACATCCCACTGTCGAAGCAAAGCTCCGGAATCA
GCTAGAGAAGCATATCTCTGAGCGCACTATTCAAGATTCCAACTATGGTGGCAAGATCCCC
ATTGTGTGTTTTGCCCAAGGAGGTGGAAAAGAGACTTTGAAAGCCATCAATACCTCCATCA
AAAATAAAATTCCTTGTGTGGTGGTGGAAGGCTCGGGCCGGATCGCTGATGTGATCGCTAG
CCTGGTGGAGGTGGAGGATGCCCCGACATCTTCTGCCGTCAAGGAGAAGCTGGTGCGCTTT
TTACCCCGCACGGTGTCCCGGCTGTCTGAGGAGGAGACTGAGAGTTGGATCAAATGGCTCA
AAGAAATTCTCGAATGTTCTCACCTATTAACAGTTATTAAAATGGAAGAAGCTGGGGATGA
AATTGTGAGCAATGCCATCTCCTACGCTCTATACAAAGCCTTCAGCACCAGTGAGCAAGAC
AAGGATAACTGGAATGGGCAGCTGAAGCTTCTGCTGGAGTGGAACCAGCTGGACTTAGCCA
ATGATGAGATTTTCACCAATGACCGCCGATGGGAGTCTGCTGACCTTCAAGAAGTCATGTT
TACGGCTCTCATAAAGGACAGACCCAAGTTTGTCCGCCTCTTTCTGGAGAATGGCTTGAAC
CTACGGAAGTTTCTCACCCATGATGTCCTCACTGAACTCTTCTCCAACCACTTCAGCACGC
TTGTGTACCGGAATCTGCAGATCGCCAAGAATTCCTATAATGATGCCCTCCTCACGTTTGT
CTGGAAACTGGTTGCGAACTTCCGAAGAGGCTTCCGGAAGGAAGACAGAAATGGCCGGGAC
GAGATGGACATAGAACTCCACGACGTGTCTCCTATTACTCGGCACCCCCTGCAAGCTCTCT
TCATCTGGGCCATTCTTCAGAATAAGAAGGAACTCTCCAAAGTCATTTGGGAGCAGACCAG
GGGCTGCACTCTGGCAGCCCTGGGAGCCAGCAAGCTTCTGAAGACTCTGGCCAAAGTGAAG
AACGACATCAATGCTGCTGGGGAGTCCGAGGAGCTGGCTAATGAGTACGAGACCCGGGCTG
TTGAGCTGTTCACTGAGTGTTACAGCAGCGATGAAGACTTGGCAGAACAGCTGCTGGTCTA
TTCCTGTGAAGCTTGGGGTGGAAGCAACTGTCTGGAGCTGGCGGTGGAGGCCACAGACCAG
CATTTCACCGCCCAGCCTGGGGTCCAGAATTTTCTTTCTAAGCAATGGTATGGAGAGATTT
CGCCTACGACGAAGAACTGGAAGATTATCCTGTGTCTGTTTATTATACCCTTGGTGGGCTG

TGGCTTTGTATCATTTAGGAAGAAACCTGTCGACAAGCACAAGAAGCTGCTTTGGTACTAT
GTGGCGTTCTTCACCTCCCCCTTCGTGGTCTTCTCCTGGAATGTGGTCTTCTACATCGCCT
TCCTCCTGCTGTTTGCCTACGTGCTGCTCATGGATTTCCATTCGGTGCCACACCCCCCCGA
GCTGGTCCTGTACTCGCTGGTCTTTGTCCTCTTCTGTGATGAAGTGAGACAGTGGTACGTA
AATGGGGTGAATTATTTTACTGACCTGTGGAATGTGATGGACACGCTGGGGCTTTTTTACT
TCATAGCAGGAATTGTATTTCGGCTCCACTCTTCTAATAAAAGCTCTTTGTATTCTGGACG
AGTCATTTTCTGTCTGGACTACATTATTTTCACTCTAAGATTGATCCACATTTTTACTGTA
AGCAGAAACTTAGGACCCAAGATTATAATGCTGCAGAGGATGCTGATGGATGTGTTCTTCT
TCCTGTTCCTCTTTGCGGTGTGGATGGTGGCCTTTGGCGTGGCCAGGCAAGGGATCCTTAG
GCAGAATGAGCAGCGCTGGAGGTGGATATTCCGTTCGGTCATCTACGAGCCCTACCTGGCC
ATGTTCGGCCAGGTGCCCAGTGACGTGGATGGTACCACGTATGACTTTGCCCACTGCACCT
TCACTGGGAATGAGTCCAAGCCACTGTGTGTGGAGCTGGATGAGCACAACCTGCCCCGGTT
CCCCGAGTGGATCACCATCCCCCTGGTGTGCATCTACATGTTATCCACCAACATCCTGCTG
GTCAACCTGCTGGTCGCCATGTTTGGCTACACGGTGGGCACCGTCCAGGAGAACAATGACC
AGGTCTGGAAGTTCCAGAGGTACTTCCTGGTGCAGGAGTACTGCA

FIGURE 17B

GCCGCCTCAATATCCCCTTCCCCTTCATCGTCTTCGCTTACTTCTACATGGTGGTGAAGAA
GTGCTTCAAGTGTTGCTGCAAGGAGAAAAACATGGAGTCTTCTGTCTGCTGTTTCAAAAAT
GAAGACAATGAGACTCTGGCATGGGAGGGTGTCATGAAGGAAAACTACCTTGTCAAGATCA
ACACAAAAGCCAACGACACCTCAGAGGAAATGAGGCATCGATTTAGACAACTGGATACAAA
GCTTAATGATCTCAAGGGTCTTCTGAAAGAGATTGCTAATAAAATCAAATAAAACTGTATG
AAACTCTAATGGAGAAAAATCTAATTATAGCAAGATCATATTAAGGAATGCTGATGAACAA
TTTTGCTATCGACTACTAAATGAGAGATTTTCAGACCCCTGGGTACATGGTGGATGATTTT
AAATCACCCTAGTGTGCTGAGACCTTGAGAATAAAGTGTGTGATTGGTTTCATACTTGAAG
ACGGATATAAAGGAAGAATATTTCCTTTATGTGTTTCTCCAGAATGGTGCCTGTTTCTCTC
TGTGTCTCAATGCCTGGGACTGGAGGTTGATAGTTTAAGTGTGTTCTTACCGCCTCCTTTT
TCCTTTAATCTTATTTTTGATGAACACATATATAGGAGAACATCTATCCTATGAATAAGAA
CCTGGTCATGCTTTACTCCTGTATTGTTATTTTGTTCATTTCCAATTGATTCTCTACTTTT
CCCTTTTTTGTATTATGTGACTAATTAGTTGGCATATTGTTAAAAGTCTCTCAAATTAGGC
CAGATTCTAAAACATGCTGCAGCAAGAGGACCCCGCTCTCTTCAGGAAAAGTGTTTTCATT
TCTCAGGATGCTTCTTACCTGTCAGAGGAGGTGACAAGGCAGTCTCTTGCTCTCTTGGACT
CACCAGGCTCCTATTGAAGGAACCACCCCCATTCCTAAATATGTGAAAAGTCGCCCAAAAT
GCAACCTTGAAAGGCACTACTGACTTTGTTCTTATTGGATACTCCTCTTATTTATTATTTT
TCCATTAAAAATAATAGCTGGCTATTATAGAAAATTTAGACCATACAGAGATGTAGAAAGA
ACATAAATTGTCCCCATTACCTTAAGGTAATCACTGCTAACAATTTCTGGATGGTTTTTCA
AGTCTATTTTTTTTCTATGTATGTCTCAATTCTCTTTCAAAATTTTACAGAATGTTATCAT
ACTACATATATACTTTTTATGTAAGCTTTTTCACTTAGTATTTTATCAAATATGTTTTTAT
TATATTCATAGCCTTCTTAAACATTATATCAATAATTGCATAATAGGCAACCTCTAGCGAT
TACCATAATTTTGCTCATTGAAGGCTATCTCCAGTTGATCATTGGGATGAGCATCTTTGTG
CATGAATCCTATTGCTGTATTTGGGAAAATTTTCCAAGGTTAGATTCCAATAAATATCTAT
TTATTATTAAATATTAAAATATCGATTTATTATTAAAACCATTTATAAGGCTTTTTCATAA
ATGTATAGCAAATAGGAATTATTAACTTGAGCATAAGATATGAGATACATGAACCTGAACT
ATTAAAATAAAATATTATATTTAACCCTAGTTTAAGAAGAAGTCAATATGCTTATTTAAAT
ATTATGGATGGTGGGCAGATCACTTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGG
CAAAACCACATCTCTACTAAAAATAAAAAAATTAGCTGGGTGTGGTGGTGCACTCCTGTAA
TCCCAGCTACTCAGAAGGCTGAGGTACAAGAATTGCTGGAACCTGGGAGGCGGAGGTTGCA
GTGAACCAAGATTGCACCACTGCACTCCAGCCGGGGTGACAGAGTGAGACTCCGACTGAAA
ATAAATAAATAAATAAATAAATAAATAAATAAATAAATATTATGGATGGTGAAGGGAATGG
TATAGAATTGGAGAGATTATCTTACTGAACACCTGTAGTCCCAGCTTTCTCTGGAAGTGGT
GGTATTTGAGCAGGATGTGCACAAGGCAATTGAAATGCCCATAATTAGTTTCTCAGCTTTG
[illegible]TTAGAAGGAAGCTACTAAAAGATCT

AATTTGAAAAACTACAAAAGCATTAACTAAAAAAGTTTATTTTCCTTTTGTCTGGGCAGTA
GTGAAAATAACTACTCACAACATTCACTATGTTTGCAAGGAATTAACACAAATAAAAGATG
CCTTTTTACTTAAACGCCAAGACAGAAAACTTGCCCAATACTGAGAAGCAACTTGCATTAG
AGAGGGAACTGTTAAATGTTTTCAACCCAGTTCATCTGGTGGATGTTTTTGCAGGTTACTC
TGAGAATTTTGCTTATGAAAAATCATTATTTTTAGTGTAGTTCACAATAATGTATTGAACA
TACTTCTAATCAAAGGTGCTATGTCCTTGTGTATGGTACTAAATGTGTCCTGTGTACTTTT
GCACAACTGAGAATCCTGCGGCTTGGTTTAATGAGTGTGTTCATGAAATAAATAATGGAGG
AATTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAA

FIGURE 18

GAATTCGCCCTTGGCTCAGCATGAGGAACAGAAGGAATGACACTCTGGACAGCACCCGGAC
CCTGTACTCCAGCGCGTCTCGGAGCACAGACTTGTCTTACAGTGAAAGCGACTTGGTGAAT
TTTATTCAAGCAAATTTTAAGAAACGAGAATGTGTCTTCTTTACCAAAGATTCCAAGGCCA
CGGAGAATGTGTGCAAGTGTGGCTATGCCCAGAGCCAGCACATGGAAGGCACCCAGATCAA
CCAAAGTGAGAAATGGAACTACAAGAAACACACCAAGGAATTTCCTACCGACGCCTTTGGG
GATATTCAGTTTGAGACACTGGGGAAGAAAGGGAAGTATATACGTCTGTCCTGCGACACGG
ACGCGGAAATCCTTTACGAGCTGCTGACCCAGCACTGGCACCTGAAAACACCCAACCTGGT
CATTTCTGTGACCGGGGGCGCCAAGAACTTCGCCCTGAAGCCGCGCATGCGCAAGATCTTC
AGCCGGCTCATCTACATCGCGCAGTCCAAAGGTGCTTGGATTCTCACGGGAGGCACCCATT
ATGGCCTGATGAAGTACATCGGGGAGGTGGTGAGAGATAACACCATCAGCAGGAGTTCAGA
GGAGAATATTGTGGCCATTGGCATAGCAGCTTGGGGCATGGTCTCCAACCGGGACACCCTC
ATCAGGAATTGCGATGCTGAGGGCTATTTTTTAGCCCAGTACCTTATGGATGACTTCACAA
GAGATCCACTGTATATCCTGGACAACAACCACACACATTTGCTGCTCGTGGACAATGGCTG
TCATGGACATCCCACTGTCGAAGCAAAGCTCCGGAATCAGCTAGAGAAGTATATCTCTGAG
CGCACTATTCAAGATTCCAACTATGGTGGCAAGATCCCCATTGTGTGTTTTGCCCAAGGAG
GTGGAAAAGAGACTTTGAAAGCCATCAATACCTCCATCAAAAATAAAATTCCTTGTGTGGT
GGTGGAAGGCTCGGGCCAGATCGCTGATGTGATCGCTAGCCTGGTGGAGGTGGAGGATGCC
CTGACATCTTCTGCCGTCAAGGAGAAGCTGGTGCGCTTTTTACCCCGCACGGTGTCCCGGC
TGCCTGAGGAGGAGACTGAGAGTTGGATCAAATGGCTCAAAGAAATTCTCGAATGTTCTCA
CCTATTAACAGTTATTAAAATGGAAGAAGCTGGGGATGAAATTGTGAGCAATGCCATCTCC
TACGCTCTATACAAAGCCTTCAGCACCAGTGAGCAAGACAAGGATAACTGGAATGGGCAGC
TGAAGCTTCTGCTGGAGTGGAACCAGCTGGACTTAGCCAATGATGAGATTTTCACCAATGA
CCGCCGATGGGAGTCTGCTGACCTTCAAGAAGTCATGTTTACGGCTCTCATAAAGGACAGA
CCCAAGTTTGTCCGCCTCTTTCTGGAGAATGGCTTGAACCTACGGAAGTTTCTCACCCATG
ATGTCCTCACTGAACTCTTCTCCAACCACTTCAGCACGCTTGTGTACCGGAATCTGCAGAT
CGCCAAGAATTCCTATAATGATGCCCTCCTCACGTTTGTCTGGAAACTGGTTGCGAACTTC
CGAAGAGGCTTCCGGAAGGAAGACAGAAATGGCCGGGACGAGATGGACATAGAACTCCACG
ACGTGTCTCCTATTACTCGGCACCCCCTGCAAGCTCTCTTCATCTGGGCCATTCTTCAGAA
TAAGAAGGAACTCTCCAAAGTCATTTGGGAGCAGACCAGGGGCTGCACTCTGGCAGCCCTG
GGAGCCAGCAAGCTTCTGAAGACTCTGGCCAAAGTGAAGAACGACATCAATGCTGCTGGGG
AGTCCGAGGAGCTGGCTAATGAGTACGAGACCCGGGCTGTTGAGCTGTTCACTGAGTGTTA
CAGCAGCGATGAAGACTTGGCAGAACAGCTGCTGGTCTATTCCTGTGAAGCTTGGGGTGGA
AGCAACTGTCTGGAGCTGGCGGTGGAGGCCACAGACCAGCATTTCATCGCCCAGCCTGGGG
TCCAGAATTTTCTTTCTAAGCAATGGTATGGAGAGATTTCCCGAGACACCAAGAACTGGAA
[illegible]GGGCTGTGGCTTTGTATCATTTAGGAAG

AAACCTGTCGACAAGCACAAGAAGCTGCTTTGGTACTATGTGGCGTTCTTCACCTCCCCCT
TCGTGGTCTTCTCCTGGAATGTGGTCTTCTACATCGCCTTCCTCCTGCTGTTTGCCTACGT
GCTGCTCATGGATTTCCATTCGGTGCCACACCCCCCCGAGCTGGTCCTGTACTCGCTGGTC
TTTGTCCTCTTCTGTGATGAAGTGAGACAGTGGTACGTAAATGGGGTGAATTATTTTACTG
ACCTGTGGAATGTGATGGACACGCTGGGGCTTTTTTACTTCATAGCAGGAATTGTATTTCG
GCTCCACTCTTCTAATAAAAGCTCTTTGTATTCTGGACGAGTCATTTTCTGTCTGGACTAC
ATTATTTTCACTCTAAGATTGATCCACATTTTTACTGTAAGCAGAAACTTAGGACCCAAGA
TTATAATGCTGCAGAGGATGCTGATCGATGTGTTCTTCTTCCTGTTCCTCTTTGCGGTGTG
GATGGTGGCCTTTGGCGTGGCCAGGCAAGGGATCCTTAGGCAGAATGAGCAGCGCTGGAGG
TGGATATTCCGTTCGGTCATCTACGAGCCCTACCTGGCCATGTTCGGCCAGGTGCCCAGTG
ACGTGGATGGTACCACGTATGACTTTGCCCACTGCACCTTCACTGGGAATGAGTCCAAGCC
ACTGTGTGTGGAGCTGGATGAGCACAACCTGCCCCGGTTCCCCGAGTGGATCACCATCCCC
CTGGTGTGCATCTACATGTTATCCACCAACATCCTGCTGGTCAACCTGCTGGTCGCCATGT
TTGGCTACACGGTGGGCACCGTCCAGGAGAACAATGACCAGGTCTGGAAGTTCCAGAGGTA
CTTCCTGGTGCAGGAGTACTGCAGCCGCCTCAATATCCCCTTCCCCTTCATTGTCTTCGCT
TACTTCTACATGGTGGTGAAGAAGTGCTTCAAGTGTTGCTGCAAGGAGAAAAACATGGAGT
CTTCTGTCTGCTGTTTCAAAAATGAAGACAATGAGACTCTGGCATGGGAGGGTGTCATGAA
GGAAAACTACCTTGTCAAGATCAACACAAAAGCCAACGACACCTCAGAGGAAATGAGGCAT
CGATTTAGACAACTGGATACAAAGCTTAATGATCTCAAGGGTCTTCTGAAAGAGATTGCTA
ATAAAATCAAATAAAACTGTATGAACTCTAATGGAGAAAAATCTAATTATAGCAAGATCAT
ATTAAGGAATGCTGATGAACAATTTTGCTATCGACTACTAAATGAGAGATTTTCAGACCCC
TGGGTACATGGTGGATGATTTTAAATCACCCTAGTGTGCTGAGACCTTGAGAATAAAGTGT
GTGATTGGTTTCATACTTGAAGACGGATATAAAGGAAGAATATTTCCTTTATGTGTTTCTC
CAGAATGGTGCCTGTTTCTCTCTGTGTCTCAATGCCTGGGACTGGAGGTTGATAGTTTAAG
TGTGTTCTTACCGCCTCCTTTTTCCTTTAATCTTATTTTTGATGAACACATATATAGGAGA
ACATCTATCCTATGAATAAGAACCTGGTCATGCTTTAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAA

FIGURE 19

GCGGCGGGGGCGAGGTGAGGTGTTGGCAGTGGAAAGGGGTTCGGGCTCGGGGGGCGGGGGG

ACGCGGAGCGGATGGCCCGCGCCGGCCGCAGGGGCGGATAAAAAGCCGTCGCGCTGCGGGA

GTGGGCGGGAGGGAGAGGGGGTGTCCGAGGGCCACAAGAGTATGACGGGGCTGTACGAGCT

GGTGTGGCGGGTGCTGCACGCGCTGCTCTGTCTGCACCGCACGCTCACCTCCTGGCTCCGC

GTTCGGTTCGGCACCTGGAACTGGATCTGGCGGCGCTGCTGCCGCGCCGCCTCTGCCGCGG

TCCTAGCGCCGCTCGGCTTCACGCTCCGCAAGCCCCCGGCAGTCGGCAGGAACCGCCGTCA

CCACCGGCACCCGCGCGGGGGGTCGTGCCTGGCAGCCGCACACCACCGGATGCGCTGGCGC

GCGGACGGTCGTTCCTTGGAGAAGCTGCCTGTGCATATGGGCCTGGTGATCACCGAGGTGG

AGCAGGAACCCAGCTTCTCGGACATCGCGAGCCTCGTGGTGTGGTGTATGGCCGTGGGCAT

CTCCTACATTAGCGTCTACGACCACCAAGGTATTTTCAAAAGAAATAATTCCAGATTGATG

GATGAAATTTTAAAACAACAGCAAGAACTTCTGGGCCTAGATTGTTCAAAATACTCACCAG

AATTTGCAAATAGTAATGACAAAGATGATCAAGTTTTAAATTGCCATTTGGCAGTGAAGGT

GCTGTCTCCGGAAGATGGAAAAGCAGATATTGTAAGAGCTGCTCAGGACTTTTGCCAGTTA

GTAGCCCAGAAGCAAAGGAGACCCACAGATTTGGATGTAGATACGTTAGCCAGTTTACTTA

GTTCAAATGGTTGTCCTGATCCTGATTTAGTATTGAAGTTCGGTCCTGTGGACAGCACATT

AGGCTTTCTTCCCTGGCACATCAGATTGACTGAGATTGTCTCTTTGCCTTCCCACCTAAAC

ATCAGTTATGAGGACTTTTTCTCTGCCCTTCGTCAATATGCAGCCTGTGAACAGCGTCTGG

GAAAGTAGTGGTCATTGGTTGCATAATTTGATTTGAGGCTTGTGGAGGAAAGGAACCAAGT

GACTCTGATGTTTACAAAGCACCTATGAAACCCTGTACACACCTATGAAACCCTGTACACA

CCTAGTTCATAATCCTCATAATTTATCAACAAACACAAAGAAGTGTCTTACTTGAGAGTGA

GTGTGTGCGTGTGTGCGTGCACACATGTGCACGTTTGTATGTGTGGAAATAAACATAAATG

GGGACGTGTTGGAGAAGGAAATACATAGACCTACAACTTTGAGCATATAGCAGTGATGTTT

TAGGAACTGAAATGTCACACTTAATAAAGTCTTCAGCCCAGCTACTTCCCTGTTTTCGTGG

GGAGAAGGGGGCCTGATTAGAACTGTTCTGGTTGTGTTTGGCGGGAGGGGAATAAATTTTG

TTCAGTCCTTCTTAGTGACCAAACTTTAATTTTTAGGAATAACATATTGACTTACTGAACT

GAAGCATTCTGAGTTGAAAGGAGCCCCAGAGGAAAGGAGTTCTGTGTTGCTCACATGTTAA

AAGCTTGCTCACCTTCAGAGCAGAGGGAATACCTATCTTCAGATATCCGCCCATTTTCATC

TCTTCATTATAGTCAAACAGTGTGACTTGAGAGTGTTGCTCTGGTGTCTGTATTCTGGCTT

ATGAAGATTATTTGAAAAAGAACTCTTACTACATTGAAATGCAGACTTTTAAAAATTTAAA

TATTGGATTAGGCAGTCAAAAAACCAAACAAGCATAAAAGGTCAGTAAGTTGTAATCTTAA

AAGTAAAGGTGGAAAACTCATTATAAATGGAAGAAAAGTTTTGATTTCCTTTTTTGTTTGA

TGGGCAGTATGCCATATTATATCCAAAGTTGGTTTAAAAAATACTTCCATCAACTATTTTT

ATTTAAAATAAACATTTGAGGGAAGTTACCAAGGCAGCTTTTTTCCTCAAAAGTAACCTGT

TCCTCTTTGGAACAGCACATTTTAGGGGCATGGTTAATACCTGAGATTTTTACTCAGTAAA

[illegible]AAGTAGGATTGAAGGCCTCTGTGGGGAATAT

AATGTTACCAAAGCCTATAAAAATAAATTTTACATGTTCTTTTTTATGACAGAGAGCAGCA
CTGGTTCTGTTATTTTTTAAAATGAATAACTGATTTCTTGACAGGTGTTTAGTATTTCTTC
CCTCACTGCTGATTCTTGGATAGAAACCATTCTTTATATTTGATAGACTGTTTTCAGAAAA
CTCTTATCAACAAGTGTACAATAGTTATCTAAAACTGTACATTTAGAATGGAGCAGTTTAA
TACTAGATCTCAGAAATTTGAAAAATAGCAAAGAAGACTGGATTTGGAAAGCATGGTCTAC
AATTGCTTGTCAATTCTGAAGCTATGAAGAATAAATGTTTCAGCTTTGGATTATGAAACCC
CGTTTATGATTTTTAAAACACACTTGAAATAAAAATGATTGAACTAAATTTTGGTCCGGTG
ACATCATTTTGCACTGCATAGCCCATTATACATTGTAAGTCTTTTTGTTTTAATTTATTAC
TGAGAGTTTTCTGTGAAGCTACAGCATATCTAATCAGAGAATTTCTGATTTGTTTCCTTAT
ACTGTGATTATGTTATATTAAGGCATTTGTAGTGCAGCTTAAGACTGAATTTATGCTTTTT
GTAAACATGATCGGTGTAAATGTCTTATAAACATTCTGGAGTATGTATAACTTTAATGAAT
GAAATTTAATGGACCTGATTAAAATAAAGGGATTTAATCTTTGTTAAAGTGAAGTTAGTGG
AATAAATTACCTACTGGAATATAGTCCAAGTCACTAAAGGTTTAATATTTGCATTTTTGTG
CTTTTATTTTCTCCTTCCATTCATAATTATATACTTGAAAGTACATCTGTAGCCTATGACT
TGAGTCTCTTGAACTTCTAGGAAGAGGCAAACTACAAACTACTAGGATTCTGATTTCAGAT
ATAGGCATTCCAGAATCTTCTCTTTACGAGTTCACCTGCTAGTATAATCTCCACAACTTGA
ATGGCCTTGGTTGTTCTGTAATTGCTGCCAAAATCATCACAAGCTGTACGTCATCAAGGCT
CCCTTTGCACTCCCAAGAAGAACTGTTCATTTTAAACAAAAGTGTATGTCTTTATTTGTAT
TGGAAAATATTGTCTTTAAATTGTTTCTTGTTGACACTTCCCACAATGGAAAAATTACCAA
ATTAAACCTGTTTTATGGATGGCAGCTTGGAGCATAGCAAGAATTTGGAGGATTTGAATTC
CATTCCCAGTTCTCATTGCGTTTTGTTTCTTAAAACTATAATAATTGGTTACTGTTATAAA
GTTTAAAAGGTGGTTTTAATATGAATAGCAAATTCTGGTATATTGTGACTAACGCTTAAGA
ATGCCTGTCTTTGAGAGGAAGGTGTTATAATATTAATCAACAGTGCCAAATACACTGTGCA
TATCTACAATTTAATCTTTGAATGTTTGTTACTGGATTAGCTCCCTCCTCCTTCTGTGTGA
TGGTACCATGCATAGAGTCAATCAAATCCTTGTGATGTTTTGTATGGACTTTGACAACATG
TAACTAATGTGTAAAGCAAGTTTTTATGATTAAGGAATCAAATTTATTGAATTTTATTATT
GAAAGTTGAAACTTAACATGTATAAACAAAAAACAATAAAATAATAAACTATTTTCATTGA
CTATAAAAAAAAAGCCCACGCGTCCGCC

FIGURE 20

GAGGTGAGGTGTTGGCAGTGGAAAGGGGTTCGGGCTCGGGGGGCGGGGGGACGCGGAGCGA
TGGCCCGCGCCGGCCGCAGGGGCGGATAAAAAGCCGTCGCGCTGCGGGAGTGGGCGGGAGG
GAGAGGGGGTGTCCGAGGGCCACAAGAGTATGACGGGGCTGTACGAGCTGGTGTGGCGGGT
GCTGCACGCGCTGCTCTGTCTGCACCGCACGCTCACCTCCTGGCTCCGCGTTCGGTTCGGC
ACCTGGAACTGGATCTGGCGGCGCTGCTGCCGCGCCGCCTCTGCCGCGGTCCTAGCGCCGC
TCGGCTTCACGCTCCGCAAGCCCCCGGCAGTCGGCAGGAACCGCCGTCACCACCGGCACCC
GCGCGGGGGGTCGTGCCTGGCAGCCGCACACCACCGGATGCGCTGGCGCGCGGACGGTCGT
TCCTTGGAGAAGCTGCCTGTGCATATGGGCCTGGTGATCACCGAGGTGGAGCAGGAACCCA
GCTTCTCGGACATCGCGAGCCTCGTGGTGTGGTGTATGGCCGTGGGCATCTCCTACATTAG
CGTCTACGACCACCAAGGTATTTTCAAAAGAAATAATTCCAGATTGATGGATGAAATTTTA
AAACAACAGCAAGAACTTCTGGGCCTAGATTGTTCAAAATACTCACCAGAATTTGCAAATA
GTAATGACAAAGATGATCAAGTTTTAAATTGCCATTTGGCAGTGAAGGTGCTGTCTCCGGA
AGATGGAAAAGCAGATATTGTAAGAGCTGCTCAGGACTTTTGCCAGTTAGTAGCCCAGAAG
CAAAAGAGACCCACAGATTTGGATGTAGATACGTTAGCCAGTTTACTTAGTTCAAATGGTT
GTCCTGATCCTGATTTAGTATTGAAGTTCGGTCCTGTGGACAGCACATTAGGCTTTCTTCC
CTGGCACATCAGATTGACTGAGATTGTCTCTTTGCCTTCCCACCTAAACATCAGTTATGAG
GACTTTTTCTCTGCCCTTCGTCAATATGCAGCCTGTGAACAGCGTCTGGGAAAGTAGTGGT
CATTGGTTGCA

FIGURE 21

GGGGGGCGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGA
TCTGAGTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGC
TGAGAAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCA
GTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTC
TGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGT
GGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT
TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTG
GGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCT
GTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC
CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGC
TGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGG
GCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTC
TGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATC
AGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGG
CCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCAC
CTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCA
GAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTT
TCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCAC
CCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTG
TTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCA
CCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCA
GTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACT
CGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT
CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCT
GCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCC
AAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC
CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCT
GCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTG
GTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCA
TCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGT
CCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCC
ATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGA
AAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGT
TAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTC
[illegible]CAGCTGGGGGCTGGGGCGTCCCT

CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCT
GGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGTGGGGACTCTGCAGGTGG
ATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTT
TTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCT
TCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGA
ACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCC
CTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTG
GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTA
TTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGT
AGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTG
CCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAAT
TCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAG
GTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAG
CCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACT
GCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACA
ACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCC
ATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCC
CCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGTTTAAGTGCCGTTTGCAATAATG
TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAAT
GTTTATGGTGACAAAATTAAAGCCTTTCTTATATGTTT

FIGURE 22

AACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGCCCCTGGTTCGGCCCACCTCT
GAAGGTTCCAGAATCGATAGTGAATTCGTGGGGAAGCGCAGCCATGGCTCTGCGGAGGCTG
GGGGCCGCGCTGCTGCTGCTGCCGCTGCTCGCCGCCGTGGAAGAAACGCTAATGGACTCCA
CTACAGCGACTGCTGAGCTGGGCTGGATGGTGCATCCTCCATCAGGGTGGGAAGAGGTGAG
TGGCTACGATGAGAACATGAACACGATCCGCACGTACCAGGTGTGCAACGTGTTTGAGTCA
AGCCAGAACAACTGGCTACGGACCAAGTTTATCCGGCGCCGTGGCGCCCACCGCATCCACG
TGGAGATGAAGTTTTCGGTGCGTGACTGCAGCAGCATCCCCAGCGTGCGTGGCTCCTGCAA
GGAGACCTTCAACCTCTATTACTATGAGGCTGACTTTGACTCGGCCACCAAGACCTTCCCC
AACTGGATGGAGAATCCATGGGTGAAGGTGGATACCATTGCAGCCGACGAGAGCTTCTCCC
AGGTGGACCTGGGTGGCCGCGTCATGAAAATCAACACCGAGGTGCGGAGCTTCGGACCTGT
GTCCCGCAGCGGCTTCTACCTGGCCTTCCAGGACTATGGCGGCTGCATGTCCCTCATCGCC
GTGCGTGTCTTCTACCGCAAGTGCCCCCGCATCATCCAGAATGGCGCCATCTTCCAGGAAA
CCCTGTCGGGGGCTGAGAGCACATCGCTGGTGGCTGCCCGGGGCAGCTGCATCGCCAATGC
GGAAGAGGTGGATGTACCCATCAAGCTCTACTGTAACGGGGACGGCGAGTGGCTGGTGCCC
ATCGGGCGCTGCATGTGCAAAGCAGGCTTCGAGGCCGTTGAGAATGGCACCGTCTGCCGAG
GTTGTCCATCTGGGACTTTCAAGGCCAACCAAGGGGATGAGGCCTGTACCCACTGTCCCAT
CAACAGCCGGACCACTTCTGAAGGGGCCACCAACTGTGTCTGCCGCAATGGCTACTACAGA
GCAGACCTGGACCCCCTGGACATGCCCTGCACAACCATCCCCTCCGCGCCCCAGGCTGTGA
TTTCCAGTGTCAATGAGACCTCCCTCATGCTGGAGTGGACCCCTCCCCGCGACTCCGGAGG
CCGAGAGGACCTCGTCTACAACATCATCTGCAAGAGCTGTGGCTCGGGCCGGGGGTGCCTGC
ACCCGCTGCGGGGACAATGTACAGTACGCACCACGCCAGCTAGGCCTGACCGAGCCACGCA
TTTACATCAGTGACCTGCTGGCCCACACCCAGTACACCTTCGAGATCCAGGCTGTGAACGG
CGTTACTGACCAGAGCCCCTTCTCGCCTCAGTTCGCCTCTGTGAACATCACCACCAACCAG
GCAGCTCCATCGGCAGTGTCCATCATGCATCAGGTGAGCCGCACCGTGGACAGCATTACCC
TGTCGTGGTCCCAGCCAGACCAGCCCAATGGCGTGATCCTGGACTATGAGCTGCAGTACTA
TGAGAAGGAGCTCAGTGAGTACAACGCCACAGCCATAAAAAGCCCCACCAACACGGTCACC
GTGCAGGGCCTCAAAGCCGGCGCCATCTATGTCTTCCAGGTGCGGGCACGCACCGTGGCAG
GCTACGGGCGCTACAGCGGCAAGATGTACTTCCAGACCATGACAGAAGCCGAGTACCAGAC
AAGCATCCAGGAGAAGTTGCCACTCATCATCGGCTCCTCGGCCGCTGGCCTGGTCTTCCTC
ATTGCTGTGGTTGTCATCGCCATCGTGTGTAACAGAAGACGGGGGTTTGAGCGTGCTGACT
CGGAGTACACGGACAAGCTGCAACACTACACCAGTGGCCACATGACCCCAGGCATGAAGAT
CTACATCGATCCTTTCACCTACGAGGACCCCAACGAGGCAGTGCGGGAGTTTGCCAAGGAA
ATTGACATCTCCTGTGTCAAAATTGAGCAGGTGATCGGAGCAGGGGAGTTTGGCGAGGTCT
GCAGTGGCCACCTGAAGCTGCCAGGCAAGAGAGAGATCTTTGTGGCCATCAAGACGCTCAA
[illegible]TGAGCGAAGCCTCCATCATGGGCCAG

TTCGACCATCCCAACGTCATCCACCTGGAGGGTGTCGTGACCAAGAGCACACCTGTGATGA
TCATCACCGAGTTCATGGAGAATGGCTCCCTGGACTCCTTTCTCCGGCAAAACGATGGGCA
GTTCACAGTCATCCAGCTGGTGGGCATGCTTCGGGGCATCGCAGCTGGCATGAAGTACCTG
GCAGACATGAACTATGTTCACCGTGACCTGGCTGCCCGCAACATCCTCGTCAACAGCAACC
TGGTCTGCAAGGTGTCGGACTTTGGGCTCTCACGCTTTCTAGAGGACGATACCTCAGACCC
CACCTACACCAGTGCCCTGGGCGGAAAGATCCCCATCCGCTGGACAGCCCCGGAAGCCATC
CAGTACCGGAAGTTCACCTCGGCCAGTGATGTGTGGAGCTACGGCATTGTCATGTGGGAGG
TGATGTCCTATGGGGAGCGGCCCTACTGGGACATGACCAACCAGGATGTAATCAATGCCAT
TGAGCAGGACTATCGGCTGCCACCGCCCATGGACTGCCCGAGCGCCCTGCACCAACTCATG
CTGGACTGTTGGCAGAAGGACCGCAACCACCGGCCCAAGTTCGGCCAAATTGTCAACACGC
TAGACAAGATGATCCGCAATCCCAACAGCCTCAAAGCCATGGCGCCCCTCTCCTCTGGCAT
CAACCTGCCGCTGCTGGACCGCACGATCCCCGACTACACCAGCTTTAACACGGTGGACGAG
TGGCTGGAGGCCATCAAGATGGGGCAGTACAAGGAGAGCTTCGCCAATGCCGGCTTCACCT
CCTTTGACGTCGTGTCTCAGATGATGATGGAGGACATTCTCCGGGTTGGGCTCACTTTGGC
TGGCCACCAGAAAAAAATCCTGAACAGTATCCAGGTGATGCGGGCGCAGATGAACCAGATT
CAGTCTGTGGAGGTTTGACATTCACCTGCCTCGGCTCACCTCTTCCTCCAAGCCCCGCCCC
CTCTGCCCCACGTGCCGGCCCTCCTGGTGCTCTATCCACTGCAGGGCCAGCCACTCGCCAG
GAGGCCACGGGCCACGGGAAGAACCAAGCGGTGCCAGCCACGAGACGTCACCAAGAAAACA
TGCAACTCAAACGACGGAAAAAAAAAGGGAATGGGAAAAAAGAAAACAGATCCTGGGAGGG
GGCGGGAAATACAAGGAATATTTTTTAAAGAGGATTCTCATAAGGAAAGCAATGACTGTTC
TTGCGGGGGATAAAAAAGGGCTTGGGAGATTCATGCGATGTGTCCAATCGGAGACAAAAGC
AGTTTCTCTCCAACTCCCTCTGGGAAGGTGACCTGGCCAGAGCCAAGAAACACTTTCAGAA
AAACAAATGTGAAGGGGAGAGACAGGGGCCGCCCTTGGCTCCTGTCCCTGCTGCTCCTCTA
GGCCTCACTCAACAACCAAGCGCCTGGAGGACGGGACAGATGGACAGACAGCCACCCTGAG
AACCCCTCTGGGAAAATCTATTCCTGCCACCACTGGGCAAACAGAAGAATTTTTCTGTCTT
TGGAGAGTATTTTAGAAACTCCAATGAAAGACACTGTTTCTCCTGTTGGCTCACAGGGCTG
AAAGGGGCTTTTGTCCTCCTGGGTCAGGGAGAACGCGGGGACCCCAG

FIGURE 23

CGGGAAGCGCAGCCATGGCTCTGCGGAGGCTGGGGGCCGCGCTGCTGCTGCTGCCGCTGCT
CGCCGCCGTGGAAGAAACGCTAATGGACTCCACTACAGCGACTGCTGAGCTGGGCTGGATG
GTGCATCCTCCATCAGGGTGGGAAGAGGTGAGTGGCTACGATGAGAACATGAACACGATCC
GCACGTACCAGGTGTGCAACGTGTTTGAGTCAAGCCAGAACAACTGGCTACGGACCAAGTT
TATCCGGCGCCGTGGCGCCCACCGCATCCACGTGGAGATGAAGTTTTCGGTGCGTGACTGC
AGCAGCATCCCCAGCGTGCCTGGCTCCTGCAAGGAGACCTTCAACCTCTATTACTATGAGG
CTGACTTTGACTCGGCCACCAAGACCTTCCCCAACTGGATGGAGAATCCATGGGTGAAGGT
GGATACCATTGCAGCCGACGAGAGCTTCTCCCAGGTGGACCTGGGTGGCCGCGTCATGAAA
ATCAACACCGAGGTGCGGAGCTTCGGACCTGTGTCCCGCAGCGGCTTCTACCTGGCCTTCC
AGGACTATGGCGGCTGCATGTCCCTCATCGCCGTGCGTGTCTTCTACCGCAAGTGCCCCCG
CATCATCCAGAATGGCGCCATCTTCCAGGAAACCCTGTCGGGGGCTGAGAGCACATCGCTG
GTGGCTGCCCGGGGCAGCTGCATCGCCAATGCGGAAGAGGTGGATGTACCCATCAAGCTCT
ACTGTAACGGGGACGGCGAGTGGCTGGTGCCCATCGGGCGCTGCATGTGCAAAGCAGGCTT
CGAGGCCGTTGAGAATGGCACCGTCTGCCGAGGTTGTCCATCTGGGACTTTCAAGGCCAAC
CAAGGGGATGAGGCCTGTACCCACTGTCCCATCAACAGCCGGACCACTTCTGAAGGGGCCA
CCAACTGTGTCTGCCGCAATGGCTACTACAGAGCAGACCTGGACCCCCTGGACATGCCCTG
CACAACCATCCCCTCCGCGCCCCAGGCTGTGATTTCCAGTGTCAATGAGACCTCCCTCATG
CTGGAGTGGACCCCTCCCCGCGACTCCGGAGGCCGAGAGGACCTCGTCTACAACATCATCT
GCAAGAGCTGTGGCTCGGGCCGGGGTGCCTGCACCCGCTGCGGGGACAATGTACAGTACGC
ACCACGCCAGCTAGGCCTGACCGAGCCACGCATTTACATCAGTGACCTGCTGGCCCACACC
CAGTACACCTTCGAGATCCAGGCTGTGAACGGCGTTACTGACCAGAGCCCCTTCTCGCCTC
AGTTCGCCTCTGTGAACATCACCACCAACCAGGCAGCTCCATCGGCAGTGTCCATCATGCA
TCAGGTGAGCCGCACCGTGGACAGCATTACCCTGTCGTGGTCCCAGCCGGACCAGCCCAAT
GGCGTGATCCTGGACTATGAGCTGCAGTACTATGAGAAGGAGCTCAGTGAGTACAACGCCA
CAGCCATAAAAAGCCCCACCAACACGGTCACCGTGCAGGGCCTCAAAGCCGGCGCCATCTA
TGTCTTCCAGGTGCGGGCACGCACCGTGGCAGGCTACGGGCGCTACAGCGGCAAGATGTAC
TTCCAGACCATGACAGAAGCCGAGTACCAGACAAGCATCCAGGAGAAGTTGCCACTCATCA
TCGGCTCCTCGGCCGCTGGCCTGGTCTTCCTCATTGCTGTGGTTGTCATCGCCATCGTGTG
TAACAGACGGGGGTTTGAGCGTGCTGACTCGGAGTACACGGACAAGCTGCAACACTACACC
AGTGGCCACATGACCCCAGGCATGAAGATCTACATCGATCCTTTCACCTACGAGGACCCCA
ACGAGGCAGTGCGGGAGTTTGCCAAGGAAATTGACATCTCCTGTGTCAAAATTGAGCAGGT
GATCGGAGCAGGGGAGTTTGGCGAGGTCTGCAGTGGCCACCTGAAGCTGCCAGGCAAGAGA
GAGATCTTTGTGGCCATCAAGACGCTCAAGTCGGGCTACACGGAGAAGCAGCGCCGGGACT
TCCTGAGCGAAGCCTCCATCATGGGCCAGTTCGACCATCCCAACGTCATCCACCTGGAGGG
[illegible]GAGTTCATGGAGAATGGCTCCCTG

GACTCCTTTCTCCGGCAAAACGATGGGCAGTTCACAGTCATCCAGCTGGTGGGCATGCTTC
GGGGCATCGCAGCTGGCATGAAGTACCTGGCAGACATGAACTATGTTCACCGTGACCTGGC
TGCCCGCAACATCCTCGTCAACAGCAACCTGGTCTGCAAGGTGTCGGACTTTGGGCTCTCA
CGCTTTCTAGAGGACGATACCTCAGACCCCACCTACACCAGTGCCCTGGGCGGAAAGATCC
CCATCCGCTGGACAGCCCCGGAAGCCATCCAGTACCGGAAGTTCACCTCGGCCAGTGATGT
GTGGAGCTACGGCATTGTCATGTGGGAGGTGATGTCCTATGGGGAGCGGCCCTACTGGGAC
ATGACCAACCAGGATGTAATCAATGCCATTGAGCAGGACTATCGGCTGCCACCGCCCATGG
ACTGCCCGAGCGCCCTGCACCAACTCATGCTGGACTGTTGGCAGAAGGACCGCAACCACCG
GCCCAAGTTCGGCCAAATTGTCAACACGCTAGACAAGATGATCCGCAATCCCAACAGCCTC
AAAGCCATGGCGCCCCTCTCCTCTGGCATCAACCTGCCGCTGCTGGACCGCACGATCCCCG
ACTACACCAGCTTTAACACGGTGGACGAGTGGCTGGAGGCCATCAAGATGGGGCAGTACAA
GGAGAGCTTCGCCAATGCCGGCTTCACCTCCTTTGACGTCGTGTCTCAGATGATGATGGAG
GACATTCTCCGGGTTGGGGTCACTTTGGCTGGCCACCAGAAAAAAATCCTGAACAGTATCC
AGGTGATGCGGGCGCAGATGAACCAGATTCAGTCTGTGGAGGTTTGACATTCACCTGCCTC
GGCTCACCTCTTCCTCCAAGCCCCGCCCCCTCTGCCCCACGTGCCGGCCCTCCTGGTGCTC
TATCCACTGCAGGGCCAGCCACTCGCCAGGAGGCCACGGGCCACGGGAAGAACCAAGCGGT
GCCAGCCACGAGACGTCACCAAGAAAACATGCAACTCAAACGACGGAAAAAAAAAAGGGAAT
GGGAAAAAAGAAAACAGATCCTGGGAGGGGGCGGGAAATACAAGGAATATTTTTTAAAGAG
GATTCTCATAAGGAAAGCAATGACTGTTCTTGCGGGGGATAAAAAAGGGCTTGGGAGATTC
ATGCGATGTGTCCAATCGGAGACAAAAGCAGTTTCTCTCCAACTCCCTCTGGGAAGGTGAC
CTGGCCAGAGCCAAGAAACACTTTCAGAAAAACAAATGTGAAGGGGAGAGACAGGGGCCGC
CCTTGGCTCCTGTCCCTGCTGCTCCTCTAGGCCTCACTCAACAACCAAGCGCCTGGAGGAC
GGGACAGATGGACAGACAGCCACCCTGAGAACCCCTCTGGGAAAATCTATTCCTGCCACCA
CTGGGCAAACAGAAGAATTTTTCTGTCTTTGGAGAGTATTTTAGAAACTCCAATGAAAGAC
ACTGTTTCTCCTGTTGGCTCACAGGGCTGAAAGGGGCTTTTGTCCTCCTGGGTCAGGGAGA
ACGCGGGGACCCCAGAAAGGTCAGCCTTCCTGAGGATGGGCAACCCCCAGGTCTGCAGCTC
CAGGTACATATCACGCGCACAGCCTGGCAGCCTGGCCCTCCTGGTGCCCACTCCCGCCAGC
CCCTGCCTCGAGGTCGAC

FIGURE 24

CCGCCAGCTCCTTCGGCAATGAACTTCTCCACCAGCAGCAGCAGCTTCGCCTACGACCGGG
AGTTCCTCCGCACCCTGCCCGGCTTCCTCATCGTGGCCGAGATCGTTCTGGGGCTGCTGGT
ATGGACGCTTATTGCTGGAACTGAGTACTTCCGGGTCCCCGCATTTGGCTGGGTCATGTTT
GTAGCTGTATTTTACTGGGTCCTCACCGTCTTCTTCCTCATTATCTACATAACAATGACCT
ACACCAGGATTCCCCAGGTGCCCTGGACAACAGTGGGCCTGTGCTTTAACGGCAGTGCCTT
CGTCTTGTACCTCTCTGCCGCTGTTGTAGATGCATCTTCCGTCTCCCCTGAGAGGGACAGT
CACAACTTCAACAGCTGGGCGGCCTCATCGTTCTTTGCCTTCCTGGTCAACATCTGCTACG
CTGGAAATACATATTTCAGTTTTATAGCATGGAGATCCAGGACCATACAGTGATTTACCAT
TTTGATAATTAAAAGGAAAAAAAAAGGAAGACTCTCACTGTAAAAACAGCTGTAGGTATAA
TGTATATTCCCAGAGAATTGTATTTAACTAATTAATGTTTTTTATATTCTTAAATTTGCTC
ACAAATTGTGGTTTGTTACAATTAAACTGGATACTTATTTGCAAAGTGTTGTAGCTTATAA
TGAACTCTTAAGTATCTTATTAATGTATTAATGTCTTCATAGATCATATTTTCTTAGACAA
TGTTTAAATAGATAAATTGCTAATATTGAGAATGTGTCAAGTTTGTAAACCTAACTTTTAA
GATGCCAGATTCTTTTTTGATTAAATGTTGCAAAATCCCAAAAAAAAAAAAAAAAAAAAA

FIGURE 25

GCCCTTGGCAATGAACTTCTCCACCAGCAGCAGCAGCTTCGCCTACGACCGGGAGTTCCTC
CGCACCCTGCCCGGCTTCCTCATCGTGGCCGAGATCGTTCTGGGGCTGCTGGTATGGACGC
TTATTGCTGGAACTGAGTACTTCCGGGTCCCCGCATTTGGCTGGGTCATGTTTGTAGCTGT
ATTTTACTGGGTCCTCACCGTCTTCTTCCTCATTATCTACATAACAATGACCTACACCAGG
ATTCCCCAGGTGCCCTGGACAACAGTGGGCCTGTGCTTTAACGGCAGTGCCTTCGTCTTGT
ACCTCTCTGCCGCTGTTGTAGATGCATCTTCCGTCTCCCCTGAGAGGGACAGTCACAACTT
CAACAGCTGGGCGGCCTCATCGTTCTTTGCCTTCCTGGTCACCATCTGCTACGCTGGAAAT
ACATATTTCAGTTTTATAGCATGGAGATCCAGGACCATACAGTGATTTACCAAAGGGC

FIGURE 26

GGCCCTGCCCTGAAGGCTGGTCACTTGCAGAGGTAAACTCCCCTCTTTGACTTCTGGCCAG
GGTTTGTGCTGAGCTGGCTGCAGCCGCTCTCAGCCTCGCTCCGGGCACGTCGGGCAGCCTC
GGGCCCTCCTGCCTGCAGGATCATGCCCACCACCGTGGACGATGTCCTGGAGCATGGAGGG
GAGTTTCACTTTTTCCAGAAGCAAATGTTTTTCCTCTTGGCTCTGCTCTCGGCTACCTTCG
CGCCCATCTACGTGGGCATCGTCTTCCTGGGCTTCACCCCTGACCACCGCTGCCGGAGCCC
CGGAGTGGCCGAGCTGAGTCTGCGCTGCGGCTGGAGTCCTGCAGAGGAACTGAACTACACG
GTGCCGGGCCCAGGACCTGCGGGCGAAGCCTCCCCAAGACAGTGTAGGCGCTACGAGGTGG
ACTGGAACCAGAGCACCTTTGACTGCGTGGACCCCCTGGCCAGCCTGGACACCAACAGGAG
CCGCCTGCCACTGGGCCCCTGCCGGGACGGCTGGGTGTACGAGACGCCTGGCTCGTCCATC
GTCACCGAGTTTAACCTGGTATGTGCCAACTCCTGGATGTTGGACCTATTCCAGTCATCAG
TGAATGTAGGATTCTTTATTGGCTCTATGAGTATCGGCTACATAGCAGACAGGTTTGGCCG
TAAGCTCTGCCTCCTAACTACAGTCCTCATAAATGCTGCAGCTGGAGTTCTCATGGCCATT
TCCCCAACCTATACGTGGATGTTAATTTTTCGCTTAATCCAAGGACTGGTCAGCAAAGCAG
GCTGGTTAATAGGCTACATCCTGATTACAGAATTTGTTGGGCGGAGATATCGGAGAACAGT
GGGGATTTTTTACCAAGTTGCCTATACAGTTGGGCTCCTGGTGCTAGCTGGGGTGGCTTAC
GCACTTCCTCACTGGAGGTGGTTGCAGTTCACAGTTGCTCTGCCCAACTTCTTCTTCTTGC
TCTATTACTGGTGCATACCTGAGTCTCCCAGGTGGCTGATCTCCCAGAATAAGAATGCTGA
AGCCATGAGAATCATTAAGCACATCGCAAAGAAAAATGGAAAATCTCTACCCGCCTCCCTT
CAGCGCCTGAGACTTGAAGAGGAAACTGGCAAGAAATTGAACCCTTCATTTCTTGACTTGG
TCAGAACTCCTCAGATAAGGAAACATACTATGATATTGATGTACAACTGGTTCACGAGCTC
TGTGCTCTACCAGGGCCTCATCATGCACATGGGCCTTGCAGGTGACAATATCTACCTGGAT
TTCTTCTACTCTGCCCTGGTTGAATTCCCAGCTGCCTTCATGATCATCCTCACCATCGACC
GCATCGGACGCCGTTACCCTTGGGCTGCATCAAATATGGTTGCAGGGGCAGCCTGTCTGGC
CTCAGTTTTTATACCTGGTGATCTACAATGGCTAAAAATTATTATCTCATGCTTGGGAAGA
ATGGGGATCACAATGGCCTATGAGATAGTCTGCCTGGTCAATGCTGAGCTGTACCCCACAT
TCATTAGGAATCTTGGCGTCCACATCTGTTCCTCAATGTGTGACATTGGTGGCATCATCAC
GCCATTCCTGGTCTACCGGCTCACTAACATCTGGCTTGAGCTCCCGCTGATGGTTTTCGGC
GTACTTGGCTTGGTTGCTGGAGGTCTGGTGCTGTTGCTTCCAGAAACTAAAGGGAAAGCTT
TGCCTGAGACCATCGAGGAAGCCGAAAATATGCAAAGACCAAGAAAAAATAAAGAAAAGAT
GATTTACCTCCAAGTTCAGAAACTAGACATTCCATTGAACTAAGAAGAGAGACCGTTGCTG
CTGTCATGACCTAGCTTTGATGGCAGCAAGACCAAAAGTAGAAATCCCTGCACTCATCACA
AAGCCCATACAACTCAACCAAACTTACCCCTGAGCCCTATCAACCTAGGTCTACAGCCAGT
GGAGTCTATTGTACACTGTGGAAAAATACCCATGGGACCAGATCCTGCCAAATTCTTCCAG
CTCACTTTATTCTCAGCATTCCTAGGACATTGGACATTGGTTTTCTGGAGGGTTTTTTTTC
[illegible]TGCAATGCTAGCAACCAGAATACAT

AGGGGAACTGTGGGCTAGGCAAANAAAATAGAAAAAGTGTGAAAAACAGTAAAGTTGGGAG

AGGAGCATCTATTTTCTTAAAGAAATAAAACACCNAAAACAAAAAAAAAAAAAAAAAAAAA

FIGURE 27

GCCCTTCTCCTGCCTGCAGGATCATGCCCACCACCGTGGACGATGTCCTGGAGCATGGAGG
GGAGTTTCACTTTTTCCAGAAGCAAATGTTTTTCCTCTTGGCTCTGCTCTCGGCTACCTTC
GCGCCCATCTACGTGGGCATCGTCTTCCTGGGCTTCACCCCTGACCACCGCTGCCGGAGCC
CCGGAGTGGCCGAGCTGAGTCTGCGCTGCGGCTGGAGTCCTGCAGAGGAACTGAACTACAC
GGTGCCGGGCCCAGGACCTGCGGGCGAAGCCTCCCCAAGACAGTGTAGGCGCTACGAGGTG
GACTGGAACCAGAGCACCTTCGACTGCGTGGACCCCCTGGCCAGCCTGGACACCAACAGGA
GCCGCCTGCCACTGGGCCCCTGCCGGGACGGCTGGGTGTACGAGACGCCTGGCTCGTCCAT
CGTCACCGAGTTTAACCTGGTATGTGCCAACTCCTGGATGTTGGACCTATTCCAGTCATCA
GTGAATGTAGGATTCTTTATTGGCTCTATGAGTATCGGCTACATAGCAGACAGGTTTGGCC
GTAAGCTCTGCCTCCTAACTACAGTCCTCATAAATGCTGCAGCTGGAGTTCTCATGGCCAT
TTCCCCAACCTATACGTGGATGTTAATTTTTCGCTTAATCCAAGGACTGGTCAGCAAAGCA
GGCTGGTTAATAGGCTACATCCTGATTACAGAATTTGTTGGGCGGAGATATCGGAGAACAG
TGGGGATTTTTTACCAAGTTGCCTATACAGTTGGGCTCCTGGTGCTAGCTGGGGTGGCTTA
CGCACTTCCTCACTGGAGGTGGTTGCAGTTCACAGTTTCTCTGCCCAACTTCTTCTTCTTG
CTCTATTACTGGTGCATACCTGAGTCTCCCAGGTGGCTGATCTCCCAGAATAAGAATGCTG
AAGCCATGAGAATCATTAAGCACATCGCAAAGAAAAATGGAAAATCTCTACCCGCCTCCCT
TCAGCGCCTGAGACTTGAAGAGGAAACTGGCAAGAAATTGAACCCTTCATTTCTTGACTTG
GTCAGAACTCCTCAGATAAGGAAACATACTATGATATTGATGTACAACTGGTTCACGAGCT
CTGTGCTCTACCAGGGCCTCATCATGCACATGGGCCTTGCAGGTGACAATATCTACCTGGA
TTTCTTCTACTCTGCCCTGGTTGAATTCCCAGCTGCCTTCATGATCATCCTCACCATCGAC
CGCATCGGACGCCGTTACCCTTGGGCTGCATCAAATATGGTTGCAGGGGCAGCCTGTCTGG
CCTCAGTTTTTATACCTGGTGATCTACAATGGCTAAAAATTATTATCTCATGCTTGGGAAG
AATGGGGATCACAATGGCCTATGAGATAGTCTGCCTGGTCAATGCTGAGCTGTACCCCACA
TTCATTAGGAATCTTGGCGTCCACATCTGTTCCTCAATGTGTGACATTGGTGGCATCATCA
CGCCATTCCTGGTCTACCGGCTCACTAACATCTGGCTTGAGCTCCCGCTGATGGTTTTCGG
CGTGCTTGGCTTGGTTGCTGGAGGTCTGGTGCTGTTGCTTCCAGAAACTAAAGGGAAAGCT
TTGCCTGAGACCATCGAGGAAGCCGAAAATATGCAAAGACCAAGAAAAAATAAAGAAAAGA
TGATTTACCTCCAAGTTCAGAAACTAGACATTCCATTGAACTAAGAAGAGAGACCGTTGCT
GCTGTCATGACAAGGGC

FIGURE 28

CGGGCCAGGTTTCCAGGCTCGGCCGCCGCCTCCATCCCAGCACCTGCGGAGGGAGCGCTGA
CCATGGCTCCCTGGCCTGAATTGGGAGATGCCCAGCCCAACCCCGATAAGTACCTCGAAGG
GGCCGCAGGTCAGCAGCCCACTGCCCCTGATAAAAGCAAAGAGACCAACAAAACAGATAAC
ACTGAGGCACCTGTAACCAAGATTGAACTTCTGCCGTCCTACTCCACGGCTACACTGATAG
ATGAGCCCACTGAGGTGGATGACCCCTGGAACCTACCCACTCTTCAGGACTCGGGGATCAA
GTGGTCAGAGAGAGACACCAAAGGGAAGATTCTCTGTTTCTTCCAAGGGATTGGGAGATTG
ATTTTACTTCTCGGATTTCTCTACTTTTTCGTGTGCTCCCTGGATATTCTTAGTAGCGCCT
TCCAGCTGGTTGGAGGAAAAATGGCAGGACAGTTCTTCAGCAACAGCTCTATTATGTCCAA
CCCTTTGTTGGGGCTGGTGATCGGGGTGCTGGTGACCGTCTTGGTGCAGAGCTCCAGCACC
TCAACGTCCATCGTTGTCAGCATGGTGTCCTCTTCATTGCTCACTGTTCGGGCTGCCATCC
CCATTATCATGGGGGCCAACATTGGAACGTCAATCACCAACACTATTGTTGCGCTCATGCA
GGTGGGAGATCGGAGTGAGTTCAGAAGAGCTTTTGCAGGAGCCACTGTCCATGACTTCTTC
AACTGGCTGTCCGTGTTGGTGCTCTTGCCCGTGGAGGTGGCCACCCATTACCTCGAGATCA
TAACCCAGCTTATAGTGGAGAGCTTCCACTTCAAGAATGGAGAAGATGCCCCAGATCTTCT
GAAAGTCATCACTAAGCCCTTCACAAAGCTCATTGTCCAGCTGGATAAAAAAGTTATCAGC
CAAATTGCAATGAACGATGAAAAAGCGAAAAACAAGAGTCTTGTCAAGATTTGGTGCAAAA
CTTTTACCAACAAGACCCAGATTAACGTCACTGTTCCCTCGACTGCTAACTGCACCTCCCC
TTCCCTCTGTTGGACGGATGGCATCCAAAACTGGACCATGAAGAATGTGACCTACAAGGAG
AACATCGCCAAATGCCAGCATATCTTTGTGAATTTCCACCTCCCGGATCTTGCTGTGGGCA
CCATCTTGCTCATACTCTCCCTGCTGGTCCTCTGTGGTTGCCTGATCATGATTGTCAAGAT
CCTGGGCTCTGTGCTCAAGGGGCAGGTCGCCACTGTCATCAAGAAGACCATCAACACTGAT
TTCCCCTTTCCCTTTGCATGGTTGACTGGCTACCTGGCCATCCTCGTCGGGGCAGGCATGA
CCTTCATCGTACAGAGCAGCTCTGTGTTCACGTCGGCCTTGACCCCCCTGATTGGAATCGG
CGTGATAACCATTGAGAGGGCTTATCCACTCACGCTGGGCTCCAACATCGGCACCACCACC
ACCGCCATCCTGGCCGCCTTAGCCAGCCCTGGCAATGCATTGAGGAGTTCACTCCAGATCG
CCCTGTGCCACTTTTTCTTCAACATCTCCGGCATCTTGCTGTGGTACCCGATCCCGTTCAC
TCGCCTGCCCATCCGCATGGCCAAGGGGCTGGGCAACATCTCTGCCAAGTATCGCTGGTTC
GCCGTCTTCTACCTGATCATCTTCTTCTTCCTGATCCCGCTGACGGTGTTTGGCCTCTCGC
TGGCCGGCTGGCGGGTGCTGGTTGGTGTCGGGGTTCCCGTCGTCTTCATCATCATCCTGGT
ACTGTGCCTCCGACTCCTGCAGTCTCGCTGCCCACGCGTCCTGCCGAAGAAACTCCAGAAC
TGGAACTTCCTGCCGCTGTGGATGCGCTCGCTGAAGCCCTGGGATGCCGTCGTCTCCAAGT
TCACCGGCTGCTTCCAGATGCGCTGCTGCTACTGCTGCCGCGTGTGCTGCCGCGCGTGCTG
CTTGCTGTGTGGCTGCCCCAAGTGCTGCCGCTGCAGCAAGTGCTGCGAGGACTTGGAGGAG
GCGCAGGAGGGGCAGGATGTCCCTGTCAAGGCTCCTGAGACCTTTGATAACATAACCATTA
[illegible]CAAAGACCGAATGCACGGCCTTGTA

GGGGACGCCCCAGATTGTCAGGGATGGGGGGATGGTCCTTGAGTTTTGCATGCTCTCCTCC
CTCCCACTTCTGCACCCTTTCACCACCTCGAGGAGATTTGCTCCCCATTAGCGAATGAAAT
TGATGCAGTCCTAAAAAAAAAAA

FIGURE 29

CAGCCCAGCACCTGCGGAGGGAGCGCTGACCATGGCTCCCTGGCCTGAATTGGGAGATGCC
CAGCCCAACCCCGATAAGTACCTCGAAGGGGCCGCAGGTCAGCAGCCCACTGCCCCTGATA
AAAGCAAAGAGACCAACAAAACAGATAACACTGAGGCACCTGTAACCAAGATTGAACTTCT
GCCGTCCTACTCCACGGCTACACTGATAGATGAGCCCACTGAGGTGGATGACCCCTGGAAC
CTACCCACTCTTCAGGACTCGGGGATCAAGTGGTCAGAGAGAGACACCAAAGGGAAGATTC
TCTGTTTCTTCCAAGGGATTGGGAGATTGATTTTACTTCTCGGATTTCTCTACTTTTTCGT
GTGCTCCCTGGATATTCTTAGTAGCGCCTTCCAGCTGGTTGGAGGAAAAATGGCAGGACAG
TTCTTCAGCAACAGCTCTATTATGTCCAACCCTTTGTTGGGGCTGGTGATCGGGGTGCTGG
TGACCGTCTTGGTGCAGAGCTCCAGCACCTCAACGTCCATCGTTGTCAGCATGGTGTCCTC
TTCATTGCTCACTGTTCGGGCTGCCATCCCCATTATCATGGGGGCCAACATTGGAACGTCA
ATCACCAACACTATTGTTGCGCTCATGCAGGTGGGAGATCGGAGTGAGTTCAGAAGAGCTT
TTGCAGGAGCCACTGTCCATGACTTCTTCAACTGGCTGTCCGTGTTGGTGCTCTTGCCCGT
GGAGGTGGCCACCCATTACCTCGAGATCATAACCCAGCTTATAGTGGAGAGCTTCCACTTC
AAGAATGGAGAAGATGCCCCAGATCTTCTGAAAGTCATCACTAAGCCCTTCACAAAGCTCA
TTGTCCAGCTGGATAAAAAAGTTATCAGCCAAATTGCAATGAACGATGAAAAAGCGAAAAA
CAAGAGTCTTGTCAAGATTTGGTGCAAAACTTTTACCAACAAGACCCAGATTAACGTCACT
GTTCCCTCGACTGCTAACTGCACCTCCCCTTCCCTCTGTTGGACGGATGGCATCCAAAACT
GGACCATGAAGAATGTGACCTACAAGGAGAACATCGCCAAATGCCAGCATATCTTTGTGAA
TTTCCACCTCCCGGATCTTGCTGTGGGCACCATCTTGCTCATACTCTCCCTGCTGGTCCTC
TGTGGTTGCCTGATCATGATTGTCAAGATCCTGGGCTCTGTGCTCAAGGGGCAGGTCGCCA
CTGTCATCAAGAAGACCATCAACACTGATTTCCCCTTTCCCTTTGCATGGTTGACTGGCTA
CCTGGCCATCCTCGTCGGGGCAGGCATGACCTTCATCGTACAGAGCAGCTCTGTGTTCACG
TCGGCCTTGACCCCCCTGATTGGAATCGGCGTGATAACCATTGAGAGGGCTTATCCACTCA
CGCTGGGCTCCAACATCGGCACCACCACCACCGCCATCCTGGCCGCCTTAGCCAGCCCTGG
CAATGCATTGAGGAGTTCACTCCAGATCGCCCTGTGCCACTTTTTCTTCAACATCTCCGGC
ATCTTGCTGTGGTACCCGATCCCGTTCACTCGCCTGCCCATCCGCATGGCCAAGGGGCTGG
GCAACATCTCTGCCAAGTATCGCTGGTTCGCCGTCTTCTACCTGATCATCTTCTTCTTCCT
GATCCCGCTGACGGTGTTTGGCCTCTCGCTGGCCGGCTGGCGGGTGCTGGTTGGTGTCGGG
GTTCCCGTCGTCTTCATCATCATCCTGGTACTGTGCCTCCGACTCCTGCAGTCTCGCTGCC
CACGCGTCCTGCCGAAGAAACTCCAGAACTGGAACTTCCTGCCGCTGTGGATGCGCTCGCT
GAAGCCCTGGGATGCCGTCGTCTCCAAGTTCACCGGCTGCTTCCAGATGCGCTGCTGCTGC
TGCTGCCGCGTGTGCTGCCGCGCGTGCTGCTTGCTGTGTGGCTGCCCCAAGTGCTGCCGCT
GCAGCAAGTGCTGCGAGGACTTGGAGGAGGCGCAGGAGGGGCAGGATGTCCCTGTCAAGGC
TCCTGAGACCTTTGATAACATAACCATTAGCAGAGAGGCTCAGGGTGAGGTCCCTGCCTCG
[illegible]CAGATTGTCAGGGATGGGGGGA

TGGTCCTTGAGTTTTGCATGCTCTCCTCCCTCCCACTTCTGCACCCTTTCACCACCTCGAG
GAGATTTGCTCCCCATTAGCGAATGAAATTGATGCAGTCCTACCTAACTCGATTCCCTTTG
GCTTGGTGGGTAGGCCTGCAGGGCACTTTTATTCCAACCCATGGCCTCCATGACTTTTTCA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

GGCGTGGGAGCTGCCTAGGCCGGGCCCTGCCAGGGAGCAAGTCTGCATCGCAGAGGCCAGG
GCAGAGTGGCTCCTCACAGCCTGAAGCTCATCCTTCTGCACGGGCCAGCCAGGCCAGCACA
GAGGCACCAGGGCAGCAGTGCACACAGGTCCCCGGGGACCCCACCATGTGGAGCGGATGGT
GGCTGTGGCCCCTTGTGGCCGTCTGCACTGCAGACTTCTTTCGGGACGAGGCAGAGAGGAT
CATGAGGGACTCCCCTGTCATTGATGGGCACAATGACCTCCCCTGGCAGCTGCTGGATATG
TTCAACAACCGGCTGCAGGACGAGAGGGCCAACCTGACCACCTTGGCCGGCACACACACCA
ACATCCCCAAGCTGAGGGCCGGCTTTGTGGGAGGCCAGTTCTGGTCCGTGTACACGCCCTG
CGACACCCAGAACAAAGACGCCGTGCGGAGGACGCTGGAGCAGATGGACGTGGTCCACCGC
ATGTGCCGGATGTACCCGGAGACCTTCCTGTATGTCACCAGCAGTGCAGGCATTCGGCAGG
CCTTCCGGGAAGGGAAGGTGGCCAGCCTGATCGGCGTGGAGGGCGGCCACTCCATTGACAG
CAGTTTGGGCGTCCTGCGGGCACTCTATCAGCTGGGCATGCGGTACCTGACCCTCACCCAC
AGCTGCAACACGCCCTGGGCTGACAACTGGCTGGTGGACACGGGAGACAGCGAGCCCCAGA
GCCAAGGCTTGTCACCCTTTGGGCAGCGTGTGGTGAAGGAGCTGAACCGTCTGGGGGTCCT
CATCGACTTGGCTCACGTGTCTGTGGCCACCATGAAGGCCACCCTGCAGCTGTCCAGAGCC
CCGGTCATCTTCAGCCACTCCTCGGCCTACAGCGTGTGCGCAAGCCGGCGCAACGTGCCTG
ACGACGTCCTGAGGCTGGTGAAACAGACAGACAGCCTGGTGATGGTGAACTTCTACAACAA
TTACATTTCCTGCACCAACAAGGCCAACCTGTCCCAAGTGGCCGACCATCTGGATCACATC
AAGGAGGTGGCAGGAGCCAGAGCCGTGGGTTTTGGTGGGGACTTTGATGGTGTTCCAAGGG
TCCCTGAGGGGCTGGAGGACGTCTCCAAGTATCCAGACCTGATCGCTGAGCTGCTCAGGAG
GAACTGGACGGAGGCGGAGGTCAAGGGCGCACTGGCTGACAACCTGCTGAGGGTCTTCCAG
GCTGTGGAACAGGCCAGCAACCTCACACAGGCTCCCGAGGAGGAGCCCATCCCGCTGGACC
AGCTGGGTGGCTCCTGCAGGACCCATTACGGCTACTCCTCTGGGGCTTCCAGCCTCCATCG
CCACTGGGGGCTCCTGCTGGCCTCCCTCGCTCCCCTGGTCCTCTGTCTGTCTCTCCTGTGA
AACCTGGGAGACCAGAGTCCCCTTTAGGGTTCCCGGAGCTCCGGGAAGACCCGCCCATCCC
AGGACTCCAGATGCCAGGAGCCCTGCTGCCCACATGCAAGGACCAGCATCTCCTGAGAGGA
CGCCTGGGCTTACCTGGGGGGCAGGATGCCTGGGGACAGTTCAGGACACACACACAGTAGG
CCCGCAATAAAAGCAACACCCCTTCAAAAAAAAAAAAAAAAAAAAA

FIGURE 31

CATCTAAAGCCTCCTCAGCCTTCTGAGTCAGCCTGAAAGGAACAGGCCGAACTGCTGTATG
GGCTCTACTGCCAGTGTGACCTCACCCTCTCCAGTCACCCCTCCTCAGTTCCAGCTATGAG
TTCCTGCAACTTCACACATGCCACCTTTGTGCTTATTGGTATCCCAGGATTAGAGAAAGCC
CATTTCTGGGTTGGCTTCCCCCTCCTTTCCATGTATGTAGTGGCAATGTTTGGAAACTGCA
TCGTGGTCTTCATCGTAAGGACGGAACGCAGCCTGCACGCTCCGATGTACCTCTTTCTCTG
CATGCTTGCAGCCATTGACCTGGCCTTATCCACATCCACCATGCCTAAGATCCTTGCCCTT
TTCTGGTTTGATTCCCGAGAGATTAGCTTTGAGGCCTGTCTTACCCAGATGTTCTTTATTC
ATGCCCTCTCAGCCATTGAATCCACCATCCTGCTGGCCATGGCCTTTGACCGTTATGTGGC
CATCTGCCACCCACTGCGCCATGCTGCAGTGCTCAACAATACAGTAACAGCCCAGATTGGC
ATCGTGGCTGTGGTCCGCGGATCCCTCTTTTTTTTCCCACTGCCTCTGCTGATCAAGCGGC
TGGCCTTCTGCCACTCCAATGTCCTCTCGCACTCCTATTGTGTCCACCAGGATGTAATGAA
GTTGGCCTATGCAGACACTTTGCCCAATGTGGTATATGGTCTTACTGCCATTCTGCTGGTC
ATGGGCGTGGACGTAATGTTCATCTCCTTGTCCTATTTTCTGATAATACGAACGGTTCTGC
AACTGCCTTCCAAGTCAGAGCGGGCCAAGGCCTTTGGAACCTGTGTGTCACACATTGGTGT
GGTACTCGCCTTCTATGTGCCACTTATTGGCCTCTCAGTTGTACACCGCTTTGGAAACAGC
CTTCATCCCATTGTGCGTGTTGTCATGGGTGACATCTACCTGCTGCTGCCTCCTGTCATCA
ATCCCATCATCTATGGTGCCAAAACCAAACAGATCAGAACACGGGTGCTGGCTATGTTCAA
GATCAGCTGTGACAAGGACTTGCAGGCTGTGGGAGGCAAGTGACCCTTAACACTACACTTC
TCCTTATCTTTATTGGCTTGATAAACATAATTATTTCTAACACTAGCTTATTTCCAGTTGC
CCATAAGCACATCAGTACTTTTCTCTGGCTGGAATAGTAAACTAAAGTATGGTACATCTAC
CTAAAGGACTATTATGTGGAATAATACATACTAATGAAGTATTACATGATTTAAAGACTAC
AATAAAACCAAACATGCTTATAACATTAAGAAAAACAATAAAGATACATGATTGAAACCAA
GTTGAAAAATAGCATATGCCTTGGAGGAAATGTGCTCAAATTACTAATGATTTAGTGTTGT
CCCTACTTTCTCTCTCTTTTTTCTTTCTTTTTTTTTTATTATGGTTAGCTGTCACATACAA
CTTTTTTTTTTTTTTGAGATGGGGTCTCCAGCCTGGGCAACAGAGCAAGACCCTGTCTCAAA
GCATAAAATGGAATAACATATCAAATGAAACAGGGAAAATGAAGCTGACAATTTATGGGAG
CCA

FIGURE 32A

CACACATACGCACGCACGATCTCACTTCGATCTATACACTGGAGGATTAAAACAAACAAAC
AAAAAAAACATTTCCTTCGCTCCCCCTCCCTCTCCACTCTGAGAAGCAGAGGAGCCGCACG
GCGAGGGGCCGCAGACCGTCTGGAAATGCGAATCCTAAAGCGTTTCCTCGCTTGCATTCAG
CTCCTCTGTGTTTGCCGCCTGGATTGGGCTAATGGATACTACAGACAACAGAGAAAACTTG
TTGAAGAGATTGGCTGGTCCTATACAGGAGCACTGAATCAAAAAAATTGGGGAAAGAAATA
TCCAACATGTAATAGCCCAAAACAATCTCCTATCAATATTGATGAAGATCTTACACAAGTA
AATGTGAATCTTAAGAAACTTAAATTTCAGGGTTGGGATAAAACATCATTGGAAAACACAT
TCATTCATAACACTGGGAAAACAGTGGAAATTAATCTCACTAATGACTACCGTGTCAGCGG
AGGAGTTTCAGAAATGGTGTTTAAAGCAAGCAAGATAACTTTTCACTGGGGAAAATGCAAT
ATGTCATCTGATGGATCAGAGCATAGTTTAGAAGGACAAAAATTTCCACTTGAGATGCAAA
TCTACTGCTTTGATGCGGACCGATTTTCAAGTTTTGAGGAAGCAGTCAAAGGAAAAGGGAA
GTTAAGAGCTTTATCCATTTTGTTTGAGGTTGGGACAGAAGAAAATTTGGATTTCAAAGCG
ATTATTGATGGAGTCGAAAGTGTTAGTCGTTTTGGGAAGCAGGCTGCTTTAGATCCATTCA
TACTGTTGAACCTTCTGCCAAACTCAACTGACAAGTATTACATTTACAATGGCTCATTGAC
ATCTCCTCCCTGCACAGACACAGTTGACTGGATTGTTTTTAAAGATACAGTTAGCATCTCT
GAAAGCCAGTTGGCTGTTTTTTGTGAAGTTCTTACAATGCAACAATCTGGTTATGTCATGC
TGATGGACTACTTACAAAACAATTTTCGAGAGCAACAGTACAAGTTCTCTAGACAGGTGTT
TTCCTCATACACTGGAAAGGAAGAGATTCATGAAGCAGTTTGTAGTTCAGAACCAGAAAAT
GTTCAGGCTGACCCAGAGAATTATACCAGCCTTCTTGTTACATGGGAAAGACCTCGAGTCG
TTTATGATACCATGATTGAGAAGTTTGCAGTTTTGTACCAGCAGTTGGATGGAGAGGACCA
AACCAAGCATGAATTTTTGACAGATGGCTATCAAGACTTGGGTGCTATTCTCAATAATTTG
CTACCCAATATGAGTTATGTTCTTCAGATAGTAGCCATATGCACTAATGGCTTATATGGAA
AATACAGCGACCAACTGATTGTCGACATGCCTACTGATAATCCTGAACTTGATCTTTTCCC
TGAATTAATTGGAACTGAAGAAATAATCAAGGAGGAGGAAGAGGGAAAAGACATTGAAGAA
GGCGCTATTGTGAATCCTGGTAGAGACAGTGCTACAAACCAAATCAGGAAAAAGGAACCCC
AGATTTCTACCACAACACACTACAATCGCATAGGGACGAAATACAATGAAGCCAAGACTAA
CCGATCCCCAACAAGAGGAAGTGAATTCTCTGGAAAGGGTGATGTTCCCAATACATCTTTA
AATTCCACTTCCCAACCAGTCACTAAATTAGCCACAGAAAAAGATATTTCCTTGACTTCTC
AGACTGTGACTGAACTGCCACCTCACACTGTGGAAGGTACTTCAGCCTCTTTAAATGATGG
CTCTAAAACTGTTCTTAGATCTCCACATATGAACTTGTCGGGGACTGCAGAATCCTTAAAT
ACAGTTTCTATAACAGAATATGAGGAGGAGAGTTTATTGACCAGTTTCAAGCTTGATACTG
GAGCTGAAGATTCTTCAGGCTCCAGTCCCGCAACTTCTGCTATCCCATTCATCTCTGAGAA
CATATCCCAAGGGTATATATTTTCCTCCGAAAACCCAGAGACAATAACATATGATGTCCTT
ATACCAGAATCTGCTAGAAATGCTTCCGAAGATTCAACTTCATCAGGTTCAGAAGAATCAC
TAAAGGATCCTTTCTATGGAGGGAAATGTGTGGTTTTCCTAGCTCTTACAGACATAACAGCACA

GCCCGATGTTGGATCAGGCAGAGAGAGCTTTCTCCAGACTAATTACACTGAGATACGTGTT
GATGAATCTGAGAAGACAACCAAGTCCTTTTCTGCAGGCCCAGTGATGTCACAGGGTCCCT
CAGTTACAGATCTGGAAATGCCACATTATTCTACCTTTGCCTACTTCCCAACTGAGGTAAC
ACCTCATGCTTTTACCCCATCCTCCAGACAACAGGATTTGGTCTCCACGGTCAACGTGGTA
TACTCGCAGACAACCCAACCGGTATACAATGGTGAGACACCTCTTCAACCTTCCTACAGTA
GTGAAGTCTTTCCTCTAGTCACCCCTTTGTTGCTTGACAATCAGATCCTCAACACTACCCC
TGCTGCTTCAAGTAGTGATTCGGCCTTGCATGCTACGCCTGTATTTCCCAGTGTCGATGTG
TCATTTGAATCCATCCTGTCTTCCTATGATGGTGCACCTTTGCTTCCATTTTCCTCTGCTT
CCTTCAGTAGTGAATTGTTTCGCCATCTGCATACAGTTTCTCAAATCCTTCCACAAGTTAC
TTCAGCTACCGAGAGTGATAAGGTGCCCTTGCATGCTTCTCTGCCAGTGGCTGGGGGTGAT
TTGCTATTAGAGCCCAGCCTTGCTCAGTATTCTGATGTGCTGTCCACTACTCATGCTGCTT
CAGAGACGCTGGAATTTGGTAGTGAATCTGGTGTTCTTTATAAAACGCTTATGTTTTCTCA
AGTTGAACCACCCAGCAGTGATGCCATGATGCATGCACGTTCTTCAGGGCCTGAACCTTCT
TATGCCTTGTCTGATAATGAGGGCTCCCAACACATCTTCACTGTTTCTTACAGTTCTGCAA
TACCTGTGCATGATTCTGTGGGTGTAACTTATCAGGGTTCCTTATTTAGCGGCCCTAGCCA
TATACCAATACCTAAGTCTTCGTTAATAACCCCAACTGCATCATTACTGCAGCCTACTCAT
GCCCTCTCTGGTGATGGGGAATGGTCTGGAGCCTCTTCTGATAGTGAATTTCTTTTACCTG
ACACAGATGGGCTGACAGCCCTTAACATTTCTTCACCTGTTTCTGTAGCTGAATTTACATA
TACAACATCTGTGTTTGGTGATGATAATAAGGCGCTTTCTAAAAGTGAAATAATATATGGA
AATGAGACTGAACTGCAAATTCCTTCTTTCAATGAGATGGTTTACCCTTCTGAAAGCACAG
TCATGCCCAACATGTATGATAATGTAAATAAGTTGAATGCGTCTTTACAAGAAACCTCTGT
TTCCATTTCTAGCACCAAGGGCATGTTTCCAGGGTCCCTTGCTCATACCACCACTAAGGTT
TTTGATCATGAGATTAGTCAAGTTCCAGAAAATAACTTTTCAGTTCAACCTACACATACTG
TCTCTCAAGCATCTGGTGACACTTCGCTTAAACCTGTGCTTAGTGCAAACTCAGAGCCAGC
ATCCTCTGACCCTGCTTCTAGTGAAATGTTATCTCCTTCAACTCAGCTCTTATTTTATGAG
ACCTCAGCTTCTTTTAGTACTGAAGTATTGCTACAACCTTCCTTTCAGGCTTCTGATGTTG
ACACCTTGCTTAAAACTGTTCTTCCAGCTGTGCCCAGTGATCCAATATTGGTTGAAACCCC
CAAAGTTGATAAAATTAGTTCTACAATGTTGCATCTCATTGTATCAAATTCTGCTTCAAGT
GAAAACATGCTGCACTCTACATCTGTACCAGTTTTTGATGTGTCGCCTACTTCTCATATGC
ACTCTGCTTCACTTCAAGGTTTGACCATTTCCTATGCAAGTGAGAAATATGAACCAGTTTT
GTTAAAAAGTGAAAGTTCCCACCAAGTGGTACCTTCTTTGTACAGTAATGATGAGTTGTTC
CAAACGGCCAATTTGGAGATTAACCAGGCCCATCCCCCAAAAGGAAGGCATGTATTTGCTA
CACCTGTTTTATCAATTGATGAACCATTAAATACACTAATAAATAAGCTTATACATTCCGA
TGAAATTTTAACCTCCACCAAAAGTTCTGTTACTGGTAAGGTATTTGCTGGTATTCCAACA
GTTGCTTCTGATACATTTGTATCTACTGATCATTCTGTTCCTATAGGAAATGGGCATGTTG
CCATTACAGCTGTTTCTCCCCACAGAGATGGTTCTGTAACCTCAACAAAGTTGCTGTTTCC

GGTGAAGATGGTGACACTGATGATGATGGTGATGATGATGATGACAGAGATAGTGATGGCT
TATCCATTCATAAGTGTATGTCATGCTCATCCTATAGAGAATCACAGGAAAAGGTAATGAA

FIGURE 32B

TGATTCAGACACCCACGAAAACAGTCTTATGGATCAGAATAATCCAATCTCATACTCACTA
TCTGAGAATTCTGAAGAAGATAATAGAGTCACAAGTGTATCCTCAGACAGTCAAACTGGTA
TGGACAGAAGTCCTGGTAAATCACCATCAGCAAATGGGCTATCCCAAAAGCACAATGATGG
AAAAGAGGAAAATGACATTCAGACTGGTAGTGCTCTGCTTCCTCTCAGCCCTGAATCTAAA
GCATGGGCAGTTCTGACAAGTGATGAAGAAAGTGGATCAGGGCAAGGTACCTCAGATAGCC
TTAATGAGAATGAGACTTCCACAGATTTCAGTTTTGCAGACACTAATGAAAAAGATGCTGA
TGGGATCCTGGCAGCAGGTGACTCAGAAATAACTCCTGGATTCCCACAGTCCCCAACATCA
TCTGTTACTAGCGAGAACTCAGAAGTGTTCCACGTTTCAGAGGCAGAGGCCAGTAATAGTA
GCCATGAGTCTCGTATTGGTCTAGCTGAGGGGTTGGAATCCGAGAAGAAGGCAGTTATACC
CCTTGTGATCGTGTCAGCCCTGACTTTTATCTGTCTAGTGGTTCTTGTGGGTATTCTCATC
TACTGGAGGAAATGCTTCCAGACTGCACACTTTTACTTAGAGGACAGTACATCCCCTAGAG
TTATATCCACACCTCCAACACCTATCTTTCCAATTTCAGATGATGTCGGAGCAATTCCAAT
AAAGCACTTTCCAAAGCATGTTGCAGATTTACATGCAAGTAGTGGGTTTACTGAAGAATTT
GAGACACTGAAAGAGTTTTACCAGGAAGTGCAGAGCTGTACTGTTGACTTAGGTATTACAG
CAGACAGCTCCAACCACCCAGACAACAAGCACAAGAATCGATACATAAATATCGTTGCCTA
TGATCATAGCAGGGTTAAGCTAGCACAGCTTGCTGAAAAGGATGGCAAACTGACTGATTAT
ATCAATGCCAATTATGTTGATGGCTACAACAGACCAAAAGCTTATATTGCTGCCCAAGGCC
CACTGAAATCCACAGCTGAAGATTTCTGGAGAATGATATGGGAACATAATGTGGAAGTTAT
TGTCATGATAACAAACCTCGTGGAGAAAGGAAGGAGAAAATGTGATCAGTACTGGCCTGCC
GATGGGAGTGAGGAGTACGGGAACTTTCTGGTCACTCAGAAGAGTGTGCAAGTGCTTGCCT
ATTATACTGTGAGGAATTTTACTCTAAGAAACACAAAAATAAAAAAGGGCTCCCAGAAAGG
AAGACCCAGTGGACGTGTGGTCACACAGTATCACTACACGCAGTGGCCTGACATGGGAGTA
CCAGAGTACTCCCTGCCAGTGCTGACCTTTGTGAGAAAGGCAGCCTATGCCAAGCGCCATG
CAGTGGGGCCTGTTGTCGTCCACTGCAGTGCTGGAGTTGGAAGAACAGGCACATATATTGT
GCTAGACAGTATGTTGCAGCAGATTCAACACGAAGGAACTGTCAACATATTTGGCTTCTTA
AAACACATCCGTTCACAAAGAAATTATTTGGTACAAACTGAGGAGCAATATGTCTTCATTC
ATGATACACTGGTTGAGGCCATACTTAGTAAAGAAACTGAGGTGCTGGACAGTCATATTCA
TGCCTATGTTAATGCACTCCTCATTCCTGGACCAGCAGGCAAAACAAAGCTAGAGAAACAA
TTCCAGCTCCTGAGCCAGTCAAATATACAGCAGAGTGACTATTCTGCAGCCCTAAAGCAAT
GCAACAGGGAAAAGAATCGAACTTCTTCTATCATCCCTGTGGAAAGATCAAGGGTTGGCAT
TTCATCCCTGAGTGGAGAAGGCACAGACTACATCAATGCCTCCTATATCATGGGCTATTAC
CAGAGCAATGAATTCATCATTACCCAGCACCCTCTCCTTCATACCATCAAGGATTTCTGGA
GGATGATATGGGACCATAATGCCCAACTGGTGGTTATGATTCCTGATGGCCAAAACATGGC
ACAAGATGAATTTTGTTTACTGGCCAAATAAAGATGAGCCTATAAATTGTGAGAGCTTTAAG

GTCACTCTTATGGCTGAAGAACACAAATGTCTATCTAATGAGGAAAAACTTATAATTCAGG
ACTTTATCTTAGAAGCTACACAGGATGATTATGTACTTGAAGTGAGGCACTTTCAGTGTCC
TAAATGGCCAAATCCAGATAGCCCCATTAGTAAAACTTTTGAACTTATAAGTGTTATAAAA
GAAGAAGCTGCCAATAGGGATGGGCCTATGATTGTTCATGATGAGCATGGAGGAGTGACGG
CAGGAACTTTCTGTGCTCTGACAACCCTTATGCACCAACTAGAAAAAGAAAATTCCGTGGA
TGTTTACCAGGTAGCCAAGATGATCAATCTGATGAGGCCAGGAGTCTTTGCTGACATTGAG
CAGTATCAGTTTCTCTACAAAGTGATCCTCAGCCTTGTGAGCACAAGGCAGGAAGAGAATC
CATCCACCTCTCTGGACAGTAATGGTGCAGCATTGCCTGATGGAAATATAGCTGAGAGCTT
AGAGTCTTTAGTTTAACACAGAAAGGGGTGGGGGGACTCACATCTGAGCATTGTTTTCCTC
TTCCTAAAATTAGGCAGGAAAATCAGTCTAGTTCTGTTATCTGTTGATTTCCCATCACCTG
ACAGTAACTTTCATGACATAGGATTCTGCCGCCAAATTTATATCATTAACAATGTGTGCCT
TTTTGCAAGACTTGTAATTTACTTATTATGTTTGAACTAAAATGATTGAATTTTACAGTAT
TTCTAAGAATGGAATTGTGGTATTTTTTTCTGTATTGATTTTAACAGAAAATTTCAATTTA
TAGAGGTTAGGAATTCCAAACTACAGAAAATGTTTGTTTTTAGTGTCAAATTTTTAGCTGT
ATTTGTAGCAATTATCAGGTTTGCTAGAAATATAACTTTTAATACAGTAGCCTGTAAATAA
AACACTCTTCCATATGATATTCAACATTTTACAACTGCAGTATTCACCTAAAGTAGAAATA
ATCTGTTACTTATTGTAAATACTGCCCTAGTGTCTCCATGGACCAAATTTATATTTATAAT
TGTAGATTTTTATATTTTACTACTGAGTCAAGTTTTCTAGTTCTGTGTAATTGTTTAGTTT
AATGACGTAGTTCATTAGCTGGTCTTACTCTACCAGTTTTCTGACATTGTATTGTGTTACC
TAAGTCATTAACTTTGTTTCAGCATGTAATTTTAACTTTTGTGGAAAATAGAAATACCTTC
ATTTTGAAAGAAGTTTTTATGAGAATAACACCTTACCAAACATTGTTCAAATGGTTTTTAT
CCAAGGAATTGCAAAAATAAATATAAATATTGCCATTAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAA

FIGURE 33

ATTGTCTGGGAATTGCAGCCGCGGGGCGGGCGGCGGCGGCGGCGGCGGCCGGGACCCA
GCGGGCCAGGTGGGGACGGCGCGGAGCGGGTGCGGGAGATGCCGTGCGGGACTGGGGCCAC
CTGAGCCGCCCGCCTCGTCCCCGCCTTCTGTGGGAAGGATGTGCGCGCGGATGGCCGGTCG
CACAAGAGCGGCCCCTCGGGGGCCCTACGGCCCCTGGCTCTGCCTCCTGGTGGCCCTCGCC
CTGGACGTCGTGAGAGTGGACTGTGGCCAGGCTCCCCTGGACCCTGTCTACCTGCATGTGA
CAGCCGCCCGCCCAGCCCAGCCCACACTCTGGACTGCCAAGCTAGACCGCTTCAAGGGCTC
CAGGCACCACACCACCCTCATCACCTGCCACCGTGCTGGGCTCACAGAGCCAGATTCCAGC
AGTCCCCTTGAACTGTCTGAGTTCCTATGGGTGGACTTTGTGGTGGAGAATAGCACTGGTG
GGGGCGTAGCGGTCACTCGCCCCGTCACGTGGCAGCTGGAGTACCCAGGCCAGGCCCCTGA
AGCAGAGAAGGACAAAATGGTGTGGGAAATCCTGGTGTCTGAGCGGGACATCAGAGCCCTT
ATCCCACTGGCCAAGGCTGAGGAGCTGGTGAATACAGCACCACTGACTGGAGTGCCCCAGC
ATGTCCCCGTGCGCCTTGTCACTGTGGACGGCGGGGGGGCCTTGGTGGAGGTGACAGAGCA
TGTCGGCTGCGAGTCTGCCAACACACAGGTCCTGCAGGTGTCTGAGGCCTGTGATGCCGTG
TTCGTGGCTGGCAAGGAGAGCCGGGGCGCCCGGGGGGTGCGAGTGGACTTCTGGTGGCGCC
GGCTCCGCGCCTCGCTGCGGCTGACCGTGTGGGCCCCGCTGCTACCGCTGCGTATCGAGCT
CACCGACACCACCCTCGAGCAGGTCCGCGGCTGGAGGGTACCTGGCCCTGCTGAAGGGCCT
GCGGAACCCGCTGCAGAGGCGTCAGATGAGGCCGAGCGGCGCGCCCGTGGCTGCCACCTGC
AGTACCAGCGGGCCGGTGTGCGCTTCCTCGCCCCCTTCGCGGCCCACCCGCTGGACGGCGG
CCGCCGCCTCACGCACCTGCTTGGCCCCGACTGGCTGCTAGACGTGTCCCACCTCGTGGCG
CCACACGCCCGCGTGCTGGACTCGCGTGTAGCCTCTCTGGAGGGTGGCCGTGTCGTGGTGG
GCCGGGAGCCCGGTGTCACCTCCATTGAGGTGCGTTCCCCACTGTCTGACTCCATCCTGGG
GGAGCAGGCGCTGGCTGTGACGGACGACAAGGTCTCAGTGCTGGAGCTGAGGGTGCAGCCA
GTGATGGGCATCTCGCTGACCTTGAGCCGGGGCACTGCCCACCCCGGGGAGGTCACAGCTA
CGTGCTGGGCACAGTCAGCCCTTCCCGCCCCAAAGCAGGAGGTGGCCCTCTCCCTATGGCT
GTCCTTCTCTGATCACACTGTGGCCCCAGCTGAGCTCTACGACCGCCGTGACCTGGGACTG
TCCGTCTCAGCCGAGGAGCCTGGTGCCATCCTGCCAGCTGAGGAGCAGGGTGCCCAGCTCG
GGGTGGTGGTGAGTGGGGCAGGCGCCGAGGGGCTGCCGCTGCATGTGGCTCTGCACCCGCC
CGAGCCCTGCCGCCGGGGCCGCCACCGTGTGCCTCTGGCCTCTGGCACCGCCTGGCTGGGG
CTGCCCCCTGCCTCCACTCCAGCCCCTGCTCTCCCATCCAGCCCTGCTTGGAGCCCACCAG
CCACAGAAGCCACCATGGGTGGTAAACGGCAGGTGGCAGGCAGTGTCGGGGGCAACACAGG
TGTGAGGGGCAAGTTTGAGCGGGCAGAGGAGGAGGCCAGGAAGGAGGAGACCGAAGCCAGG
GAGGAGGAGGAGGAAGAGGAGGAGGAGATGGTCCCTGCCCCTCAGCATGTCACTGAGCTAG
AGCTGGGCATGTACGCCCTGCTGGGAGTCTTCTGCGTGGCCATCTTCATCTTCTTGGTCAA
TGGTGTGGTCTTCGTCCTGCGCTATCAGCGCAAAGAACCTCCCGACAGTGCCAGTGACCCC

AGCTGGACCGGCAGTCCCCTGGCCCGCCCAAGGGGGAGGGGAGCTGCCCCTGTGAGAGTGG
GGGAGGAGGGGAGGCCCCTACCCTGGCCCCTGGCCCTCCTGGGGGCACCACCAGCTCCTCA
AGCACCCTGGCCCGAAAGGAGGCTGGGGGGCGGCGGAAGCGAGTAGAGTTTGTGACATTTG
TGCCAGCCCCCTCCAGCCCAGTCACCTGAGGAGCCTGTAGGGGCCCCTGCTGTGCAGTCCAT
CCTTGTGGCAGGCGAGGAGGACATCCGCTGGGTGTGTGAGGACATGGGGCTGAAGGACCCT
GAGGAGCTTCGCAACTACATGGAGAGGATCCGGGGCAGCTCCTGACCCTCCACAGCCACCT
GGTCAGCCACCAGCTGGGGCAACGAGGGTGGAGGTCCCACTGAGCCTCTCGCCTGCCCCCG
CCACTCGTCTGGTGCTTGTTGATCCAAGTCCCCTGCCTGGTCCCCCACAAGGACTCCCATC
CAGGCCCCCTCTGCCCTGCCCCTTGTCATGGACCATGGTCGTGAGGAAGGGCTCATGCCCC
TTATTTATGGGAACCATCTCATTCTAACAGAATAAACCGAGAAGGAAACCAGAAAAAAAAA
AAAAAAAAAA

CAAACTTGCTGTCAATTCCGAGATCTAATCTCCCCCTACGCTCTGCCAGGAATTCTTTCAG
ACCTCACTAGCACAAGCCCGGTTGCTCCTTGTCAGGAGAATTTGTAGATCATTCTCACTTC
AAATTCCTGGGGCTGATACTTCTCTCATCTTGCACCCCAACCTCTGTAAATAGATTTACCG
CATTTACGGCTGCATTCTGTAAGTGGGCATGGTCTCCTAATGGAGGAGTGTTCATTGTATA
ATAAGTTATTCACCTGAGTATGCAATAAAGATGTGGTGGCCACTCTTTCATGGTGGTGGCA
GCAAAAAAAAAAAAAA

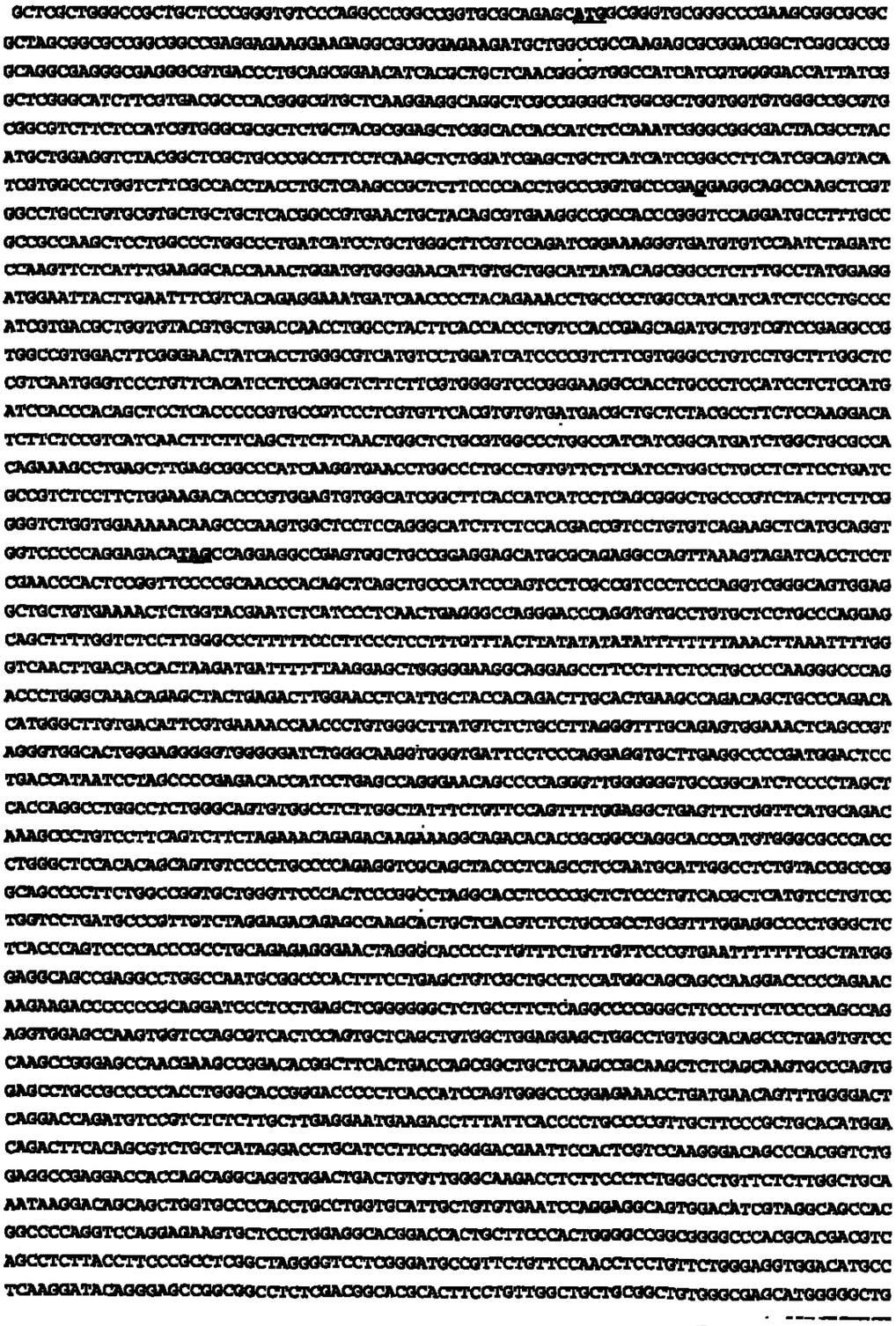
FIGURE 34

ACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGGTCGGGGCCAGAGCGCAGGTGTACCTGG
CGGCCGTGCTGGAGCACCTGACCGCCGAGATCCTGGAGCTGGCTGGCAACCCGGCCCGCGA
CAAGAAGACCCGCATCATCCTGCGCCACCTGTAGCTGGCCATTCGCAACGGCGAGGAGCTT
AACAAGCTGCTGGGCGAAGTCACCATCGCGCAGGGCGGTGTCCTGCCCAACATTCAGGGCG
TGCTTCTGCCCCAGAAGACCAAGAGCCACCACAAGGCCAAGGGTGAAAACCATTCACTAGG
AGAGGAGAAACACAATGGCCACCAAGACAGAGTTGAGTCCCACAGCAAGGGAGAGCAAGAA
CGCACAAGATATGCAAGTGGATGAGACACTGATCCCCAGGAAAGGTCCAAGTTTATGTTCT
GCTCGCTATGGAATAGCCCTCGTCTTACATTTCTGCAATTTCACAACGATAGCACAAAATG
TCATCATGAACATCACCATGGTAGCCATGGTCAACAGCACAAGCCCTCAATCCCAGCTCAA
TGATTCCTCTGAGGTGCTGCCTGTTGACTCATTTGGTGGCCTAAGTAAAGCCCCAAAGAGT
CTTCCTGCAAAGTCCTCAATACTTGGGGGTCAGTTTGCAATTTGGGAAAAGTGGGGCCCTC
CACAAGAACGAAGCAGACTCTGCAGCATTGCTTTATCAGGAATGTTACTGGGATGCTTTAC
TGCCATCCTCATAGGTGGCTTCATTAGTGAAACCCTTGGGTGGCCCTTTGTCTTCTATATC
TTTGGAGGTGTTGGCTGTGTCTGCTGCCTTCTCTGGTTTGTTGTGATTTATGATGACCCCT
TTTCCTATCCATGGATAAGCACCTCAGAAAAAGAATACATCATATCCTCCTTGAAACAACA
GGTCGGGTCTTCTAAGCAGCCTCTTCCCATCAAAGCTATGCTCAGATCTCTACCCATTTGG
TCCATATGTTTAGGCTGTTTCAGCCATCAATGGTTAGTTAGCACAATGGTTGTATACATAC
CAACTTACATCAGCTCTGTGTACCATGTTAACATCAGAGACAATGGACTTCTATCTGCCCT
TCCTTTTATTGTTGCCTGGGTCATAGGCATGGTGGGAGGCTATCTGGCAGATTTCCTTCTA
ACCAAAAAGTTTAGACTCATCACTGTGAGGAAAATTGCCACAATTTTAGGAAGTCTCCCCT
CTTCAGCACTCATTGTGTCTCTGCCTTACCTCAATTCCGGCTATATCACAGCAACTGCCTT
GCTGACGCTCTCTTGCGGATTAAGCACATTGTGTCAGTCAGGGATTTATATCAATGTCTTA
GATATTGCTCCAAGGTATTCCAGTTTTCTCATGGGAGCATCAAGAGGATTTTCGAGCATAG
CACCTGTCATTGTACCCACTGTCAGCGGATTTCTTCTTAGTCAGGACCCTGAGTTTGGGTG
GAGGAATGTCTTCTTCTTGCTGTTTGCCGTTAACCTGTTAGGACTACTCTTCTACCTCATA
TTTGGAGAAGCAGATGTCCAAGAATGGGCTAAAGAGAGAAAACTCACTCGTTTATGAAGTT
ATCCCACCTTGGATGGAAAAGTCATTAGGCACCGTATTGCATAAAATAGAAGGCTTCCGTG
ATGAAAATACCAGTGAAAAGATTTTTTTTCCTGTGGCTCTTTTCAATTATGAGATCAGTT
CATTATTTTATTCAGACTTTTTTTTGAGAGAAATGTAAGATGAATAAAAATTCAAATAAAA
TGATAACTAAGAAAAAAAAAAAAAAA

FIGURE 35

GTGAAAACCATTCACTAGGAGAGGAGAAACACAATGGCCACCAAGACAGAGTTGAGTCCCA
CAGCAAGGGAGAGCAAGAACGCACAAGATATGCAAGTGGATGAGACACTGATCCCCAGGAA
AGTTCCAAGTTTATGTTCTGCTCGCTATGGAATAGCCCTCGTCTTACATTTCTGCAATTTC
ACAACGATAGCACAAAATGTCATCATGAACATCACCATGGTAGCCATGGTCAACAGCACAA
GCCCTCAATCCCAGCTCAATGATTCCTCTGAGGTGCTGCCTGTTGACTCATTTGGTGGCCT
AAGTAAAGCCCCAAAGAGTCTTCCTACAAAGTCCTCAATACTTGGGGGTCAGTTTGCAATT
TGGGAAAAGTGGGGCCCTCCACAAGAACGAAGCAGACTCTGCAGCATTGCTTTATCAGGAA
TGTTACTGGGATGCTTTACTGCCATCCTCATAGGTGGCTTCATTAGTGAAACCCTTGGGTG
GCCCTTTGTCTTCTATATCTTTGGAGGTGTTGGCTGTGTCTGCTGCCTTCTCTGGTTTGTT
GTGATTTATGATGACCCCGTTTCCTATCCATGGATAAGCACCTCAGAAAAAGAATACATCA
TATCCTCCTTGAAACAACAGGTCAGGTCTTCTAAGCAGCCTCTTCCCATCAAAGCTATGCT
CAGATCTCTACCCATTTGGTCCATATGTTTAGGCTGTTTCAGCCATCAATGGTTAGTTAGC
ACAATGGTTGTATACATACCAACTTACATCAGCTCTGTGTACCATGTTAACATCAGAGACA
ATGGACTTCTATCTGCCCTTCCTTTTATTGTTGCCTGGGTCATAGGCATGGTGGGAGGCTA
TCTGGCAGATTTCCTTCTAACCAAAAAGTTTAGACTCATCACTGTGAGGAAAATTGCCACA
ATTTTAGGAAGTCTCCCCTCTTCAGCACTCATTGTGTCTCTGCCTTACCTCAATTCCGGCT
ATATCACAGCAACTGCCTTGCTGACGCTCTCTTGCGGATTAAGCACATTGTGTCAGTCAGG
GATTTATATCAATGTCTTAGATATTGCTCCAAGGTATTCCAGTTTTCTCATGGGAGCATCA
AGAGGATTTTCGAGCATAGCACCTGTCATTGTACCCACTGTCAGCGGATTTCTTCTTAGTC
AGGACCCTGAGTTTGGGTGGAGGAATGTCTTCTTCTTGCTGTTTGCCGTTAACCTGTTAGG
ACTACTCTTCTACCTCATATTTGGAGAAGCAGATGTCCAAGAATGGGCTAAAGAGAGAAAA
CTCACTCGTTTATGAAGTTATCCCACCTTGGATGGAAAAGTCATTAGGCACCGTATTGCAT
AAAATAGAAGGCTTCCGTGATGAAAATACCAGTGAAAAGATTTTTTTTTCCTGTGGCTCTT
TTCAATTATGAGATCAGTTCATTATTTTATTCAGACTTTTTTTTGAGAGAAATGTAAGATG
AATAAAAATTCAAATAAAATGATAACTAAGAAAAAAAAAAAAAA

FIGURE 36



GGGCGTCTGAGCTGGGGTGGGAGCTCCTCCGCAGAAGGTTGGTGGGGGGTCCAGTCTGTGATCCTTGGTGCTGTGTGCCCCA
CTCCAGCCTGGGGACCCCACTTCAGAAGGTAGGGGCCGTGTCCCGCGGTGCTGACTGAGGCCTGGTTCCCCCTCCCCCTCCT
GCTGTGCTGGAATTCCACAGGGACCAGGGCCACCGCAGGGGACTGTCTCAGAAGACTTGATTTTTCCGTCCCTTTTTCTCCA
CACTCCACTGACAAACGTCCCCAGCGGTTTCCACTTGTGGGCTTCAGGTGTTTTCAAGCACAACCCACCACAACAAGCAAGT
GCATTTTCAGTCGTTGTGCTTTTTTGTTTTGTGCTAACGTCTTACTAATTTAAAGATGCTGTCGGCACCATGTTTATTTATT
TCCAGTGGTCATGCTCAGCCTTGCTGCTCTGCGTGGCGCAGGTGCCATGCCTGCTCCCTGTCTGTGTCCCAGCCACGCAGGG
CCATCCACTGTGACGTCGGCCGACCAGGCTGGACACCCTCTGCCGAGTAATGACGTGTGTGGCTGGGACCTTCTTTATTCTG
TGTTAATGGCTAACCTGTTACACTGGGCTGGGTTGGGTAGGGTGTTCTGGCTTTTTTGTGGGGTTTTTATTTTAAAGAAAC
ACTCAATCATCCTAG

FIGURE 37

ATTCGCCCTTTGCTCCCGGGTGTCCCAGGCCCGGCCGGTGCGCAGAGCATGGCGGGTGCGG
GCCCGAAGCGGCGCGCGCTAGCGGCGCCGGCGGCCGAGGAGAAGGAAGAGGCGCGGGAGAA
GATGCTGGCCGCCAAGAGCGCGGACGGCTCGGCGCCGGCAGGCGAGGGCGAGGGCGTGACC
CTGCAGCGGAACATCACGCTGCTCAACGGCGTGGCCATCATCGTGGGGACCATTATCGGCT
CGGGCATCTTCGTGACGCCCACGGGCGTGCTCAAGGAGGCAGGCTCGCCGGGGCTGGCGCT
GGTGGTGTGGGCCGCGTGCGGCGTCTTCTCCATCGTGGGCGCGCTCTGCTACGCGGAGCTC
GGCACCACCATCTCCAAATCGGGCGGCGACTACGCCTACATGCTGGAGGTCTACGGCTCGC
TGCCCGCCTTCCTCAAGCTCTGGATCGAGCTGCTCATCATCCGGCCTTCATCGCAGTACAT
CGTGGCCCTGGTCTTCGCCACCTACCTGCTCAAGCCGCTCTTCCCCACCTGCCCGGTGCCC
GAGGAGGCAGCCAAGCTCGTGGCCTGCCTCTGCGTGCTGCTGCTCACGGCCGTGAACTGCT
ACAGCGTGAAGGCCGCCACCCGGGTCCAGGATGCCTTTGCCGCCGCCAAGCTCCTGGCCCT
GGCCCTGATCATCCTGCTGGGCTTCGTCCAGATCGGGAAGGGTGATGTGTCCAATCTAGAT
CCCAACTTCTCATTTGAAGGCACCAAACTGGATGTGGGGAACATTGTGCTGGCATTATACA
GCGGCCTCTTTGCCTATGGAGGATGGAATTACTTGAATTTCGTCACAGAGGAAATGATCAA
CCCCTACAGAAACCTGCCCCTGGCCATCATCATCTCCCTGCCCATCGTGACGCTGGTGTAC
GTGCTGACCAACCTGGCCTACTTCACCACCCTGTCCACCGAGCAGATGCTGTCGTCCGAGG
CCGTGGCCGTGGACTTCGGGAACTATCACCTGGGCGTCATGTCCTGGATCATCCCCGTCTT
CGTGGGCCTGTCCTGCTTCGGCTCCGTCAATGGGTCCCTGTTCACATCCTCCAGGCTCTTC
TTCGTGGGGTCCCGGGAAGGCCACCTGCCCTCCATCCTCTCCATGATCCACCCACAGCTCC
TCACCCCCGTGCCGTCCCTCGTGTTCACGTGTGTGATGACGCTGCTCTACGCCTTCTCCAA
GGACATCTTCTCCGTCATCAACTTCTTCAGCTTCTTCAACTGGCTCTGCGTGGCCCTGGCC
ATCATCGGCATGATCTGGCTGCGCCACAGAAAGCCTGAGCTTGAGCGGCCCATCAAGGTGA
ACCTGGCCCTGCCTGTGTTCTTCATCCTGGCCTGCCTCTTCCTGATCGCCGTCTCCTTCTG
GAAGACACCCGTGGAGTGTGGCATCGGCTTCACCATCATCCTCAGCGGGCTGCCCGTCTAC
TTCTTCGGGGTCTGGTGGAAAAACAAGCCCAAGTGGCTCCTCCAGGGCATCTTCTCCACGA
CCGTCCTGTGTCAGAAGCTCATGCAGGTGGTCCCCCAGGAGACATAGCCAGGAGGCCGAAG
GGCG

FIGURE 38

CCAGGCTCTCCACCCCCACTTCCCAATTGAGGAAACCGAGGCAGAGGAGGCTCAGGTGTGG
CCAATCACCCTGCACATCAGAGTTACCCTGGGCAGGGCCCACTGAGACCTGGGAGGGGCCA
CTCGGGACCTGGAGGGCTGGGGGCTGCCCGGGCGTTAGGGGTAAAGCTCCCTACCCAACTG
CGCAGAAGGCCTCAGAGGCCTGGGGGCTGGGCTTCCCCTTTCACATCGCCCTTTAGAGGCC
CACGTGTGGGCATTGGCCCGCGATCTGAAAGGGGCTGTCCTGTTCCTCATGGGCGCTGCCA
GCGCCACGCACTCCTCTTTCTGCCTGGCCGGCCACTCCCGTCTGCTGTGACGCGCGGACAG
AGAGCTACCGGTGGACCCACGGTGCCTCCCTCCCTGGGATCTACACAGACCATGGCCTTGC
CAACGGCTCGACCCCTGTTGGGGTCCTGTGGGACCCCCGCCCTCGGCAGCCTCCTGTTCCT
GCTCTTCAGCGTCGGATGGGTGCAGCCCTCGAGGACCCTGGCTGGAGAGACAGGGCAGGAG
GCTGCACCCCTGGACGGAGTCCTGGCCAACCCACCTAACATTTCCAGCCTCTCCCCTCGCC
AACTCCTTGGCTTCCCGTGTGCGGAGGTGTCCGGCCTGAGCACGGAGCGTGTCCGGGAGCT
GGCTGTGGCCTTGGCACAGAAGAATGTCAAGCTCTCAACAGAGCAGCTGCGCTGTCTGGCT
CACCGGCTCTCTGAGCCCCCCGAGGACCTGGACGCCCTCCCATTGGACCTGCTGCTATTCC
TCAACCCAGATGCGTTCTCGGGGCCCCAGGCCTGCACCCGTTTCTTCTCCCGCATCACGAA
GGCCAATGTGGACCTGCTCCCGAGGGGGGCTCCCGAGCGACAGCGGCTGCTGCCTGCGGCT
CTGGCCTGCTGGGGTGTGCGGGGGTCTCTGCTGAGCGAGGCTGATGTGCGGGCTCTGGGAG
GCCTGGCTTGCGACCTGCCTGGGCGCTTTGTGGCCGAGTCGGCCGAAGTGCTGCTACCCCG
GCTGGTGAGCTGCCCGGGACCCCTGGACCAGGACCAGCAGGAGGCAGCCAGGGCGGCTCTG
CAGGGCGGGGGACCCCCCTACGGCCCCCCGTCGACATGGTCTGTCTCCACGATGGACGCTC
TGCGGGGCCTGCTGCCCGTGCTGGGCCAGCCCATCATCCGCAGCATCCCGCAGGGCATCGT
GGCCGCGTGGCGGCAACGCTCCTCTCGGGACCCATCCTGGCGGCAGCCTGAACGGACCATC
CTCCGGCCGCGGTTCCGGCGGGAAGTGGAGAAGACAGCCTGTCCTTCAGGCAAGAAGGCCC
GCGAGATAGACGAGAGCCTCATCTTCTACAAGAAGTGGGAGCTGGAAGCCTGCGTGGATGC
GGCCCTGCTGGCCACCCAGATGGACCGCGTGAACGCCATCCCCTTCACCTACGAGCAGCTG
GACGTCCTAAAGCATAAACTGGATGAGCTCTACCCACAAGGTTACCCCGAGTCTGTGATCC
AGCACCTGGGCTACCTCTTCCTCAAGATGAGCCCTGAGGACATTCGCAAGTGGAATGTGAC
GTCCCTGGAGACCCTGAAGGCTTTGCTTGAAGTCAACAAAGGGCACGAAATGAGTCCTCAG
GTGGCCACCCTGATCGACCGCTTTGTGAAGGGAAGGGGCCAGCTAGACAAAGACACCCTAG
ACACCCTGACCGCCTTCTACCCTGGGTACCTGTGCTCCCTCAGCCCCGAGGAGCTGAGCTC
CGTGCCCCCCAGCAGCATCTGGGCGGTCAGGCCCCAGGACCTGGACACGTGTGACCCAAGG
CAGCTGGACGTCCTCTATCCCAAGGCCCGCCTTGCTTTCCAGAACATGAACGGGTCCGAAT
ACTTCGTGAAGATCCAGTCCTTCCTGGGTGGGGCCCCCACGGAGGATTTGAAGGCGCTCAG
TCAGCAGAATGTGAGCATGGACTTGGCCACGTTCATGAAGCTGCGGACGGATGCGGTGCTG
CCGTTGACTGTGGCTGAGGTGCAGAAACTTCTGGGACCCCACGTGGAGGGCCTGAAGGCGG

GCTGGGGCTGGGGCTACAGGGCGGCATCCCCAACGGCTACCTGGTCCTAGACCTCAGCGTG
CAAGAGGCCCTCTCGGGGACGCCCTGCCTCCTAGGACCTGGACCTGTTCTCACCGTCCTGG
CACTGCTCCTAGCCTCCACCCTGGCCTGAGGGCCCCACTCCCTTGCTGGCCCCAGCCCTGC
TGGGGATCCCCGCCTGGCCAGGAGCAGGCACGGGTGATCCCCGTTCCACCCCAAGAGAACT
CGCGCTCAGTAAACGGGAACATGCCCCCTGCAGACACGTAAAAAAAAAAAAAAAAAA

FIGURE 39

TGGGATCTACACAGACCATGGCCTTGCCAACGGCTCGACCCCTGTTGGGGTCCTGTGGGAC
CCCCGCCCTCGGCAGC

CTCCTGTTCCTGCTCTTCAGCCTCGGATGGGTGCAGCCCTCGAGGACCCTGGCTGGAGAGA
CAGGGCAGGAGGCTGCGCCCCTGGACGGAGTCCTGGCCAACCCACCTAACATTTCCAGCCT
CTCCCCTCGCCAACTCCTTGGCTTCCCGTGTGCGGAGGTGTCCGGCCTGAGCACGGAGCGT
GTCCGGGAGCTGGCTGTGGCCTTGGCACAGAAGAATGTCAAGCTCTCAACAGAGCAGCTGC
GCTGTCTGGCTCACCGGCTCTCTGAGCCCCCCGAGGACCTGGACGCCCTCCCATTGGACCT
GCTGCTATTCCTCAACCCAGATGCGTTCTCGGGGCCCCAGGCCTGCACCCGTTTCTTCTCC
CGCATCACGAAGGCCAATGTGGACCTGCTCCCGAGGGGGGCTCCCGAGCGACAGCGGCTGC
TGCCTGCGGCTCTGGCCTGCTGGGGTGTGCGGGGGTCTCTGCTGAGCGAGGCTGATGTGCG
GGCTCTGGGAGGCCTGGCTTGCGACCTGCCTGGGCGCTTTGTGGCCGAGTCGGCCGAAGTG
CTGCTACCCCGGCTGGTGAGCTGCCCGGGACCCCTGGACCAGGACCAGCAGGAGGCAGCCA
GGGCGGCTCTGCAGGGCGGGGGACCCCCCTACGGCCCCCCGTCGACATGGTCTGTCTCCAC
GATGGACGCTCTGCGGGGCCTGCTGCCCGTGCTGGGCCAGCCCATCATCCGCAGCATCCCG
CAGGGCATCGTGGCCGCGTGGCGGCAACGCTCCTCTCGGGACCCATCCTGGCGGCAGCCTG
AACGGACCATCCTCCGGCCGCGGTTCCGGCGGGAAGTGGAGAAGACAGCCTGTCCTTCAGG
CAAGAAGGCCCGCGAGATAGACGAGAGCCTCATCTTCTACAAGAAGTGGGAGCTGGAAGCC
TGCGTGGATGCGGCCCTGCTGGCCACCCAGATGGACCGCGTGAACGCCATCCCCTTCACCT
ACGAGCAGCTGGACGTCCTAAAGCATAAACTGGATGAGCTCTACCCACAAGGTTACCCCGA
GTCTGTGATCCAGCACCTGGGCTACCTCTTCCTCAAGATGAGCCCTGAGGACATTCGCAAG
TGGAATGTGACGTCCCTGGAGACCCTGAAGGCTTTGCTTGAAGTCAACAAAGGGCACGAAA
TGAGTCCTCAGGTGGCCACCCTGATCGACCGCTTTGTGAAGGGAAGGGGCCAGCTAGACAA
AGACACCCTAGACACCCTGACCGCCTTCTACCCTGGGTACCTGTGCTCCCTCAGCCCCGAG
GAGCTGAGCTCCGTGCCCCCCAGCAGCATCTGGGCGGTCAGGCCCCAGGACCTGGACACGT
GTGACCCAAGGCAGCTGGACGTCCTCTATCCCAAGGCCCGCCTTGCTTTCCAGAACATGAA
CGGGTCCGAATACTTCGTGAAGATCCAGTCCTTCCTGGGTGGGGCCCCCACGGAGGATTTG
AAGGCGCTCAGTCAGCAGAATGTGAGCATGGACTTGGCCACGTTCATGAAGCTGCGGACGG
ATGCGGTGCTGCCGTTGACTGTGGCTGAGGTGCAGAAACTTCTGGGACCCCACGTGGAGGG
CCTGAAGGCGGAGGAGCGGCACCGCCCGGTGCGGGACTGGATCCTACGGCAGCGGCAGGAC
GACCTGGACACGCTGGGGCTGGGGCTACAGGGCGGCATCCCCAACGGCTACCTGGTCCTAG
ACCTCAGCGTGCAAGAGGCCCTCTCGGGGACGCCCTGCCTCCTAGGACCTGGACCTGTTCT
CACCGTCCTGGCACTGCTCCTAGCCTCCACCCTGGCCTGAGGGCCCCACTCCCTTGCTGGC
CCCAGCCCTGCTGGGGATCCCCGCCTGGCCAGGAGCAGGCACGGGTGATCCCCGTTCCACC
CCAAGAGAACT

FIGURE 40

ATAACAGCATGAAGTGCCGTGGAACTGGAATAGGCGTGTCCTCTCCCTCGACCCTCCCCCT
CCTTGTCCCTCTGCTCACCCCTCGCTCGTTCCCTCCCTCCGGCGAGGGCCGCCTTTATAAC
AACTGCTCAGAGTGCGAGGGCGGGATAGCTGTCCAAGGTCTCCCCCAGCACTGAGGAGCTC
GCCTGCTGCCCTCTTGCGCGCGGGAAGCAGCACCAAGTTCACGGCCAACGCCTTGGCACTA
GGGTCCAGAATGGCTACAACAGTCCCTGATGGTTGCCGCAATGGCCTGAAATCCAAGTACT
ACAGACTTTGTGATAAGGCTGAAGCTTGGGGCATCGTCCTAGAAACGGTGGCCACAGCCGG
GGTTGTGACCTCGGTGGCCTTCATGCTCACTCTCCCGATCCTCGTCTGCAAGGTGCAGGAC
TCCAACAGGCGAAAAATGCTGCCTACTCAGTTTCTCTTCCTCCTGGGTGTGTTGGGCATCT
TTGGCCTCACCTTCGCCTTCATCATCGGACTGGACGGGAGCACAGGGCCCACACGCTTCTT
CCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCCTGCCTGCTGGCTCATGCTGTCAGTCTG
ACCAAGCTCGTCCGGGGGAGGAAGCCCCTTTCCCTGTTGGTGATTCTGGGTCTGGCCGTGG
GCTTCAGCCTAGTCCAGGATGTTATCGCTATTGAATATATTGTCCTGACCATGAATAGGAC
CAACGTCAATGTCTTTTCTGAGCTTTCCGCTCCTCGTCGCAATGAAGACTTTGTCCTCCTG
CTCACCTACGTCCTCTTCTTGATGGCGCTGACCTTCCTCATGTCCTCCTTCACCTTCTGTG
GTTCCTTCACGGGCTGGAAGAGACATGGGGCCCACATCTACCTCACGATGCTCCTCTCCAT
TGCCATCTGGGTGGCCTGGATCACCCTGCTCATGCTTCCTGACTTTGACCGCAGGTGGGAT
GACACCATCCTCAGCTCCGCCTTGGCTGCCAATGGCTGGGTGTTCCTGTTGGCTTATGTTA
GTCCCGAGTTTTGGCTGCTCACAAAGCAACGAAACCCCATGGATTATCCTGTTGAGGATGC
TTTCTGTAAACCTCAACTCGTGAAGAAGAGCTATGGTGTGGAGAACAGAGCCTACTCTCAA
GAGGAAATCACTCAAGGTTTTGAAGAGACAGGGGACACGCTCTATGCCCCCTATTCCACAC
ATTTTCAGCTGCAGAACCAGCCTCCCCAAAAGGAATTCTCCATCCCACGGGCCCACGCTTG
GCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGGCAGCTAACTCTGTCCTGAAGAG
TGGGACAAATGCAGCCGGGCGGCAGATCTAGCGGGAGCTCAAAGGGATGTGGGCGAAATCT
TGAGTCTTCTGAGAAAACTGTACAAGACACTACGGGAACAGTTTGCCTCCCTCCCAGCCTC
AACCACAATTCTTCCATGCTGGGGCTGATGTGGGCTAGTAAGACTCCAGTTCTTAGAGGCG
CTGTAGTATTTTTTTTTTTTTGTCTCATCCTTTGGATACTTCTTTTAAGTGGGAGTCTCAG
GCAACTCAAGTTTAGACCCTTACTCTTTTTGTTTGTTTTTTGAAACAGGATCTTGCTCTGT
CACCCAGGCTTGAGTGCAGTGGTGCGATCACAGCCCAGTGCAGCCTCGACCACCTGTGCTC
AAGCAATCCTCCCATCTCCATCTCCCAAAGTGCTGGGATGACAGGCGTGAGCCACAGCTCC
CAGCCTAGGCCCTTAATCTTGCTGTTATTTTCCATGGACTAAAGGTCTGGTCATCTGAGCT
CACGCTGGCTCACACAGCTCTAGGGGCCTGCTCCTCTAACTCACAGTGGGTTTTGTGAGGC
TCTGTGGCCCAGAGCAGACCTGCATATCTGAGCAAAAATAGCAAAAGCCTCTCTCAGCCCA
CTGGCCTGAATCTACACTGGAAGCCAACTTGCTGGCACCCCCGCTCCCCAACCCTTCTTGC
CTGGGTAGGAGAGGCTAAAGATCACCCTAAATTTACTCATCTCTCTAGTGCTGCCTCACAT

CAAACTTGCTGTCAATTCCGAGATCTAATCTCCCCCTACGCTCTGCCAGGAATTCTTTCAG
ACCTCACTAGCACAAGCCCGGTTGCTCCTTGTCAGGAGAATTTGTAGATCATTCTCACTTC
AAATTCCTGGGGCTGATACTTCTCTCATCTTGCACCCCAACCTCTGTAAATAGATTTACCG
CATTTACGGCTGCATTCTGTAAGTGGGCATGGTCTCCTAATGGAGGAGTGTTCATTGTATA
ATAAGTTATTCACCTGAGTATGCAATAAAGATGTGGTGGCCACTCTTTCATGGTGGTGGCA
GCAAAAAAAAAAAAAA

FIGURE 41

AGAGACTCAAGATGATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTGTTAA
CCCTATAAACGCCAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGGTCGG
GACCAAGGCCCAAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATACTTCA
GCACTTGTAAGAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTTTTATATGA
ATGTTGCCCTGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCCCATT
GACCATGTTTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCTGACG
CCTCAAAACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGAGTAA
TGAGGCTTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAATGTT
GAATTACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGACTTAA
AAAATGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATTATCC
TAATGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAACAAAT
GGTGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGACTTCA
TTGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATATTGGA
GGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGAGAAA
CTTCCACGAGGTGTCCTAGAAAGGTTCATGGGAGACAAAGTGGCTTCCGAAGCTCTTATGA
AGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCTTTGA
GACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAATGGA
ATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGATTGATC
AGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCAAACCAC
CTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGGAGAATAC
ACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGTTCAGCGCC
TCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAATGAGCTTTA
CAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGTATATCGTACA
GCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAGAAACGGTGCGA
TTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGAAAAGTTAAAACA
AGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTTGAAAGAGCTCCTG
ACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTTTAAGGGAATGACTA
GTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAACATCATTCTTTATCA
CCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGTTACTAACATTTTAAAG
ACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATACACTTCTGGTGAATGAAT
TGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAATTCATGTTGTAGATAAACT
CCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAAATTA
ATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTGACTG
TCTATACAACTAAAATTATAACCAAAGTTGTGGAACCAAAAATTAAAGTGATTGAAGGCAG
TCTTTCAGCCTATTATCAAAACTGAAGGACCCACACTAACAAAAGTCAAAATTGAAGGTGAA

CCTGAATTCAGACTGATTAAAGAAGGTGAAACAATAACTGAAGTGATCCATGGAGAGCCAA
TTATTAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGAC
ACGAGAAGAACGAATCATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGT
GGAGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGAGGTCACCAAGGTCACCAAAT
TCATTGAAGGTGGTGATGGTCATTTATTTGAAGATGAAGAAATTAAAAGACTGCTTCAGGG
AGACACACCCGTGAGGAAGTTGCAAGCCAACAAAAAAGTTCAAGGTTCTAGAAGACGATTA
AGGGAAGGTCGTTCTCAGTGAAAATCCAAAAACCAGAAAAAAATGTTTATACAACCCTAAG
TCAATAACCTGACCTTAGAAAATTGTGAGAGCCAAGTTGACTTCAGGAACTGAAACATCAG
CACAAAGAAGCAATCATCAAATAATTCTGAACACAAATTTAATATTTTTTTTTCTGAATGA
GAAACATGAGGGAAATTGTGGAGTTAGCCTCCTGTGGTAAAGGAATTGAAGAAAATATAAC
ACCTTACACCCTTTTTCATCTTGACATTAAAAGTTCTGGCTAACTTTGGAATCCATTAGAG
AAAAATCCTTGTCACCAGATTCATTACAATTCAAATCGAAGAGTTGTGAACTGTTATCCCA
TTGAAAAGACCGAGCCTTGTATGTATGTTATGGATACATAAAATGCACGCAAGCCATTATC
TCTCCATGGGAAGCTAAGTTATAAAAATAGGTGCTTGGTGTACAAAACTTTTTATATCAAA
AGGCTTTGCACATTTCTATATGAGTGGGTTTACTGGTAAATTATGTTATTTTTTACAACTA
ATTTTGTACTCTCAGAATGTTTGTCATATGCTTCTTGCAATGCATATTTTTTAATCTCAAA
CGTTTCAATAAAACCATTTTTCAGATATAAAGAGAATTACTTCAAATTGAGTAATTCAGAA
AAACTCAAGATTTAAGTTAAAAAGTGGTTTGGACTTGGGAA

FIGURE 42

CCCTTGAGACTCAAGATGATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTG
TTAACCCTATAAACGCCAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGG
TCGGGACCAAGGCCCAAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATAC
TTCAGCACTTGTAAGAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTGTTAT
ATGAATGTTGCCCTGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCC
CATTGACCATGTTTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCT
GACGCCTCAAAACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGA
GTAATGAGGCTTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAA
TGTTGAATTACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGAC
TTAAAAAATGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATT
ATCCTAATGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAAC
AAATGGTGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGAC
TTCATTGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATAT
TGGAGGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGA
GAAACTTCCACGAGGTGTCCTAGAAAGGATCATGGGAGACAAAGTGGCTTCCGAAGCTCTT
ATGAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCT
TTGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
TGGAATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGATT
GATCAGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCAAA
CCACCTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGGAGA
ATACACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGATCAG
CGCCTCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAATGAGC
TTTACAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGTATATCG
TACAGCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAGAAACGGT
GCGATTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGAAAAGTTAA
AACAAGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTTGAAAGAGCT
CCTGACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTTTAAGGGAATG
ACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAACATCATTCTTT
ATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGTTACTAACATTTT
AAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATACACTTCTGGTGAAT
GAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAATTCATGTTGTAGATA
AACTCCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAA
ATTAATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTG
ACTGTCTATAGACCCACACTAACAAAAGTCAAAATTGAAGGTGAACCTGAATTCAGACTGA

CAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGAATC
ATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAA
CTCTGAAGAAATTGTTACAAGAAGAGGTCACCAAGGTCACCAAATTCATTGAAGGTGGTGA
TGGTCATTTATTTGAAGATGAAGAAATTAAAAGACTGCTTCAGGGAGACACACCCGTGAGG
AAGTTGCAAGCCAACAAAAAAGTTCAAGGATCTAGAAGACGATTAAGGGAAGGTCGTTCTC
AGTGAAGGGC

FIGURE 43

CCCTTGAGACTCAAGATGATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTG
TTAACCCTATAAACGCCAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGG
TCGGGACCAAGGCCCAAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATAC
TTCAGCACTTGTAAGAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTGTTAT
ATGAATGTTGCCCTGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCC
CATTGACCATGTTTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCT
GACGCCTCAAAACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGA
GTAATGAGGCTTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAA
TGTTGAATTACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGAC
TTAAAAAATGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATT
ATCCTAATGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAAC
AAATGGTGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGAC
TTCATTGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATAT
TGGAGGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGA
GAAACTTCCACGAGGTGTCCTAGAAAGGATCATGGGAGACAAAGTGGCTTCCGAAGCTCTT
ATGAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCT
TTGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
TGGAATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGATT
GATCAGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCAAA
CCACCTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGGAGA
ATACACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGATCAG
CGCCTCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAATGAGC
TTTACAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGTATATCG
TACAGCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAGAAACGGT
GCGATTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGAAAAGTTAA
AACAAGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTTGAAAGAGCT
CCTGACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTTTAAGGGAATG
ACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAACATCATTCTTT
ATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGTTACTAACATTTT
AAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATACACTTCTGGTGAAT
GAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAATTCATGTTGTAGATA
AACTCCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAA
ATTAATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTG
ACTGTCTATAAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAA
[illegible]GAATCATTACAGGTCCTGAAATAAAATACACTAG

GATTTCTACTGGAGGTGGAGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGACACA
CCCGTGAGGAAGTTGCAAGCCAACAAAAAAGTTCAAGGATCTAGAAGACGATTAAGGGAAG
GTCGTTCTCAGTGAAGGGC

FIGURE 44

CCCTTGAGACTCAAGATGATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTG
TTAACCCTATAAACGCCAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGG
TCGGGACCAAGGCCCAAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATAC
TTCAGCACTTGTAAGAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTGTTAT
ATGAATGTTGCCCTGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCC
CATTGACCATGTTTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCT
GACGCCTCAAAACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGA
GTAATGAGGCTTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAA
TGTTGAATTACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGAC
TTAAAAAATGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATT
ATCCTAATGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAAC
AAATGGTGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGAC
TTCATTGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATAT
TGGAGGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGA
GAAACTTCCACGAGGTGTCCTAGAAAGGATCATGGGAGACAAAGTGGCTTCCGAAGCTCTT
ATGAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCT
TTGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
TGGAATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGATT
GATCAGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCAAA
CCACCTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGGAGA
ATACACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGATCAG
CGCCTCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAATGAGC
TTTACAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGTATATCG
TACAGCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAGAAACGGT
GCGATTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGAAAAGTTAA
AACAAGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTTGAAAGAGCT
CCTGACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTTTAAGGGAATG
ACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAACATCATTCTTT
ATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGTTACTAACATTTT
AAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATACACTTCTGGTGAAT
GAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAATTCATGTTGTAGATA
AACTCCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAA
ATTAATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTG
ACTGTCTATAGACCCACACTAACAAAAGTCAAAATTGAAGGTGAACCTGAATTCAGACTGA
TTTAAAGAAGCTGAAAGAATAACTGAAGTGATCCATGGAGAGCCAATTATTAAAAAATACAC

CAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGAATC
ATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAA
CTCTGAAGAAATTGTTACAAGAAGACACACCCGTGAGGAAGTTGCAAGCCAACAAAAAAGT
TCAAGGATCTAGAAGACGATTAAGGGAAGGTCGTTCTCAGTGAAGGGC

FIGURE 45

CCCTTGAGACTCAAGATGATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTG
TTAACCCTATAAACGCCAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGG
TCGGGACCAAGGCCCAAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATAC
TTCAGCACTTGTAAGAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTGTTAT
ATGAATGTTGCCCTGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCC
CATTGACCATGTTTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCT
GACGCCTCAAAACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGA
GTAATGAGGCTTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAA
TGTTGAATTACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGAC
TTAAAAAATGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATT
ATCCTAATGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAAC
AAATGGTGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGAC
TTCATTGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATAT
TGGAGGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGA
GAAACTTCCACGAGGTGTCCTAGAAAGGATCATGGGAGACAAAGTGGCTTCCGAAGCTCTT
ATGAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCT
TTGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
TGGAATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGATT
GATCAGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCAAA
CCACCTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGGAGA
ATACACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGATCAG
CGCCTCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAATGAGC
TTTACAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGTATATCG
TACAGCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAGAAACGGT
GCGATTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGAAAAGTTAA
AACAAGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTTGAAAGAGCT
CCTGACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTTTAAGGGAATG
ACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAACATCATTCTTT
ATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGTTACTAACATTTT
AAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATACACTTCTGGTGAAT
GAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAATTCATGTTGTAGATA
AACTCCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAA
ATTAATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTG
ACTGTCTATAAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAA
TAACTTGAAAAAGAGACACGAGAAGAACGAATCATTACAGGTCCTGAAATAAAATACACTAG

GATTTCTACTGGAGGTGGAGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGAGGTC
ACCAAGGTCACCAAATTCATTGAAGGTGGTGATGGTCATTTATTTGAAGATGAAGAAATTA
AAAGACTGCTTCAGGGAGACACACCCGTGAGGAAGTTGCAAGCCAACAAAAAAGTTCAAGG
ATCTAGAAGACGATTAAGGGAAGGTCGTTCTCAGTGAAGGGC

FIGURE 46

ACACGGACCAAGGAGTCTAACACGTGCGCGAGTCGGGGGCTCGCACGAAAGCCGCCGTGGC
GCAATGAAGGTGAAGGCCGGCGCGCTCGCCGGCCGAGGTGGGATCCCGAGGCCTCTCCAGT
CCGCCGAGGGCGCACCACCGGCCCGTCTCGCCCGCCGCGCCGGGGAGGTGGAGCACGAGCG
CACGTGTTAGGACCCGAAAGATGGTGAACTATGCCTGGGCAGGGCGAAGCCAGAGGAAACT
CTGGTGGAGGTCCGTAGCGGTCCTGACGTGCAAATCGGTCGTCCGACCTGGGTATAGGGGC
GGGCTCCAGGCGAGGCGGTCGACGCTCCTGAAAACTTGCGCGCGCGCTCGCGCCACTGCGC
CCGGAGCGATGAAGATGGTCGCGCCCTGGACGCGGTTCTACTCCAACAGCTGCTGCTTGTG
CTGCCATGTCCGCACCGGCACCATCCTGCTCGGCGTCTGGTATCTGATCATCAATGCTGTG
GTACTGTTGATTTTATTGAGTGCCCTGGCTGATCCGGATCAGTATAACTTTTCAAGTTCTG
AACTGGGAGGTGACTTTGAGTTCATGGATGATGCCAACATGTGCATTGCCATTGCGATTTC
TCTTCTCATGATCCTGATATGTGCTATGGCTACTTACGGAGCGTACAAGCAACGCGCAGCC
TGGATCATCCCATTCTTCTGTTACCAGATCTTTGACTTTGCCCTGAACATGTTGGTTGCAA
TCACTGTGCTTATTTATCCAAACTCCATTCAGGAATACATACGGCAACTGCCTCCTAATTT
TCCCTACAGAGATGATGTCATGTCAGTGAATCCTACCTGTTTGGTCCTTATTATTCTTCTG
TTTATTAGCATTATCTTGACTTTTAAGGGTTACTTGATTAGCTGTGTTTGGAACTGCTACC
GATACATCAATGGTAGGAACTCCTCTGATGTCCTGGTTTATGTTACCAGCAATGACACTAC
GGTGCTGCTACCCCCGTATGATGATGCCACTGTGAATGGTGCTGCCAAGGAGCCACCGCCA
CCTTACGTGTCTGCCTAAGCCTTCAAGTGGGCGGAGCTGAGGGCAGCAGCTTGACTTTGCA
GACATCTGAGCAATAGTTCTGTTATTTCACTTTTGCCATGAGCCTCTCTGAGCTTGTTTGT
TGCTGAAATGCTACTTTTTAAAATTTAGATGTTAGATTGAAAACTGTAGTTTTCAACATAT
GCTTTGCTAGAACACTGTGATAGATTAACTGTAGAATTCTTCCTGTACGATTGGGGATATA
ATGGGCTTCACTAACCTTCCCTAGGCATTGAAACTTCCCCCAAATCTGATGGACCTAGAAG
TCTGCTTTTGTACCTGCTGGGCCCCAAAGTTGGGCATTTTTCTCTCTGTTCCCTCTCTTTT
GAAAATGTAAAATAAAACCAAAAATAGACAACTTTTTCTTCAGCCATTCCAGCATAGAGAA
CAAAACCTTATGGAAACAGGAATGTCAATTGTGTAATCATTGTTCTAATTAGGTAAATAGA
AGTCCTTATGTATGTGTTACAAGAATTTCCCCCACAACATCCTTTATGACTGAAGTTCAAT
GACAGTTTGTGTTTGGGTGGTAAAGGATTTTCTCCATGGCCTGAATTAAGACCATTAGAAA
GCACCAGGCCGTGGGAGCAGTGACCATCTGCTGACTGTTCTTGTGGATCTTGTGTCCAGGG
ACATGGGGTGACATGCCTCGTATGTGTTAGAGGGTGGAATGGATGTGTTTGGCGCTGCATG
GGATCTGGTGCCCCTCTTCTCCTGGATTCACATCCCCACCCAGGGCCCGCTTTTACTAAGT
GTTCTGCCCTAGATTGGTTCAAGGAGGTCATCCAACTGACTTTATCAAGTGGAATTGGGAT
ATATTTGATATACTTCTGCCTAACAACATGGAAAAGGGTTTTCTTTTCCCTGCAAGCTACA
TCCTACTGCTTTGAACTTCCAAGTATGTCTAGTCACCTTTTAAAATGTAAACATTTTCAGA
AAAATGAGGATTGCCTTCCTTGTATGCGCTTTTTACCTTGACTACCTGAATTGCAAGGGAT

TGTAAATTAAGTTGTTTGCAATTAAAACAAGGTTTGCCCACAAAAAAAAAA

FIGURE 47

ATTCGCCCTTCCGAAAGATGGTGAACTATGCCTGGGCAGGGCGGGCTCCAGGCGAGGCGGT
CGACGCTCCTGAAAACTTGCGCGCGCGCTCGCGCCACTGCGCCCGGAGCGATGAAGATGGT
CGCGCCCTGGACTCGGTTCTACTCCAACAGCTGCTGCTTGTGCTGCCATGTCCGCACCGGC
ACCATCCTGCTCGGCGTCTGGTATCTGATCATCAATGCTGTGGTACTGTTGATTTTATTGA
GTGCCCTGGCTGATCCGGATCAGTATAACTTTTCAAGTTCTGAACTGGGAGGTGACTTTGA
GTTCATGGATGATGCCAACATGTGCATTGCCATTGCGATTTCTCTTCTCATGATCCTGATA
TGTGCTATGGCTACTTACGGAGCGTACAAGCAACGCGCAGCCTGGATCATCCCATTCTTCT
GTTACCAGATCTTTGACTTTGCCCTGAACATGTTGGTTGCAATCACTGTGCTTATTTATCC
AAACTCCATTCAGGAATACATACGGCAACTGCCTCCTAATTTTCCCTACAGAGATGATGTC
ATGTCAGTGAATCCTACCTGTTTGGTCCTTATTATTCTTCTGTTTATTAGCATTATCTTGA
CTTTTAAGGGTTACTTGATTAGCTGTGTTTGGAACTGCTACCGATACATCAATGGTAGGAA
CTCCTCTGATGTCCTGGTTTATGTTACCAGCAATGACACTACGGTGCTGCTACCCCCGTAT
GATGATGCCACTGTGAATGGTGCTGCCAAGGAGCCACCGCCACCTTACGTGTCTGCCTAAG
CCTTCAAGTGGGCGGAGCTGAGGGCAGCAGCTTGACTTTGCAGACATCTGAGAAGGGC

FIGURE 48

AGCGCCATGGGGCGAGCCGGGGGCGGGGGCCCGGGCCGGGGGCCGCCGCCACTGCTGCTGT
TTCTGGGGGCCGCGCTGGTCCTGGCCTCTGGGGCCGTGCCGGCGCGTGAGGCGGGCAGCGC
GGTTGAGGCCGAAGAGCTGGTGAAGGGCAGCCCGGCGTGGGAGCCGCCTGCCAACGACACG
CGGGAAGAAGCCGGCCCACCAGCGGCTGGGGAAGATGAGGCGTCGTGGACGGCGCCCGGTG
GCGAGCTGGCCGGGCCAGAAGAGGTGCTGCAGGAGTCGGCTGCGGTGACCGGCACCGCCTG
GCTGGAAGCTGACAGCCCAGGCCTGGGAGGAGTGACCGCAGAGGCGGGCAGCGGCGATGCC
CAGGCCCTTCCAGCTACGCTCCAGGCTCCCCACGAGGTCCTCGGGCAGTCAATCATGCCCC
CTGCCATTCCTGAGGCTACAGAGGCCAGCGGGCCACCCTCCCCCACCCCCGGCGACAAGCT
GAGCCCAGCTTCTGAACTCCCCAAGGAGAGCCCCTTGGAGGTTTGGCTGAACCTGGGGGGC
AGCACACCCGACCCTCAAGTGCCAGAGCTGACTTACCCATTTCAGGGCACCCTGGAGCCCC
AACCGGCATCAGATATCATTGACATCGACTACTTCGAAGGACTGGATGGTGAGGGTCGTGG
CGCAGATCTGGGGAGCTTCCCAGGGTCACCAGGAACCTCAGAGAACCACCCTGATACTGAG
GGAGAGACCCCTTCCTGGAGCCTGCTTGACTTATACGATGATTTCACCCCCTTCGATGAAT
CTGATTTCTACCCCACCACATCCTTTTATGATGACTTGGATGAAGAGGAGGAGGAAGAGGA
GGATGACAAAGATGCAGTAGGAGGTGGAGACCTAGAAGATGAAAATGAGCTTCTAGTGCCC
ACTGGGAAGCCTGGTCTGGGGCCCGGGACAGGCCAGCCCACCAGTCGGTGGCATGCTGTCC
CTCCACAGCACACTCTGGGGTCGGTCCCCGGCAGCAGCATCGCCCTCAGGCCCCGCCCAGG
AGAGCCAGGCAGGGACTTGGCCTCCAGTGAAAATGGCACTGAGTGCCGCAGTGGCTTTGTG
CGGCATAACGGCTCCTGCCGGTCAGTGTGCGACCTCTTCCCAAGTTACTGTCACAATGGCG
GCCAGTGCTACCTGGTGGAGAACATAGGGGCCTTCTGCAGGTGCAACACGCAGGACTACAT
CTGGCACAAGGGGATGCGCTGCGAGTCCATCATCACCGACTTCCAGGTGATGTGCGTGGCC
GTGGGCTCGGCTGCCCTCGTCCTGCTCCTGCTCTTCATGATGACGGTGTTCTTTGCCAAGA
AGCTCTACCTGCTCAAGACGGAGAATACCAAGCTGCGTAGGACCAACAAATTCCGGACCCC
ATCTGAGCTCCACAATGATAACTTCTCCCTCTCCACCATTGCCGAGGGCTCTCACCCAAAT
GATGATCCTAGTGCTCCCCACAAAATCCAGGAGGTTCTCAAGTCCTGCCTGAAAGAGGAGG
AGTCATTTAACATCCAGAACTCCATGTCGCCCAAACTTGAGGGTGGCAAAGGTGACCAGGC
TGACTTGGATGTGAACTGTCTTCAGAATAATTTAACCTAAAGCAGAGCAAGAAGAGAGGAA
GCGGGGTAGTGGGTGGGGGTAGGGAAGAAACATTATCTCCTCTTGTACAGAGTCTATTTCT
TGTAACCATTTGTTAAACTC

FIGURE 49

GGACAGGCGTGGCGGCCGGAGCCCCAGCATCCCTGCTTGAGGTCCAGGAGCGGAGCCCGCG
GCCACCGCCGCCTGATCAGCGCGACCCCGGCCCGCGCCCGCCCCGCCCGGCAAGATGCTGC
CCGTGTACCAGGAGGTGAAGCCCAACCCGCTGCAGGACGCGAACATCTGCTCACGCGTGTT
CTTCTGGTGGCTCAATCCCTTGTTTAAAATTGGCCATAAACGGAGATTAGAGGAAGATGAT
ATGTATTCAGTGCTGCCAGAAGACCGCTCACAGCACCTTGGAGAGGAGTTGCAAGGGTTCT
GGGATAAAGAAGTTTTAAGAGCTGAGAATGACGCACAGAAGCCTTCTTTAACAAGAGCAAT
CATAAAGTGTTACTGGAAATCTTATTTAGTTTTGGGAATTTTTACGTTAATTGAGGAAAGT
GCCAAAGTAATCCAGCCCATATTTTTGGGAAAAATTATTAATTATTTTGAAAATTATGATC
CCATGGATTCTGTGGCTTTGAACACAGCGTACGCCTATGCCACGGTGCTGACTTTTTGCAC
GCTCATTTTGGCTATACTGCATCACTTATATTTTTATCACGTTCAGTGTGCTGGGATGAGG
TTACGAGTAGCCATGTGCCATATGATTTATCGGAAGGCACTTCGTCTTAGTAACATGGCCA
TGGGGAAGACAACCACAGGCCAGATAGTCAATCTGCTGTCCAATGATGTGAACAAGTTTGA
TCAGGTGACAGTGTTCTTACACTTCCTGTGGGCAGGACCACTGCAGGCGATCGCAGTGACT
GCCCTACTCTGGATGGAGATAGGAATATCGTGCCTTGCTGGGATGGCAGTTCTAATCATTC
TCCTGCCCTTGCAAAGCTGTTTTGGGAAGTTGTTCTCATCACTGAGGAGTAAAACTGCAAC
TTTCACGGATGCCAGGATCAGGACCATGAATGAAGTTATAACTGGTATAAGGATAATAAAA
ATGTACGCCTGGGAAAAGTCATTTTCAAATCTTATTACCAATTTGAGAAAGAAGGAGATTT
CCAAGATTCTGAGAAGTTCCTGCCTCAGGGGGATGAATTTGGCTTCGTTTTTCAGTGCAAG
CAAAATCATCGTGTTTGTGACCTTCACCACCTACGTGCTCCTCGGCAGTGTGATCACAGCC
AGCCGCGTGTTCGTGGCAGTGACGCTGTATGGGGCTGTGCGGCTGACGGTTACCCTCTTCT
TCCCCTCAGCCATTGAGAGGGTGTCAGAGGCAATCGTCAGCATCCGAAGAATCCAGACCTT
TTTGCTACTTGATGAGATATCACAGCGCAACCGTCAGCTGCCGTCAGATGGTAAAAAGATG
GTGCATGTGCAGGATTTTACTGCTTTTTGGGATAAGGCATCAGAGACCCCAACTCTACAAG
GCCTTTCCTTTACTGTCAGACCTGGCGAATTGTTAGCTGTGGTCGGCCCCGTGGGAGCAGG
GAAGTCATCACTGTTAAGTGCCGTGCTCGGGGAATTGGCCCCAAGTCACGGGCTGGTCAGC
GTGCATGGAAGAATTGCCTATGTGTCTCAGCAGCCCTGGGTGTTCTCGGGAACTCTGAGGA
GTAATATTTTATTTGGGAAGAAATATGAAAAGGAACGATATGAAAAAGTCATAAAGGCTTG
TGCTCTGAAAAAGGATTTACAGCTGTTGGAGGATGGTGATCTGACTGTGATAGGAGATCGG
GGAACCACGCTGAGTGGAGGGCAGAAAGCACGGGTAAACCTTGCAAGAGCAGTGTATCAAG
ATGCTGACATCTATCTCCTGGACGATCCTCTCAGTGCAGTAGATGCGGAAGTTAGCAGACA
CTTGTTCGAACTGTGTATTTGTCAAATTTTGCATGAGAAGATCACAATTTTAGTGACTCAT
CAGTTGCAGTACCTCAAAGCTGCAAGTCAGATTCTGATATTGAAAGATGGTAAAATGGTGC
AGAAGGGGACTTACACTGAGTTCCTAAAATCTGGTATAGATTTTGGCTCCCTTTTAAAGAA
GGATAATGAGGAAAGTGAACAACCTCCAGTTCCAGGAACTCCCACACTAAGGAATCGTACC
TTCTTCAGAGTCTTTCGGTTTGGTCTCAACAATCTTTCTAGACCCTCCTTTGAAAGATGGTGCTC

TGGAGAGCCAAGATACAGAGAATGTCCCAGTTACACTATCAGAGGAGAACCGTTCTGAAGG
AAAAGTTGGTTTTCAGGCCTATAAGAATTACTTCAGAGCTGGTGCTCACTGGATTGTCTTC
ATTTTCCTTATTCTCCTAAACACTGCAGCTCAGGTTGCCTATGTGCTTCAAGATTGGTGGC
TTTCATACTGGGCAAACAAACAAAGTATGCTAAATGTCACTGTAAATGGAGGAGGAAATGT
AACCGAGAAGCTAGATCTTAACTGGTACTTAGGAATTTATTCAGGTTTAACTGTAGCTACC
GTTCTTTTTGGCATAGCAAGATCTCTATTGGTATTCTACGTCCTTGTTAACTCTTCACAAA
CTTTGCACAACAAAATGTTTGAGTCAATTCTGAAAGCTCCGGTATTATTCTTTGATAGAAA
TCCAATAGGAAGAATTTTAAATCGTTTCTCCAAAGACATTGGACACTTGGATGATTTGCTG
CCGCTGACGTTTTTAGATTTCATCCAGACATTGCTACAAGTGGTTGGTGTGGTCTCTGTGG
CTGTGGCCGTGATTCCTTGGATCGCAATACCCTTGGTTCCCCTTGGAATCATTTTCATTTT
TCTTCGGCGATATTTTTTGGAAACGTCAAGAGATGTGAAGCGCCTGGAATCTACAACTCGG
AGTCCAGTGTTTTCCCACTTGTCATCTTCTCTCCAGGGGCTCTGGACCATCCGGGCATACA
AAGCAGAAGAGAGGTGTCAGGAACTGTTTGATGCACACCAGGATTTACATTCAGAGGCTTG
GTTCTTGTTTTTGACAACGTCCCGCTGGTTCGCCGTCCGTCTGGATGCCATCTGTGCCATG
TTTGTCATCATCGTTGCCTTTGGGTCCCTGATTCTGGCAAAAACTCTGGATGCCGGGCAGG
TTGGTTTGGCACTGTCCTATGCCCTCACGCTCATGGGGATGTTTCAGTGGTGTGTTCGACA
AAGTGCTGAAGTTGAGAATATGATGATCTCAGTAGAAAGGGTCATTGAATACACAGACCTT
GAAAAAGAAGCACCTTGGGAATATCAGAAACGCCCACCACCAGCCTGGCCCCATGAAGGAG
TGATAATCTTTGACAATGTGAACTTCATGTACAGTCCAGGTGGGCCTCTGGTACTGAAGCA
TCTGACAGCACTCATTAAATCACAAGAAAAGGTTGGCATTGTGGGAAGAACCGGAGCTGGA
AAAAGTTCCCTCATCTCAGCCCTTTTTAGATTGTCAGAACCCGAAGGTAAAATTTGGATTG
ATAAGATCTTGACAACTGAAATTGGACTTCACGATTTAAGGAAGAAAATGTCAATCATACC
TCAGGAACCTGTTTTGTTCACTGGAACAATGAGGAAAAACCTGGATCCCTTTAAGGAGCAC
ACGGATGAGGAACTGTGGAATGCCTTACAAGAGGTACAACTTAAAGAAACCATTGAAGATC
TTCCTGGTAAAATGGATACTGAATTAGCAGAATCAGGATCCAATTTTAGTGTTGGACAAAG
ACAACTGGTGTGCCTTGCCAGGGCAATTCTCAGGAAAAATCAGATATTGATTATTGATGAA
GCGACGGCAAATGTGGATCCAAGAACTGATGAGTTAATACAAAAAAAAATCCGGGAGAAAT
TTGCCCACTGCACCGTGCTAACCATTGCACACAGATTGAACACCATTATTGACAGCGACAA
GATAATGGTTTTAGATTCAGGAAGACTGAAAGAATATGATGAGCCGTATGTTTTGCTGCAA
AATAAAGAGAGCCTATTTTACAAGATGGTGCAACAACTGGGCAAGGCAGAAGCCGCTGCCC
TCACTGAAACAGCAAAACAGGTATACTTCAAAAGAAATTATCCACATATTGGTCACACTGA
CCACATGGTTACAAACACTTCCAATGGACAGCCCTCGACCTTAACTATTTTCGAGACAGCA
CTGTGAATCCAACCAAAATGTCAAGTCCGTTCCGAAGGCATTTTCCACTAGTTTTTGGACT
ATGTAAACCACATTGTACTTTTTTTTACTTTGGCAACAAATATTTATACATACAAGATGCT
AGTTCATTTGAATATTTCTCCC

FIGURE 50

CTCCGTCCCAGGGGAGAAGGAGAGGCGTCTGCAGGGGGCAGAGACCGCAGCTACCTGCCGG
GTGCGCCCCCCACCCAGGAGCGCTCGCTTCGCCCCCTTTCCTCCCCCCGCCCCCACCTCCT
TATTGGTGCTAGTTTGCAGCGCCCAGCTCCTGCGCCTTCGCTTCGCGTTTGAATCTGGCTC
GCCCCTTCGTATTATGTCTGCACTCCGAAGGAAATTTGGGGACGATTATCAGGTAGTGACC
ACATCGTCCAGCGGCTCGGGCTTGCAGCCCCAGGGGCCAGGCCAGGACCCTCAGCAGCAGC
TTGTGCCCAAGAAGAAGCGGCAGCGGTTCGTGGACAAGAACGGCCGGTGCAATGTACAGCA
CGGCAACCTGGGCAGCGAGACAAGCCGCTACCTCTCGGACCTCTTCACCACGCTGGTGGAC
CTCAAGTGGCGCTGGAACCTCTTCATCTTCATTCTCACCTACACCGTGGCCTGGCTTTTCA
TGGCGTCCATGTGGTGGGTGATCGCCTACACTCGGGGCGACCTGAACAAAGCCCACGTCGG
TAACTACACGCCTTGCGTGGCCAATGTCTATAACTTCCCTTCTGCCTTCCTCTTCTTCATC
GAGACGGAGGCCACCATCGGCTATGGCTACCGATACATCACAGACAAGTGCCCCGAGGGCA
TCATCCTCTTCCTCTTCCAGTCCATCCTGGGCTCCATCGTGGACGCCTTCCTCATCGGCTG
CATGTTCATCAAGATGTCCCAGCCCAAGAAGCGCGCCGAGACCCTCATGTTCAGCGAGCAC
GCGGTGATCTCCATGAGGGACGGAAAACTCACGCTTATGTTCCGGGTGGGCAACCTGCGCA
ACAGCCACATGGTCTCCGCGCAGATTCGCTGCAAGCTGCTCAAATCTCGGCAGACACCTGA
GGGTGAGTTCCTTCCCCTTGACCAACTTGAACTGGATGTAGGTTTTAGTACAGGGGCAGAT
CAACTTTTTCTTGTGTCCCCCCTCACAATTTGCCACGTGATCGATGCCAAAAGCCCCTTTT
ATGACCTATCCCAGCGAAGCATGCAAACTGAACAGTTCGAGATTGTCGTCATCCTAGAAGG
CATTGTGGAAACAACTGGGATGACTTGTCAAGCTCGAACATCATATACTGAAGATGAAGTT
CTTTGGGGTCATCGTTTTTTTCCTGTAATTTCCTTAGAAGAGGGATTCTTTAAAGTTGATT
ACTCCCAGTTCCACGCAACATTTGAAGTCCCCACCCCACCTTACAGTGTGAAAGAGCAGGA
GGAAATGCTTCTCATGTCGTCCCCTTTAATAGCACCAGCCATAACTAACAGCAAAGAAAGA
CATAATTCTGTGGAATGCTTAGATGGACTAGATGATATTACTACAAAACTACCATCTAAGC
TGCAGAAAATTACTGGAAGAGAAGACTTTCCCAAAAAACTCTTGAGGATGAGTTCTACAAC
TTCAGAAAAAGCCTACAGCTTGGGAGACTTGCCCATGAAACTTCAACGAATAAGTTCAGTT
CCGGGCAACTCAGAAGAAAAACTGGTATCTAAAACCACCAAGATGTTATCTGATCCCATGA
GCCAGTCTGTGGCTGATTTGCCACCAAAGCTTCAAAAGATGGCTGGAGGAGCAGCTAGGAT
GGAAGGGAACCTTCCAGCCAAATTAAGAAAAATGAACTCTGATCGCTTCACATAACAAAGC
ACTCCCTTAGGCATTATTTAATGTTTGATTTAGTAATAGTCCAATATTTGGCGATGAGGTA
ATTCTCCCTAAGGAATCTGAAAGTATATTTTCCTCCCAGTTCTACAGCATATTTCGAGAAC
CCTTCCTTTCCCAAGTATTGCGAATGTGCAGAAAGCAACAGTTACGGAGGGAGGACATCAT
AAGGAAGTTATTAACGGGCATGTATTATCACATCAAGCATGCAATAATGTGCAAATTTTGC
ATTTAGTTTTATGGCATGATTTATATATGGCATATTTATATTGTATATTCTGGAAAAAAAA
ATATATATATATATTTAAAGGGGAGATACTCTCCCTGACATTTCTAACATATGTATTAAGC

TATGTATACACACATATACATATATATATACACATACATACACATACATACATACATACAT

ATATATCTGATAAAATTGTGATGTTTTGTTCAAAGTTGTAGTTCTTGTGCATGTTTACTTT

ATTAGAGTAGGAAGGCTACTGGCATTAATTATTAATACCAAATATTTTAGCCTTAAATTTT

TGTCATTTTAAAATCTGATTTAATGTTTTCTGCTGTTTAAGGTCTTGGGAGGCTTTCAATT

GTATTTTATATGAGAGAATCACACAAGTTTGTGCTATCTATGGCCCTGCAAAAATATAACC

ATTACATGTTTAAATTGTAAATTTTAGAGCATACCAGTACTCAGTATAGCATTGAACATTT

CTTATGATTTTTAAAAGTTGCTAGTACTGGGGAGAAATAATTGTTGATTAATTTGAGAATT

ATTCCTTTCCTAGACTAATTAAAATCTGGAAATCTGTTTTGTATATGATCTAATACAAAGA

TGAGCTCTGAACAAACACTGAATCATGTTAATAGACAGTAGCCAAGTTATATTGAATATAT

CAGAATCTGTGTGAAGTTACACAATTAATTGTCCCTGTTTCAAACTGAGTAAATTGGAAAC

ATTTTCTTTCTTTTTCTGGAAATTTTGTCCATTTTAAAAACCAATCATTTTAAGAAGACAT

GACAATGCAATGAAACAGATGATAAATATTTATGCTTAAAATAAAAAAAAAAAAA

FIGURE 51

GAGAGGGTCCTTCAGGGTCTGCTTATGCCCTTGTTCAAGAACACCAGTGTCAGCTCTCTGT
ACTCTGGTTGCAGACTGACCTTGCTCAGGCCTGAGAAGGATGGGGCAGCCACCAGAGTGGA
TGCTGTCTGCACCCATCGTCCTGACCCCAAAAGCCCTGGACTGGACAGAGAGCGGCTGTAC
TGGAAGCTGAGCCAGCTGACCCACGGCATCACTGAGCTGGGCCCCTACACCCTGGACAGGC
ACAGTCTCTATGTCAATGGTTTCACCCATCAGAGCTCTATGACGACCACCAGAACTCCTGA
TACCTCCACAATGCACCTGGCAACCTCGAGAACTCCAGCCTCCCTGTCTGGACCTACGACC
GCCAGCCCCTCTCCTGGTGCTATTCACAATTAACTTCACCATCACTAACCTGCGGTATGAGG
AGAACATGCATCACCCTGGCTCTAGAAAGTTTAACACCACGGAGAGAGTCCTTCAGGGTCT
GCTCAGGCCTGTGTTCAAGAACACCAGTGTTGGCCCTCTGTACTCTGGCTGCAGACTGACC
TTGCTCAGGCCCAAGAAGGATGGGGCAGCCACCAAAGTGGATGCCATCTGCACCTACCGCC
CTGATCCCAAAAGCCCTGGACTGGACAGAGAGCAGCTATACTGGGAGCTGAGCCAGCTAAC
CCACAGCATCACTGAGCTGGGCCCCTACACCCTGGACAGGGACAGTCTCTATGTCAATGGT
TTCACACAGCGGAGCTCTGTGCCCACCACTAGCATTCCTGGGACCCCCACAGTGGACCTGG
GAACATCTGGGACTCCAGTTTCTAAACCTGGTCCCTCGGCTGCCAGCCCTCTCCTGGTGCT
ATTCACTCTCAACTTCACCATCACCAACCTGCGGTATGAGGAGAACATGCAGCACCCTGGC
TCCAGGAAGTTCAACACCACGGAGAGGGTCCTTCAGGGCCTGCTCAGGTCCCTGTTCAAGA
GCACCAGTGTTGGCCCTCTGTACTCTGGCTGCAGACTGACTTTGCTCAGGCCTGAAAAGGA
TGGGACAGCCACTGGAGTGGATGCCATCTGCACCCACCACCCTGACCCCAAAAGCCCTAGG
CTGGACAGAGAGCAGCTGTATTGGGAGCTGAGCCAGCTGACCCACAATATCACTGAGCTGG
GCCACTATGCCCTGGACAACGACAGCCTCTTTGTCAATGGTTTCACTCATCGGAGCTCTGT
GTCCACCACCAGCACTCCTGGGACCCCCACAGTGTATCTGGGAGCATCTAAGACTCCAGCC
TCGATATTTGGCCCTTCAGCTGCCAGCCATCTCCTGATACTATTCACCGTCAACTTCACCA
TCACTAACCTGCGGTATGAGGAGAACATGTGGCCTGGCTCCAGGAAGTTCAACACTACAGA
GAGGGTCCTTCAGGGCCTGCTAAGGCCCTTGTTCAAGAACACCAGTGTTGGCCCTCTGTAC
TCTGGCTCCAGGCTGACCTTGCTCAGGCCAGAGAAAGATGGGGAAGCCACCGGAGTGGATG
CCATCTGCACCCACCGCCCTGACCCCACAGGCCCTGGGCTGGACAGAGAGCAGCTGTATTT
GGAGCTGAGCCAGCTGACCCACAGCATCACTGAGCTGGGCCCCTACACACTGGACAGGGAC
AGTCTCTATGTCAATGGTTTCACCCATCGGAGCTCTGTACCCACCACCAGCACCGGGGTGG
TCAGCGAGGAGCCATTCACACTGAACTTCACCATCAACAACCTGCGCTACATGGCGGACAT
GGGCCAACCCGGCTCCCTCAAGTTCAACATCACAGACAACGTCATGAAGCACCTGCTCAGT
CCTTTGTTCCAGAGGAGCAGCCTGGGTGCACGGTACACAGGCTGCAGGGTCATCGCACTAA
GGTCTGTGAAGAACGGTGCTGAGACACGGGTGGACCTCCTCTGCACCTACCTGCAGCCCCT
CAGCGGCCCAGGTCTGCCTATCAAGCAGGTGTTCCATGAGCTGAGCCAGCAGACCCATGGC
ATCACCCGGCTGGGCCCCTACTCTCTGGACAAAGACAGCCTCTACCTTAACGGTTACAATG
AACCTGGTCCTAGATGAGCCTCCCTACAACTCCCGAAGCCAGCCACCACATTCCTGCCTCCTCT

GTCAGAAGCCACAACAGCCATGGGGTACCACCTGAAGACCCTCACACTCAACTTCACCATC
TCCAATCTCCAGTATTCACCAGATATGGGCAAGGGCTCAGCTACATTCAACTCCACCGAGG
GGGTCCTTCAGCACCTGCTCAGACCCTTGTTCCAGAAGAGCAGCATGGGCCCCTTCTACTT
GGGTTGCCAACTGATCTCCCTCAGGCCTGAGAAGGATGGGGCAGCCACTGGTGTGGACACC
ACCTGCACCTACCACCCTGACCCTGTGGGCCCCGGGCTGGACATACAGCAGCTTTACTGGG
AGCTGAGTCAGCTGACCCATGGTGTCACCCAACTGGGCTTCTATGTCCTGGACAGGGATAG
CCTCTTCATCAATGGCTATGCACCCCAGAATTTATCAATCCGGGGCGAGTACCAGATAAAT
TTCCACATTGTCAACTGGAACCTCAGTAATCCAGACCCCACATCCTCAGAGTACATCACCC
TGCTGAGGGACATCCAGGACAAGGTCACCACACTCTACAAAGGCAGTCAACTACATGACAC
ATTCCGCTTCTGCCTGGTCACCAACTTGACGATGGACTCCGTGTTGGTCACTGTCAAGGCA
TTGTTCTCCTCCAATTTGGACCCCAGCCTGGTGGAGCAAGTCTTTCTAGATAAGACCCTGA
ATGCCTCATTCCATTGGCTGGGCTCCACCTACCAGTTGGTGGACATCCATGTGACAGAAAT
GGAGTCATCAGTTTATCAACCAACAAGCAGCTCCAGCACCCAGCACTTCTACCCGAATTTC
ACCATCACCAACCTACCATATTCCCAGGACAAAGCCCAGCCAGGCACCACCAATTACCAGA
GGAACAAAAGGAATATTGAGGATGCGCTCAACCAACTCTTCCGAAACAGCAGCATCAAGAG
TTATTTTTCTGACTGTCAAGTTTCAACATTCAGGTCTGTCCCCAACAGGCACCACACCGGG
GTGGACTCCCTGTGTAACTTCTCGCCACTGGCTCGGAGAGTAGACAGAGTTGCCATCTATG
AGGAATTTCTGCGGATGACCCGGAATGGTACCCAGCTGCAGAACTTCACCCTGGACAGGAG
CAGTGTCCTTGTGGATGGGTATTCTCCCAACAGAAATGAGCCCTTAACTGGGAATTCTGAC
CTTCCCTTCTGGGCTGTCATCTTCATCGGCTTGGCAGGACTCCTGGGACTCATCACATGCC
TGATCTGCGGTGTCCTGGTGACCACCCGCCGGCGGAAGAAGGAAGGAGAATACAACGTCCA
GCAACAGTGCCCAGGCTACTACCAGTCACACCTAGACCTGGAGGATCTGCAATGAACTGGAA
CTTGCCGGTGCCTGGGGTGCCTTTCCCCCAGCCAGGGTCCAAAGAAGCTTGGCTGGGGCAG
AAATAAACCATATTGGTCG

FIGURE 52

GCCCGGGCAGGTGCTGATAGCACAGTTCTGTCCAGAGAAGGAAGGCGGAATAAACTTATTC
ATTCCCAGGAACTCTTGGGGTAGGTGTGTGTTTTTCACATCTTAAAGGCTCACAGACCCTG
CGCTGGACAAATGTTCCATTCCTGAAGGACCTCTCCAGAATCCGGATTGCTGAATCTTCCC
TGTTGCCTAGAAGGGCTCCAAACCACCTCTTGACAATGGGAAACTGGGTGGTTAACCACTG
GTTTTCAGTTTTGTTTCTGGTTGTTTGGTTAGGGCTGAATGTTTTCCTGTTTGTGGATGCC
TTCCTGAAATATGAGAAGGCCGACAAATACTACTACACAAGAAAAATCCTTGGGTCAACAT
TGGCCTGTGCCCGAGCGTCTGCTCTCTGCTTGAATTTTAACAGCACGCTGATCCTGCTTCC
TGTGTGTCGCAATCTGCTGTCCTTCCTGAGGGGCACCTGCTCATTTTGCAGCCGCACACTG
AGAAAGCAATTGGATCACAACCTCACCTTCCACAAGCTGGTGGCCTATATGATCTGCCTAC
ATACAGCTATTCACATCATTGCACACCTGTTTAACTTTGACTGCTATAGCAGAAGCCGACA
GGCCACAGATGGCTCCCTTGCCTCCATTCTCTCCAGCCTATCTCATGATGAGAAAAAGGGG
GGTTCTTGGCTAAATCCCATCCAGTCCCGAAACACGACAGTGGAGTATGTGACATTCACCA
GCGTTGCTGGTCTCACTGGAGTGATCATGACAATAGCCTTGATTCTCATGGTAACTTCAGC
TACTGAGTTCATCCGGAGGAGTTATTTTGAAGTCTTCTGGTATACTCACCACCTTTTTATC
TTCTATATCCTTGGCTTAGGGATTCACGGCATTGGTGGAATTGTCCGGGGTCAAACAGAGG
AGAGCATGAATGAGAGTCATCCTCGCAAGTGTGCAGAGTCTTTTGAGATGTGGGATGATCG
TGACTCCCACTGTAGGCGCCCTAAGTTTGAAGGGCATCCCCCTGAGTCTTGGAAGTGGATC
CTTGCACCGGTCATTCTTTATATCTGTGAAAGGATCCTCCGGTTTTACCGCTCCCAGCAGA
AGGTTGTGATTACCAAGGTTGTTATGCACCCATCCAAAGTTTTGGAATTGCAGATGAACAA
GCGTGGCTTCAGCATGGAAGTGGGGCAGTATATCTTTGTTAATTGCCCCTCAATCTCTCTC
CTGGAATGGCATCCTTTTACTTTGACCTCTGCTCCAGAGGAAGATTTCTTCTCCATTCATA
TCCGAGCAGCAGGGGACTGGACAGAAAATCTCATAAGGGCTTTCGAACAACAATATTCACC
AATTCCCAGGATTGAAGTGGATGGTCCCTTTGGCACAGCCAGTGAGGATGTTTTCCAGTAT
GAAGTGGCTGTGCTGGTTGGAGCAGGAATTGGGGTCACCCCCTTTGCTTCTATCTTGAAAT
CCATCTGGTACAAATTCCAGTGTGCAGACCACAACCTCAAAACAAAAAAGATCTATTTCTA
CTGGATCTGCAGGGAGACAGGTGCCTTTTCCTGGTTCAACAACCTGTTGACTTCCCTGGAA
CAGGAGATGGAGGAATTAGGCAAAGTGGGTTTTCTAAACTACCGTCTCTTCCTCACCGGAT
GGGACAGCAATATTGTTGGTCATGCAGCATTAAACTTTGACAAGGCCACTGACATCGTGAC
AGGTCTGAAACAGAAAACCTCCTTTGGGAGACCAATGTGGGACAATGAGTTTTCTACAATA
GCTACCTCCCACCCCAAGTCTGTAGTGGGAGTTTTCTTATGTGGCCCTCGGACTTTGGCAA
AGAGCCTGCGCAAATGCTGTCACCGATATTCCAGTCTGGATCCTAGAAAGGTTCAATTCTA
CTTCAACAAAGAAAATTTTTGAGTTATAGGAATAAGGACGGTAATCTGCATTTTGTCTCTT
TGTATCTTCAGTAATTGAGTTATAGGAATAAGGACGGTAATCTGCATTTTGTCTCTTTGTA
TCTTCAGTAATTTACTTGGTCTCGTCAGGTTTGAGCAGTCACTTTAGGATAAGAATGTGCC

TCAGCAACTTAAGAACTTCTGAAGTTCTTAAAGTTCTGAAGTTCTTAAAGCCCATGGATCC
TTTCTCAGAAAAATAACTGTAAATCTTTCTGGACAGCCATGACTGTAGCAAGGCTTGATAG
CAGAGGTTTGGTGGTTCAGAGTTATACAACTAATCCCAGGTGATTTTATCAATTCCAGTGT
TACCATCTCCTGAGTTTTGGTTTGTAATCTTTTGTCCCTCCCACCCCCACAGAAGATTTCT
AAGTAGGGTGACTTTTTTAAATAAAAATTTATTGAATAATTAATGATAAAACATAATAATAA
ACATAAATAATAAACAAAATTACCGAGAACCCCATCCCCATATAACACCAACAGTGTACAT
GTTTACTGTCACTTTTGATATGGTCTTATCCAGTGTGAACAGCAATTTATTATTTTTGCTC
ATCAAAAAATAAAGGATTTTCTTCTTCACTTGATGAATGAGTCTTTTGTTACTTCTTTTTG
GGCTCTCCATTTGGGAAAGAAAATCTAGAAAGTATGCTCATTGTAATTGAAATTGTATTAT
AATTGCAAGTCAGCTATACTGTCATCTTGTATTATTCTGTTAAAAGCGATTGTGAAAACTT
GAACTCCATGTTAGGATGACTTATTTAGTCAAGAGGAGCTCTTTATCACTATTAAGGGAGT
TTATTAATGTCTAGTTTAAAAAATAACAATTTATAATTTTCATAGATTTTTGTGGACTCCT
ACCTATATCGCTAAGGAATCATTCAGCTAATAACATAAGCAAAATGCCTTCATTCTCTATC
AGAATGCTTTTTATTTCTGCCATCTACCCAACCTCGGCCTACAAGGACCACACTGGCTCTG
GGATGTTAGTCGTCACAGTGAATCCCTAGAAGGAGAGAAGCCCAGGGCTCTTAAGAGAGAA
ACCTCACACTTCATTGTGCCCAGACAATTCAACTCATTATAGATTAGAAGAATGGTTGTAC
TGAATTAGCATTATATTATCTATTTTTTCTACCATTAAGTGACCACGAATGAACAAATGGT
TGTTTCTAATAGTTTAGGAACCTCTACGGCTTGAGGGATAAAACTGTTTTTGGTTTTTTTT
TGTCACTAGTCTGAGCACAATAAAGGTAGGTCATGAACTCATGGTGGCTAATCATGTTTCA
TTTATTACAACAGTAAAACAATACTGATGGTTGTTTACATAGGGAGGGATACCACTGCTGT
TGCCAATTTGTTCACCCGGATTTCAGCTAATTCAGGTTTAATAAGGCTTTGGGAAAATGCA
CAATGAAACCTTGAAGACATGATTCTTATTAGACAACTACTGACCAAATGACAACTTCATC
TTCAATGACCTTGATTTGGAATCTTGAAACTTACAACGAGTTCCTAATGAGAAAGGCATCA
CTCTTGGTCTCCATGTCTAAATATATTGACACACTCACTTCAAGATATGATGGCTGCCAAA
ATCAATATCCCAAGGCTTGAATATTTATGGGAGTAGGCATAAATTCGTGATTATATTCATG
GTAATGAATATAATAAATCCTATGCATAGCCAACATCCAAGTTGAACTGGGGATGAGCAGG
ATACACAGAAAGCTCCTCACCTGCTTTGAGCAAGTTTTTCTTGCAAAGTTGTAATATGAAA
TACCTTAATCAACAAATTTCTTATAAATGTTTATATGAGTAGAAAATCCCTCAATTTCATT
GTCACTCAAAAAGTTGATTATCT

FIGURE 53

CCCTTGGGCTCCAAACCACCTCTTGACAATGGGAAACTGGGTGGTTAACCACTGGTTTTCA
GTTTTGTTTCTGGTTGTTTGGTTAGGGCTGAATGTTTTCCTGTTTGTGGATGCCTTCCTGA
AATATGAGAAGGCCGACAAATACTACTACACAAGAAAAATCCTTGGGTCAACATTGGCCTG
TGCCCGAGCGTCTGCTGTCTGCTTGAATTTTAACAGCACGCTGATCCTGCTTCCTGTGTGT
CGCAATCTGCTGTCCTTCCTGAGGGGCACCTGCTCATTTTGCAGCCGCACACTGAGAAAGC
AATTGGATCACAACCTCACCTTCCACAAGCTGGTGGCCTATATGATCTGCCTACATACAGC
TATTCACATCATTGCACACCTGTTTAACTTTGACTGCTATAGCAGAAGCCGACAGGCCACA
GATGGCTCCCTTGCCTCCATTCTCTCCAGCCTATCTCATGATGAGAAAAAGGGGGGTTCTT
GGCTAAATCCCATCCAGTCCCGAAACACGACAGTGGAGTATGTGACATTCACCAGCATTGC
TGGTCTCACTGGAGTGATCATGACAATAGCCTTGATTCTCATGGTAACTTCAGCTACTGAG
TTCATCCGGAGGAGTTATTTTGAAGTCTTCTGGTATACTCACCACCTTTTTATCTTCTATA
TCCTTGGCTTAGGGATTCACGGCATTGGTGGAATTGTCCGGGGTCAAACAGAGGAGAGCAT
GAATGAGAGTCATCCTCGCAAGTGTGCAGAGTCTTTTGAGATGTGGGATGATCGTGACTCC
CACTGTAGGCGCCCTAAGTTTGAAGGGCATCCCCCTGAGTCTTGGAAGTGGATCCTTGCAC
CGGTCATTCTTTATATCTGTGAAAGGATCCTCCGGTTTTACCGCTCCCAGCAGAAGGTTGT
GATTACCAAGGTTGTTATGCACCCATCCAAAGTTTTGGAATTGCAGATGAACAAGCGTGGC
TTCAGCATGGAAGTGGGGCAGTATATCTTTGTTAATTGCCCCTCAATCTCTCTCCTGGAAT
GGCATCCTTTTACTTTGACCTCTGCTCCAGAGGAAGATTTCTTCTCCATTCATATCCGAGC
AGCAGGGGACTGGACAGAAAATCTCATAAGGGCTTTCGAACAACAATATTCACCAATTCCC
AGGATTGAAGTGGATGGTCCCTTTGGCACAGCCAGTGAGGATGTTTTCCAGTATGAAGTGG
CTGTGCTGGTTGGAGCAGGAATTGGGGTCACCCCCTTTGCTTCTATCTTGAAATCCATCTG
GTACAAATTCCAGTGTGCAGACCACAACCTCAAAACAAAAAAGATCTATTTCTACTGGATC
TGCAGGGAGACAGGTGCCTTTTCCTGGTTCAACAACCTGTTGACTTCCCTGGAACAGGAGA
TGGAGGAATTAGGCAAAGTGGGTTTTCTAAACTACCGTCTCTTCCTCACCGGATGGGACAG
CAATATTGTTGGTCATGCAGCATTAAACTTTGACAAGGCCACTGACATCGTGACAGGTCTG
AAACAGAAAACCTCCTTTGGGAGACCAATGTGGGACAATGAGTTTTCTACAATAGCTACCT
CCCACCCCAAGTCTGTAGTGGGAGTTTTCTTATGTGGCCCTCGGACTTTGGCAAAGAGCCT
GCGCAAATGCTGTCACCGATATTCCAGTCTGGATCCTAGAAAGGTTCAATTCTACTTCAAC
AAAGAAAATTTTTGAGTTATAGGAATAAGGACGGTAATCTGCATTTTGTCTCTTTGTATCT
TCAGTAATTTACTTGGTCTCGTCAGGTTTGAGCAGTCACTTTAGGAAG

FIGURE 54

GGGCTCCAAACCACCTCTTGACAATGGGAAACTGGGTGGTTAACCACTGGTTTTCAGTTTT
GTTTCTGGTTGTTTGGTTAGGGCTGAATGTTTTCCTGTTTGTGGATGCCTTCCTGAAATAT
GAGAAGGCCGACAAATACTACTACACAAGAAAAATCCTTGGGTCTTGGAAGTGGATCCTTG
CACCGGTCATTCTTTATATCTGTGAAAGGATCCTCCGGTTTTACCGCTCCCAGCAGAAGGT
TGTGATTACCAAGGTTGTTATGCACCCATCCAAAGTTTTGGAATTGCAGATGAACAAGCGT
GGCTTCAGCATGGAAGTGGGGCAGTATATCTTTGTTAATTGCCCCTCAATCTCTCTCCTGG
AATGGCATCCTTTTACTTTGACCTCTGCTCCAGAGGAAGATTTCTTCTCCATTCATATCCG
AGCAGCAGGGGACTGGACAGAAAATCTCATAAGGGCTTTCGAACAACAATATTCACCAATT
CCCAGGATTGAAGTGGATGGTCCCTTTGGCACAGCCAGTGAGGATGTTTTCCAGTATGAAG
TGGCTGTGCTGGTTGGAGCAGGAATTGGGGTCACCCCCTTTGCTTCTATCTTGAAATCCAT
CTGGTACAAATTCCAGTGTGCAGACCACAACCTCAAAACAAAAAAGGTTGGTCATGCAGCA
TTAAACTTTGACAAGGCCACTGACATCGTGACAGGTCTGAAACAGAAAACCTCCTTTGGGA
GACCAATGTGGGACAATGAGTTTTCTACAATAGCTACCTCCCACCCCAAGTCTGTAGTGGG
AGTTTTCTTATGTGGCCCTCGGACTTTGGCAAAGAGCCTGCGCAAATGCTGTCACCGATAT
TCCAGTCTGGATCCTAGAAAGGTTCAATTCTACTTCAACAAAGAAAATTTTTGAGTTATAG
GAATAAGGACGGTAATCTGCATTTTGTCTCTTTGTATCTTCAGTAATTTACTTGGTCTCGT
CAGGTTTGAGCAGTCACTTTAGGA

FIGURE 55

GGGCTCCAAACCACCTCTTGACAATGGGAAACTGGGTGGTTAACCACTGGTTTTCAGTTTT
GTTTCTGGTTGTTTGGTTAGGGCTGAATGTTTTCCTGTTTGTGGATGCCTTCCTGAAATAT
GAGAAGGCCGACAAATACTACTACACAAGAAAAATCCTTGGGTCAACATTGGCCTGTGCCC
GAGCGTCTGCTCTCTGCTTGAATTTTAACAGCACGCTGATCCTGCTTCCTGTGTGTCGCAA
TCTGCTGTCCTTCCTGAGGGGCACCTGCTCATTTTGCAGCCGCACACTGAGAAAGCAATTG
GATCACAACCTCACCTTCCACAAGCTGGTGGCCTATATGATCTGCCTACATACAGCTATTC
ACATCATTGCACACCTGTTTAACTTTGACTGCTATAGCAGAAGCCGACAGGCCACAGATGG
CTCCCTTGCCTCCATTCTCTCCAGCCTATCTCATGATGAGAAAAAGGGGGGTTCTTGGCTA
AATCCCATCCAGTCCCGAAACACGACAGTGGAGTATGTGACATTCACCAGCATTGCTGGTC
TCACTGGAGTGATCATGACAATAGCCTTGATTCTCATGGTAACTTCAGCTACTGAGTTCAT
CCGGAGGAGTTATTTTGAAGTCTTCTGGTATACTCACCACCTTTTTATCTTCTATATCCTT
GGCTTAGGGATTCACGGCATTGGTGGAATTGTCCGGGGTCAAACAGAGGAGAGCATGAATG
AGAGTCATCCTCGCAAGTGTGCAGAGTCTTTTGAGATGTGGGATGATCGTGACTCCCACTG
TAGGCGCCCTAAGTTTGAAGGGCATCCCCCTGAGTCTTGGAAGTGGATCCTTGCACCGGTC
ATTCTTTATATCTGTGAAAGGATCCTCCGGTTTTACCGCTCCCAGCAGAAGGTTGTGATTA
CCAAGGTTGTTATGCACCCATCCAAAGTTTTGGAATTGCAGATGAACAAGCGTGGCTTCAG
CATGGAAGTGGGGCAGTATATCTTTGTTAATTGCCCCTCAATCTCTCTCCTGGAATGGCAT
CCTTTTACTTTGACCTCTGCTCCAGAGGAAGATTTCTTCTCCATTCATATCCGAGCAGCAG
GGGACTGGACAGAAAATCTCATAAGGGCTTTCGAACAACAATATTCACCAATTCCCAGGAT
TGAAGTGGATGGTCCCTTTGGCACAGCCAGTGAGGATGTTTTCCAGTATGAAGTGGCTGTG
CTGGTTGGAGCAGGAATTGGGGTCACCCCCTTTGCTTCTATCTTGAAATCCATCTGGTACA
AATTCCAGTGTGCAGACCACAACCTCAAAACAAAAAAGGTTGGTCATGCAGCATTAAACTT
TGACAAGGCCACTGACATCGTGACAGGTCTGAAACAGAAAACCTCCTTTGGGAGACCAATG
TGGGACAATGAGTTTTCTACAATAGCTACCTCCCACCCCAAGTCTGTAGTGGGAGTTTTCT
TATGTGGCCCTCGGACTTTGGCAAAGAGCCTGCGCAAATGCTGTCACCGATATTCCAGTCT
GGATCCTAGAAAGGTTCAATTCTACTTCAACAAAGAAAATTTTTGAGTTATAGGAATAAGG
ACGGTAATCTGCATTTTGTCTCTTTGTATCTTCAGTAATTTACTTGGTCTCGTCAGGTTTG
AGCAGTCACTTTAGGAAG

FIGURE 56

GGACTGCTTTGTAACTGCTAAGATTGCAGACAGAAATAGCACACAACCACTGTGAGCTGTA
TGCGATTCAGAAACCAAGACCAAATTTTGCTCACTTTCATTAATCAGTTGCTCAGATAGAA
GGAAATGACATCTGGTTCTGTCTTCTTCTACATCTTAATTTTTGGAAAATATTTTTCTCAT
GGGGGTGGACAGGATGTCAAGTGCTCCCTTGGCTATTTCCCCTGTGGGAACATCACAAAGT
GCTTGCCTCAGCTCCTGCACTGTAACGGTGTGGACGACTGCGGGAATCAGGCCGATGAGGA
CAACTGTGGAGACAACAATGGATGGTCCATGCAATTTGACAAATATTTTGCCAGTTACTAC
AAAATGACTTCCCAATATCCTTTTGAGGCAGAAACACCTGAATGTTTGGTCGGTTCTGTGC
CAGTGCAATGTCTTTGCCAAGGTCTGGAGCTTGACTGTGATGAAACCAATTTACGAGCTGT
TCCATCGGTTTCTTCAAATGTGACTGCAATGTCACTTCAGTGGAACTTAATAAGAAAGCTT
CCTCCTGATTGCTTCAAGAATTATCATGATCTTCAGAAGCTGTACCTGCAAAACAATAAGA
TTACATCCATCTCCATCTATGCTTTCAGAGGACTGAATAGCCTTACTAAACTGTATCTCAG
TCATAACAGAATAACCTTCCTGAAGCCGGGTGTTTTTGAAGATCTTCACAGACTAGAATGG
CTGATAATTGAAGATAATCACCTCAGTCGAATTTCCCCACCAACATTTTATGGACTAAATT
CTCTTATTCTCTTAGTCCTGATGAATAACGTCCTCACCCGTTTACCTGATAAACCTCTCTG
TCAACACATGCCAAGACTACATTGGCTGGACCTTGAAGGCAACCATATCCATAATTTAAGA
AATTTGACTTTTATTTCCTGCAGTAATTTAACTGTTTTAGTGATGAGGAAAAACAAAATTA
ATCACTTAAATGAAAATACTTTTGCACCTCTCCAGAAACTGGATGAATTGGATTTAGGAAG
TAATAAGATTGAAAATCTTCCACCGCTTATATTCAAGGACCTGAAGGAGCTGTCACAATTG
AATCTTTCCTATAATCCAATCCAGAAAATTCAAGCAAACCAATTTGATTATCTTGTCAAAC
TCAAGTCTCTCAGCCTAGAAGGGATTGAAATTTCAAATATCCAACAAAGGATGTTTAGACC
TCTTATGAATCTCTCTCACATATATTTTAAGAAATTCCAGTACTGTGGGTATGCACCACAT
GTTCGCAGCTGTAAACCAAACACTGATGGAATTTCATCTCTAGAGAATCTCTTGGCAAGCA
TTATTCAGAGAGTATTTGTCTGGGTTGTATCTGCAGTTACCTGCTTTGGAAACATTTTTGT
CATTTGCATGCGACCTTATATCAGGTCTGAGAACAAGCTGTATGCCATGTCAATCATTTCT
CTCTGCTGTGCCGACTGCTTAATGGGAATATATTTATTCGTGATCGGAGGCTTTGACCTAA
AGTTTCGTGGAGAATACAATAAGCATGCGCAGCTGTGGATGGAGAGTACTCATTGTCAGCT
TGTAGGATCTTTGGCCATTCTGTCCACAGAAGTATCAGTTTTACTGTTAACATTTCTGACA
TTGGAAAAATACATCTGCATTGTCTATCCTTTTAGATGTGTGAGACCTGGAAAATGCAGAA
CAATTACAGTTCTGATTCTCATTTGGATTACTGGTTTTATAGTGGCTTTCATTCCATTGAG
CAATAAGGAATTTTTCAAAAACTACTATGGCACCAATGGAGTATGCTTCCCTCTTCATTCA
GAAGATACAGAAAGTATTGGAGCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATT
TGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAG
TGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGT
TTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTC

GCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAA
ATGATTCATCGGTTTTGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTCAGA
AAACATATGCTCCATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGA
GTTAATGAAGCCGGACCTTTTCACATACCCCTGTGAAATGTCACTGATTTCTCAATCAACG
AGACTCAATTCCTATTCATGA

FIGURE 57

TAGAGATCCCTCGACCTCGACCCACGCGTCCGAGGAAAGAAAAAAAGAGGAATGGAAAGAG
ACAGAGAAAGGAAATGGGAGTGGAAGGAGGGAGGACTGCTTTGTAACTGCTAAGATTGCAG
ACAGAAATAGCACACAACCACTGTGAGCTGTATGCGATTCAGAAACCAAGACCAAATTTTG
CTCACTTTCATTAATCAGTTGCTCAGATAGAAGGAAATGACATCTGGTTCTGTCTTCTTCT
ACATCTTAATTTTTGGAAAATATTTTTCTCATGGGGGTGGACAGGATGTCAAGTGCTCCCT
TGGCTATTTCCCCTGTGGGAACATCACAAAGTGCTTGCCTCAGCTCCTGCACTGTAACGGT
GTGGACGACTGCGGGAATCAGGCCGATGAGGACAACTGTGGAGACAACAATGGATGGTCTC
TGCAATTTGACAAATATTTTGCCAGTTACTACAAAATGACTTCCCAATATCCTTTTGAGGC
AGAAACACCTGAATGTTTGGTCGGTTCTGTGCCAGTGCAATGTCTTTGCCAAGGTCTGGAG
CTTGACTGTGATGAAACCAATTTACGAGCTGTTCCATCGGTTTCTTCAAATGTGACTGCAA
TGTCACTTCAGTGGAACTTAATAAGAAAGCTTCCTCCTGATTGCTTCAAGAATTATCATGA
TCTTCAGAAGCTGTACCTGCAAAACAATAAGATTACATCCATCTCCATCTATGCTTTCAGA
GGACTGAATAGCCTTACTAAACTGTATCTCAGTCATAACAGAATAACCTTCCTGAAGCCGG
GTGTTTTTTGAAGATCTTCACAGACTAGAATGGCTGATAATTGAAGATAATCACCTCAGTCG
AATTTCCCCACCAACATTTTATGGACTAAATTCTCTTATTCTCTTAGTCCTGATGAATAAC
GTCCTCACCCGTTTACCTGATAAACCTCTCTGTCAACACATGCCAAGACTACATTGGCTGG
ACCTTGAAGGCAACCATATCCATAATTTAAGAAATTTGACTTTTATTTCCTGCAGTAATTT
AACTGTTTTAGTGATGAGGAAAAACAAAATTAATCACTTAAATGAAAATACTTTTGCACCT
CTCCAGAAACTGGATGAATTGGATTTAGGAAGTAATAAGATTGAAAATCTTCCACCGCTTA
TATTCAAGGACCTGAAGGAGCTGTCACAATTGAATCTTTCCTATAATCCAATCCAGAAAAT
TCAAGCAAACCAATTTGATTATCTTGTCAAACTCAAGTCTCTCAGCCTAGAAGGGATTGAA
ATTTCAAATATCCAACAAAGGATGTTTAGACCTCTTATGAATCTCTCTCACATATATTTTA
AGAAATTCCAGTACTGTGGGTATGCACCACATGTTCGCAGCTGTAAACCAAACACTGATGG
AATTTCATCTCTAGAGAATCTCTTGGCAAGCATTATTCAGAGAGTATTTGTCTGGGTTGTA
TCTGCAGTTACCTGCTTTGGAAACATTTTTGTCATTTGCATGCGACCTTATATCAGGTCTG
AGAACAAGCTGTATGCCATGTCAATCATTTCTCTCTGCTGTGCCGACTGCTTAATGGGAAT
ATATTTATTCGTGATCGGAGGCTTTGACCTAAAGTTTCGTGGAGAATACAATAAGCATGCG
CAGCTGTGGATGGAGAGTACTCATTGTCAGCTTGTAGGATCTTTGGCCATTCTGTCCACAG
AAGTATCAGTTTTACTGTTAACATTTCTGACATTGGAAAAATACATCTGCATTGTCTATCC
TTTTAGATGTGTGAGACCTGGAAAATGCAGAACAATTACAGTTCTGATTCTCATTTGGATT
ACTGGTTTTATAGTGGCTTTCATTCCATTGAGCAATAAGGAATTTTTCAAAAACTACTATG
GCACCAATGGAGTATGCTTCCCTCTTCATTCAGAAGATACAGAAAGTATTGGAGCCCAGAT
TTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCC
TATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATC
AAGTTAAAAAAGAGATGATCCTTTGCCAAACGTTTTTTCTTTTATAGTATTTACTGATGCATT

ATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGT
ACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTC
TCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGGTATAACTACAG
ACAAAGAAAATCTATGGACAGCAAAGGTCAGAAAACATATGCTCCATCATTCATCTGGGTG
GAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAATGAAGCCGGACCTTTTCACATACC
CCTGTGAAATGTCACTGATTTCTCAATCAACGAGACTCAATTCCTATTCATGACTGACTCT
GAAATTCATTTCTTCGCAGAGAATACTGTGGGGGTGCTTCATGAGGGATTTACTGGTATGA
AATGAATACCACAAAATTAATTTATAATAATAGCTAAGATAAATATTTTACAAGGACATGA
GGAAAAATAAAAATGACTAATGCTCTTACAAAGGGAAGTAATTATATCAATAATGTATATA
TATTAGTAGACATTTTGCATAAGAAATTAAGAGAAATCTACTTCAGTAACATTCATTCATT
TTTCTAACATGCATTTATTGAGTACCCACTACTATGTGCATAGCATTGCAATATAGTCCTG
GAAGTAGACAGTGCAGAACCTTTCAATCTGTAGATGGTGTTTAATGACAAAAGACTATACA
AAGTCCATCTGCAGTTCCTAGTTTAAAGTAGAGCTTTACCTGTCATGTGCATCAGCAAGAA
TCATAGGCACTTTTAAATAAAGGTTTAAAGTTTTGGAAT

FIGURE 58

GATCCATTGTGCTGGCGCGGATTCTTTATCACTGGAAGGAGGAGGAGGGGGAGAGTCGCTC
CCGCCGGGCGAGCATGGGGCGCCTGGCCTCGAGGCCGCTGCTGCTGGCGCTCCTGTCGTTG
GCTCTTTGCCGAGGGCGTGTGGTGAGAGTCCCCACAGCGACCCTGGTTCGAGTGGTGGGCA
CTGAGCTGGTCATCCCCTGCAACGTCAGTGACTATGATGGCCCCAGCGAGCAAAACTTTGA
CTGGAGCTTCTCATCTTTGGGGAGCAGCTTTGTGGAGCTTGCAAGCACCTGGGAGGTGGGG
TTCCCAGCCCAGCTGTACCAGGAGCGGCTGCAGAGGGGCGAGATCCTGTTAAGGCGGACTG
CCAACGACGCCGTGGAGCTCCACATAAAGAACGTCCAGCCTTCAGACCAAGGCCACTACAA
ATGTTCAACCCCCAGCACAGATGCCACTGTCCAGGGAAACTATGAGGACACAGTGCAGGTT
AAAGTGCTGGCCGACTCCCTGCACGTGGGCCCCAGCGCGCGGCCCCCGCCGAGCCTGAGCC
TGCGGGAGGGGGAGCCCTTCGAGCTGCGCTGCACCGCCGCCTCCGCCTCGCCGCTGCACAC
GCACCTGGCGCTGCTGTGGGAGGTGCACCGCGGCCCGGCCAGGCGGAGCGTCCTCGCCCTG
ACCCACGAGGGCAGGTTCCACCCGGGCCTGGGGTACGAGCAGCGCTACCACAGTGGGGACG
TGCGCCTCGACACCGTGGGCAGCGACGCCTACCGCCTCTCAGTGTCCCGGGCTCTGTCTGC
CGACCAGGGCTCCTACAGGTGTATCGTCAGCGAGTGGATCGCCGAGCAGGGCAACTGGCAG
GAAATCCAAGAAAAGGCCGTGGAAGTTGCCACCGTGGTGATCCAGCCGACAGTTCTGCGAG
CAGCTGTGCCCAAGAATGTGTCTGTGGCTGAAGGAAAGGAACTGGACCTGACCTGTAACAT
CACAACAGACCGAGCCGATGACGTCCGGCCCGAGGTGACGTGGTCCTTCAGCAGGATGCCT
GACAGCACCCTACCTGGCTCCCGCGTGTTGGCGCGGCTTGACCGTGATTCCCTGGTGCACA
GCTCGCCTCATGTTGCTTTGAGTCATGTGGATGCACGCTCCTACCATTTACTGGTTCGGGA
TGTTAGCAAAGAAAACTCTGGCTACTATTACTGCCACGTGTCCCTGTGGGCACCCGGACAC
AACAGGAGCTGGCACAAAGTGGCAGAGGCCGTGTCTTCCCCAGCTGGTGTGGGTGTGACCT
GGCTAGAACCAGACTACCAGGTGTACCTGAATGCTTCCAAGGTCCCCGGGTTTGCGGATGA
CCCCACAGAGCTGGCATGCCGGGTGGTGGACACGAAGAGTGGGGAGGCGAATGTCCGATTC
ACGGTTTCGTGGTACTACAGGATGAACCGGCGCAGCGACAATGTGGTGACCAGCGAGCTGC
TTGCAGTCATGGACGGGGACTGGACGCTAAAATATGGAGAGAGGAGCAAGCAGCGGGCCCA
GGATGGAGACTTTATTTTTTCTAAGGAACATACAGACACGTTCAATTTCCGGATCCAAAGG
ACTACAGAGGAAGACAGAGGCAATTATTACTGTGTTGTGTCTGCCTGGACCAAACAGCGGA
ACAACAGCTGGGTGAAAAGCAAGGATGTCTTCTCCAAGCCTGTTAACATATTTTGGGCATT
AGAAGATTCCGTGCTTGTGGTGAAGGCGAGGCAGCCAAAGCCTTTCTTTGCTGCCGGAAAT
ACATTTGAGATGACTTGCAAAGTATCTTCCAAGAATATTAAGTCGCCACGCTACTCTGTTC
TCATCATGGCTGAGAAGCCTGTCGGCGACCTCTCCAGTCCCAATGAAACGAAGTACATCAT
CTCTCTGGACCAGGATTCTGTGGTGAAGCTGGAGAATTGGACAGATGCATCACGGGTGGAT
GGCGTTGTTTTAGAAAAAGTGCAGGAGGATGAGTTCCGCTATCGAATGTACCAGACTCAGG
TCTCAGACGCAGGGCTGTACCGCTGCATGGTGACAGCCTGGTCTCCTGTCAGGGGCAGCCT
TTGGCGAGAAGCAGCAACCAGTCTCTCCAATCCTATTGAGATAGACTTCCAAACCTCAGGT

CCTATATTTAATGCTTCTGTGCATTCAGACACACCATCAGTAATTCGGGGAGATCTGATCA
AATTGTTCTGTATCATCACTGTCGAGGGAGCAGCACTGGATCCAGATGACATGGCCTTTGA
TGTGTCCTGGTTTGCGGTGCACTCTTTTGGCCTGGACAAGGCTCCTGTGCTCCTGTCTTCC
CTGGATCGGAAGGGCATCGTGACCACCTCCCGGAGGGACTGGAAGAGCGACCTCAGCCTGG
AGCGCGTGAGTGTGCTGGAATTCTTGCTGCAAGTGCATGGCTCCGAGGACCAGGACTTTGG
CAACTACTACTGTTCCGTGACTCCATGGGTGAAGTCACCAACAGGTTCCTGGCAGAAGGAG
GCAGAGATCCACTCCAAGCCCGTTTTTATAACTGTGAAGATGGATGTGCTGAACGCCTTCA
AGTATCCCTTGCTGATCGGCATCGGTCTGTCCACGGTCATCGGGCTCCTGTCCTGTCTCAT
CGGGTACTGCAGCTCCCACTGGTGTTGTAAGAAGGAGGTTCAGGAGACACGGCGCGAGCGC
CGCAGGCTCATGTCGATGGAGATGGACTAGGCTGGCCCGGGAGGGGAGTGACAGAGGGACG
TTCTAGGAGCAATTGGGGCAAGAAGAGGACAGTGATATTTTAAAACAAAGTGTGTTACACT
AAAAACCAGTCCTCTCTAATCTCAGGTGGGACTTGGCGCTCTCTCTTTTCTGCATGTCAAG
TTCTGAGCGCGGACATGTTTACCAGCACACGGCTCTTCTTCCCACGGCACTTTCTGATGTA
ACAATCGAGTGTGTGTTTTCCCAACTGCCAGTGATAAAGAATCCGCGCCAGCACAATGGAT
CTCGAGGGATCTTCCATACCTACCAGTTCTGCGCCTGCAGGTC.

FIGURE 59

CCCACGCGTCCGCGGTAAATCACTTCTTGGAGGTGCCCTGCACGCCGGTCCTGGGAGCAGG
CGGCCTCCCGGGGGTGCGGGAGCCCCACTCCTCCGTGGTGTGTTCCATTTGCTTCCCACAT
CTGGAGGAGCTGACGTGCCAGCCTCCCCCAGCACCACCCAGGGACGGGAGGCATGAGCCGG
TCAAGGCACCTGGGCAAAATCCGGAAGCGTCTGGAAGATGTCAAGAGCCAGTGGGTCCGGC
CAGCCAGGGCTGACTTTAGTGACAACGAGAGTGCCCGGCTGGCCACGGACGCCCTCTTGGA
TGGGGGTTCTGAAGCCTACTGGCGGGTGCTCAGCCAGGAAGGCGAGGTGGACTTCTTGTCC
TCGGTGGAGGCCCAGTACATCCAGGCCCAGGCCAGGGAGCCCCCGTGTCCCCCAGACACCC
TGGGAGGGGCGGAAGCAGGCCCTAAGGGACTGGACTCCAGCTCCCTACAGTCCGGCACCTA
CTTCCCTGTGGCCTCAGAGGGCAGCGAGCCGGCCCTACTGCACAGCTGGGCCTCAGCTGAG
AAGCCCTACCTGAAGGAAAAATCCAGCGCCACTGTGTACTTCCAGACCGTCAAGCACAACA
ACATCAGAGACCTCGTCCGCCGCTGCATCACCCGGACTAGCCAGGTCCTGGTCATCCTGAT
GGATGTGTTCACGGATGTGGAGATCTTCTGTGACATTCTAGAGGCAGCCAACAAGCGTGGG
GTGTTCGTTTGTGTGCTCCTGGACCAGGGAGGTGTGAAGCTCTTCCAGGAGATGTGTGACA
AAGTCCAGATCTCTGACAGTCACCTCAAGAACATTTCCATCCGGAGTGTGGAAGGAGAGAT
ATACTGTGCCAAGTCAGGCAGGAAATTCGCTGGCCAAATCCGGGAGAAGTTCATCATCTCG
GACTGGAGATTTGTCCTGTCTGGATCTTACAGCTTCACCTGGCTCTGCGGACACGTGCACC
GGAACATCCTCTCCAAGTTCACAGGCCAGGCGGTGGAGCTGTTTGACGAGGAGTTCCGCCA
CCTCTACGCCTCCTCCAAGCCTGTGATGGGCCTGAAGTCCCCGCGGCTGGTCGCCCCCGTC
CCGCCCGGAGCAGCCCCGGCCAATGGCCGCCTTAGCAGCAGCAGTGGCTCCGCCAGTGACC
GCACGTCCTCCAACCCCTTCAGCGGCCGCTCGGCAGGCAGCCACCCCGGTACCCGAAGTGT
GTCCGCGTCTTCAGGGCCCTGTAGCCCCGCGGCCCCACACCCGCCTCCACCGCCCCGGTTC
CAGCCCCACCAAGGCCCTTGGGGAGCCCCGAGTCCCCAGGCCCACCTCTCCCCGCGGCCCC
ACGACGGCCCGCCCGCCGCTGTCTACAGCAACCTGGGGGCCTACAGGCCCACGCGGCTGCA
GCTGGAGCAGCTGGGCCTGGTGCCGAGGCTGACTCCAACCTGGAGGCCCTTCCTGCAGGCC
TCCCCTCACTTCTGAAGGTCCCATCCCCTGCTGCCCTCCGCAGGCCCAGGGCTGGGCACTC
CCTGAGACCCAAAGACCCACCTCAACGACGAGTGGCGTTGAGCCACTTCCCTTTGAAAAGA
CACTCAAAATCACTGCCATGGTTCAATGTTGCCCAGGCCCCAGGCCATCCACTTGCCGGCCC
CCACCAGTTCTTGGGTTCCCCGCTCTAGTTTGACCTGTGCAGCACATTCCAGAAGGTTCCA
GGGGAGGTTGTGGGGCAGCTAGAGGACAAAATCATGAAAACAGAGTCCCTGTCTTCCAGAG
ATCATCCGGGGCTTTAATATTAATGGCCCCCAAAACTCCGTAAGAAGCAGGAAATGCAGCC
CAAGTTTTACAAATGGGTAAACAGAGGCACTGAGAGATAGATGGTAGTTTGGTACTTCTGG
TTCCCAGTGCCCAGGAATGGTCCACTCCCAAGAAATTCAGGAAAGAAAGACTGAGGAGAAG
GTGTGGGAACATTCTGGATGTTTCGGGAGAGTTGGGGAAACTCCTCCTCTTAGGAAAGGCT
AATACTAGGGTATCCTTGGGCCCAATGAATTAGGGGTGAGGCCCCAGAACCCGTTATCTAT
GAGTTGTATGGGGGAGCCATCTGAAGCTGTAGCCACCAGGGATGCAGCTAGCTGAGGAGTT

TGGGGTGTTGGGTTGGACAAGGCAGGTTAGTAGACTCAGATTCTTGCTTCAAAGAGCCTTG
GGCTGGCCTGGAGGTCCCTGGAGTCTAGACTGGACCTAGGAGCTTGAGTTGTCAGGGGCCA
GGACTGGCCCCACTGCAGTGCCCAGGCCAGTCTTGAGCAGCAGGGAGGGCTCAGCTGTCCC
CAGATCCAGGTGCCTCTGACCAGCCTGGTCACCTCCTGAGGAATAAATGCTGAACCTCACA
AGCCCCATCATTCATTTCTTCTCAATTCACAGTGCCCCTCTTTGTTTCTGGGGTGGAACTA
GGTCCTGAGGGCACAGCCTAGCTGAGTGCAAAGAAATATAGGATGCTTAGAAAGCATACAG
GAGGGGCCAGGCGTGGTGGCTCATGCCTGTAATCCCAGAACTTTGGGATGCCAAGGTGGTT
GGATTACCTGAGATCAGGTGGATTACCTGGTCTCGAGACCAGCCTGACCAATATGGTGAAA
CCCCGTCTCTACTAAAAATACAAAAATTAGGCTGAGACAGGAGAATTGCTTGAACCCAGGA
AGCAGAGGTTGCAATGAGCTGAGATTGCATCACTGCACTCCAGCATGGGCAACAAAGCAAG
ACTCCGTCACAGAAAAAAAAAAAAAA

FIGURE 60

ATGGCTCTGCCGCAGCCGCCGCCGCCCGCCGGGGCGCGTGACCCGGTGACCCCGGAGCACT
GGGCATCCGGACCCCCCTCGCAGGCGCAGCCGCTATTAAGGCAGGAGGCTAAAGAGGAGGA
GGAAGGGGAGGAGACAGGCGTCCAGGGCGCCTGGGGAACCGGCACGGCGGAGCAGCGGCGG
CGGGGCTGGGGGGAGGCCGCCGAGTCGGCCGCGGCCGAGGAGGGGCAGGCGGAGGTCGGGG
GCGCCGCGGCCGCGGGGTCAGGTTCCCCGGCGGGAGGCGCAGGTGGCGGCCTGGGGAGCTG
GCGGCCGCTCCTCGCCTGGCTGCAGCGGCGGCAGCCCCAGTGCTGCCCCTGTGCGGCGCCC
CTTTCCCGCTCCGCCGCGCACTGTTGTCATGGAGGAACCAAGATGGCGGCTCTGGCCTACA
ACCTGGGCAAGCGGGAGATCAACCACTACTTCAGCGTGAGGAGCGCCAAGGTGCTGGCGCT
GGTGGCCGTGCTGCTGCTCGCAGCGTGCCACCTCGCCTCCCGCCGCTACCGAGGCAATGAT
TCGTGTGAATACCTTCTCTCAAGTGGCAGATTTCTTGGAGAGAAAGTTTGGCAACCTCACA
GTTGTATGATGCATAAATACAAAATCAGTGAAGCAAAGAACTGCCTTGTAGATAAACATAT
TGCATTTATTGGAGATTCCAGAATTCGTCAATTGTTTTATTCTTTTGTAAAAATAATTAAT
CCCCAATTCAAAGAAGAAGGAAATAAGCATGAAAACATTCCTTTTGAAGACAAGACTGCAT
CAGTTAAAGTGGATTTTCTGTGGCATCCTGAAGTTAATGGTTCTATGAAACAGTGTATCAA
AGTGTGGACTGAGGATTCCATTGCAAAGCCACATGTGATTGTAGCAGGAGCTGCCACATGG
TCCATCAAGATTCACAATGGTAGCAGTGAAGCGCTTTCTCAATATAAAATGAACATCACCT
CCATAGCACCACTTTTAGAAAAATTGGCAAAGACTAGTGATGTTTATTGGGTCTTACAAGA
TCCTGTTTATGAAGATCTATTAAGTGAAAATAGGAAGATGATCACTAATGAGAAGATAGAT
GCTTACAATGAAGCTGCAGTCAGTATTTTGAATAGTAGCACCAGAAATTCTAAATCAAATG
TTAAGATGTTCAGTGTTTCCAAATTAATTGCTCAAGAAACCATCATGGAATCTTTGGATGG
CTTACATCTTCCTGAATCGAGCAGAGAAACTACTGCAATGATTCTTATGAATGTGTATTGC
AATAAGATTTTGAAGCCTGTAGATGGGTCCTGTTGTCAACCTCGGCCTCCTGTTACTCTCA
TACAGAAGCTAGCTGCTTGTTTTTTCACTTTATCTATTATCGGATATTTAATTTTTTACAT
AATTCATCGTAATGCTCATCGGAAGAATAAGCCGTGTACTGATTTGGAAAGTGGAGAGGAA
AAGAAAAATATTATCAATACCCCTGTGTCTTCATTAGAAATACTTTTACAATCTTTCTGCA
AACTTGGCCTGATTATGGCATATTTCTATATGTGTGACCGTGCAAATCTGTTCATGAAGGA
AAACAAATTTTATACACATTCATCTTTCTTTATTCCAATTATCTACATTTTGGTTTTGGGA
GTATTTTATAATGAAAATACTAAAGAGACTAAAGTATTAAATAGAGAACAAACAGACGAAT
GGAAAGGCTGGATGCAACTTGTGATTTTGATTTATCACATTTCTGGAGCAAGTACATTTTT
GCCTGTATACATGCACATTCGAGTTCTGGTTGCTGCATATTTATTTCAGACAGGGTATGGG
CATTTCTCATACTTTTGGATAAAAGGAGATTTTGGAATCTATAGAGTATGTCAGGTTTTAT
TTCGTCTCAATTTCCTGGTAGTGGTGTTATGTATAGTAATGGATCGACCTTATCAATTCTA
TTACTTTGTCCCCTTGGTCACTGTATGGTTCATGGTCATATATGTTACTTTAGCACTATGG
CCACAAATAATCCAAAAAAAAGCAAACGGAAATTGTTTCTGGCATTTTGGCTTACTGTTGA
AACTAGGCTTTTTGCTGTTATTCATATGTTTTTTGGCATATTCTCAGGGTGCATTTGAGAA

GATCTTTTCTCTTTGGCCATTGTCCAAGTGTTTTGAACTGAAAGGGAATGTATATGAATGG
TGGTTCAGATGGAGGTTAGACCGTTATGTAGTTTTCCACGGAATGCTGTTTGCTTTTATTT
ATCTGGCTTTGCAGAAGCGTCAAATACTTTCTGAAGGAAAGGGTGAACCTCTTTTTTCAAA
CAAAATTTCAAATTTTCTGTTGTTTATTTCAGTAGTTTCTTTCTTGACCTATTCCATCTGG
GCTAGCAGTTGTAAAAACAAAGCAGAGTGCAATGAACTCCATCCGTCTGTTTCTGTGGTAC
AGATTTTAGCCTTCATCCTAATAAGAAACATCCCTGGATATGCCCGTTCAGTTTACAGTTC
ATTTTTTGCTTGGTTTGGAAAAATTTCATTAGAGCTATTTATTTGCCAGTATCACATATGG
CTGGCAGCGGACACAAGGGGTATCTTGGTACTGATACCTGGAAACCCTATGCTCAACATCA
TTGTCAGCACTTTCATATTTGTTTGTGTGGCACATGAAATTTCTCAGATCACTAATGATCT
TGCACAGATTATTATTCCTAAAGATAACTCATCTCTCTTGAAAAGGTTGGCATGTATAGCT
GCATTTTTTTGTGGACTCCTCATCTTATCATCCATTCAAGATAAATCAAAACATTAGGTTC
CAAAAATTCTAAAAAACCTAAACTCTTCAGGCTACCTTTGTGTGTCTCTAGAAGAGAAAAG
CATCTATCTGGAGATATAAATGTGTWTGTAAATATAAACGTTTGTGGCAAGAGGACAGTTC
TGTGACATCTGTTGAACATATGTGGTTGTATATATTGGAAATGTACATATCCAATATGAAA
TACTAAAACAAACAAACAAACAAAAAACCAGAATGCATTGTATAGGATTGCATGTGAAGTC
TTTTCTACTGAATCTATATTTCCATTTGTAAGTGATTTTAAGTTAACATATGAAGGCAGGG
AAATGATTACCTTTCCAGTAAAAAGTATAGATAATTTAATTAACTTAGTGACACCA:CCAA
GTGTTTTGATATAACTAAATTTGTGGTAATAAGACTGTCTGCACCTGTATTCATTGTGGAA
CTTCCTCTTTCATTGGAAACTTTCTTACTCAAGAATGACGGCAGTATTGTTTTCTTATATG
TGCAATGAAGTGGAATGATAAACAGTATGCCTTTAATTTATATGTGTTCTTGTTCTGATGT
TGTTTCCTGAAATGATTTTTCTTCCTAACTGTGGTTTTCGGGTATGCAAGCCTAAATCTTT
GTACACTTTGTCTCACAGAATAGTTCTGAGGCTCCATGACAGGGTTTTGTCATTGTTGATG
TTATTGTTGCTTCGTTTTATAAAAAAGCCAAAATTTTTTTTCCAATCCAAACGTTCACCTG
TTTCCTTTCCTCAAGCTATACCAGTGTAATACCAGTTACCCTGTGGATCCATTTAATATGT
TATCCCCACTAATTAATTTTCGTATATTATTTCCAATATTTGGAAAGCTCTTTATAGCCAT
TTGGTATTTCCTATTACCCACCTCCTATTTTAAATATTTATCAGTCTAAACTTGTGCAGTG
TAGTAAACATGCAAGTTGTTACGATTGAGCTGTATTACCATAAGTAGAATTTTAAGTAAAC
TGGTGAATTTGGGCAATAAATGTTTTTGCTTTTTGTTTGATTTTTTTTTTACAAGCTAACTG
TTAGAGGTATACATTTATTTATCTGTTGTACAGATTTGATTATGATTTTAATGTTTGAAAG
ATTGCACTTGTTTGCTTTTACTATATGTGGGGTAAAATATATTTTCTGTTCACAGTATATG
AAAATATGGAGTAATTTAAACAGTAAATAAACATTCTGTGGATGCTTATTTTTGTATTGGC
AAAGTATCAATTAAACTATATGTGTTCTTTTTCAA

FIGURE 61A

GAAAGACATACACACTTCATGTAATGCTACCTGCAAGTCTCCCTAGAAAAGCAGTTTTTGT
AGGTGAAAACAATGAAGCCAGGTAATATTGCAAGGAGGCTGTAATTTTAGCAGACCTACCA
ACAACACTGATGTAGGAAGCTCATTATTTTAATTTCTGGAGCCTTTTAATTTTTTCTTTAG
AAAGTGTATAAATAATTGCAGTGCTGCTTTGCTTCCAAAACTGGGCAGTGAGTTCAACAAC
AACGACAACAACAGCCGCAGCTCATCCTGGCCGTCATGGAGTTTCTTGAAAGAACGTATCT
TGTGAATGATAAAGCTGCCAAGATGTATGCTTTCACACTAGAAAGAAGGAGCTGCAAATGA
ACACTTCATAGCAATGTGGAACTCCAACAGAAACCGGTGAATAAAGATCAGTGTCCCAGAG
AGAGACCAGAGGAGCTGGAGTCAGGAGGCATGTACCACTGCCACAGTGGCTCCAAGCCCAC
AGAAAAGGGGGCGAATGAGTACGCCTATGCCAAGTGGAAACTCTGTTCTGCTTCAGCAATA
TGCTTCATTTTCATGATTGCAGAGGTCGTGGGTGGGCACATTGCTGGGAGTCTTGCTGTTG
TCACAGATGCTGCCCACCTCTTAATTGACCTGACCAGTTTCCTGCTCAGTCTCTTCTCCCT
GTGGTTGTCATCGAAGCCTCCCTCTAAGCGGCTGACATTTGGATGGCACCGAGCAGAGATC
CTTGGTGCCCTGCTCTCCATCCTGTGCATCTGGGTGGTGACTGGCGTGCTAGTGTACCTGG
CATGTGAGCGCCTGCTGTATCCTGATTACCAGATCCAGGCGACTGTGATGATCATCGTTTC
CAGCTGCGCAGTGGCGGCCAACATTGTACTAACTGTGGTTTTGCACCAGAGATGCCTTGGC
CACAATCACAAGGAAGTACAAGCCAATGCCAGCGTCAGAGCTGCTTTTGTGCATGCCCTTG
GAGATCTATTTCAGAGTATCAGTGTGCTAATTAGTGCACTTATTATCTACTTTAAGCCAGA
GTATAAAATAGCCGACCCAATCTGCACATTCATCTTTTCCATCCTGGTCTTGGCCAGCACC
ATCACTATCTTAAAGGACTTCTCCATCTTACTCATGGAAGGTGTGCCAAAGAGCCTGAATT
ACAGTGGTGTGAAAGAGCTTATTTTAGCAGTCGACGGGGTGCTGTCTGTGCACAGCCTGCA
CATCTGGTCTCTAACAATGAATCAAGTAATTCTCTCAGCTCATGTTGCTACAGCAGCCAGC
CGGGACAGCCAAGTGGTTCGGAGAGAAATTGCTAAAGCCCTTAGCAAAAGCTTTACGATGC
ACTCACTCACCATTCAGATGGAATCTCCAGTTGACCAGGACCCCGACTGCCTTTTCTGTGA
AGACCCCTGTGACTAGCTCAGTCACACCGTCAGTTTCCCAAATTTGACAGGCCACCTTCAA
ACATGCTGCTATGCAGTTTCTGCATCATAGAAAATAAGGAACCAAAGGAAGAAATTCATGT
CATGGTGCAATGCACATTTTATCTATTTATTTAGTTCCATTCACCATGAAGGAAGAGGCAC
TGAGATCCATCAATCAATTGGATTATATACTGATCAGTAGCTGTGTTCAATTGCAGGAATG
TGTATATAGATTATTCCTGAGTGGAGCCGAAGTAACAGCTGTTTGTAACTATCGGCAATAC
CAAATTCATCTCCCTTCCAATAATGCATCTTGAGAACACATAGGTAAATTTGAACTCAGGA
AAGTCTTACTAGAAATCAGTGGAAGGGACAAATAGTCACAAAATTTTACCAAAACATTAGA
AACAAAAAATAAGGAGAGCCAAGTCAGGAATAAAAGTGACTCTGTATGCTAACGCCACATT
AGAACTTGGTTCTCTCACCAAGCTGTAATGTGATTTTTTTTTCTACTCTGAATTGGAAATA
TGTATGAATATACAGAGAAGTGCTTACAACTAATTTTTATTTACTTGTCACATTTTGGCAA
TAAATCCCTCTTATTTCTAAATTCTAACTTGTTTATTTCAAAACTTTATATAATCACTGTT

TTGTAGAGACCCAGAAGTCTTCAAGAACAGCCGACAAAAACATTCGAGTTGACCCCACCAA
GTTGTTGCCACAGATAATTTAGATATTTACCTGCAAGAAGGAATAAAGCAGATGCAACCAA
TTCATTCAGTCCACGAGCATGATGTGAGCACTGCTTTGTGCTAGACATTGGGCTTAGCATT
GAAACTATAAAGAGGAATCAGACGCAGCAAGTGCTTCTGTGTTCTGGTAGCAACTCAACAC
TATCTGTGGAGAGTAAACTGAAGATGTGCAGGCCAACATTCTGGAAATCCTATGTCACTGG
GTTTGGTTTGGAAACCTGGACTTCTGCATTTTTAAAAGTTACCCAGAGATGCTTCTAAAGA
TGAGCCATAGTCTAGAAGATTGTCAACCACAGGAGTTCATTGAGTGGGACAGCTAGAYACA
TACATTGGCAGCTACAATAGTATCATGAATTGCAATGATGTAGTGGGGTATAAAAGGAAAG
CGATGGATATTGCCGGATGGGCATGGCCAGTGATGTTTCACGTCATTGAGGTGACAGCTCT
GCTGGACTTTGAATTACATATGGAGGCTCTCCAGGAAGACGAAGAAGAGAAGGACATTCTA
GGCAAAAAGAAGACTAGGCACAAGGCACACTTATGTTTGTCTGTTAGCTTTTAGTTGAAAA
AGCAAAATACATGATGCAAAGAAACCTCTCCACGCTGTGATTTTTAAAACTACATACTTTT
TGCAACTTTATGGTTATGAGTATTGTAGAGAACAGGAGATAGGTCTTAGATGATTTTTATG
TTGTTGTCAGACTCTAGCAAGGTACTAGAAACCTAGCAGGCATTAATAATTGTTGAGGCAA
TGACTCTGAGGCTATATCTGGGCCTTGTCATTATTTATCATTTATATTTGTATTTTTTTCT
GAAATTTGAGGGCCAAGAAAACATTGACTTTGACTGAGGAGGTCACATCTGTGCCATCTCT
GCAAATCAATCAGCACCACTGAAATAACTACTTAGCATTCTGCTGAGCTTTCCCTGCTCAG
TAGAGACAAATATACTCATCCCCCACCTCAGTGAGCTTGTTTAGGCAACCAGGATTAGAGC
TGCTCAGGTTCCCAACGTCTCCTGCCACATCGGGTTCTCAAAATGGAAAGAATGGTTTATG
CCAAATCACTTTTCCTGTCTGAAGGACCACTGAATGGTTTTGTTTTTCCATATTTTGCATA
GGACGCCCTAAAGACTAGGTGACTTGGCAAACACACAAGTGTTAGTATAATTCTTTGCTTC
TGCTTCTTTTTGAAAATCATGTTTAGATTTGATTTTAAGTCAGAAATTCACTGAATGTCAG
GTAATCATTATGGAGGGAGATTTGTGTGTCAACCAAAGTAATTGTCCCATGGCCCCAGGGT
ATTTCTGTTGTTTCCCTGAAATTCTGCTTTTTTAGTCAGCTAGATTGAAAACTCTGAACAG
TAGATGTTTATATGGCAAAATGCAAGACAATCTATAAGGGAGATTTTAAGGATTTTGAGAT
GAAAAAACAGATGCTACTCAGGGGCTTTATGGACCATCCATCAATTCTGAAGTTCTGACTC
TCCCATTACCCTTTCCCTGGTGTGGTCAGAACTCCAGGTCACTGGAAGTTAGTGGAATCAT
GTAGTTGAATTCTTTACTTCAAGACATTGTATTCTCTCCAGCTATCAAAACATTAATGATC
TTTTATGTCTTTTTTTTGTTATTGTTATACTTTAAGTTCTGGGGTACATGTGCGGAACATG
TAGGTTTGTTACATAGGTATACATGTGCCATGGTGGTTTGCTGCACTCATCAACCTGTCAT
CTACATTCTTTTATGTCTGTCTTTCAAAGCAACACTCTGTTCTTCTGAGTAGTGAAATCAG
GTCAACTTTACCACCAGCCTCCATTTTTAATATGCTTCACCATCATCCAGCACCTACTTAA
GATTTATCTAGGGCTCTGTGGTGATGTTAGGACCCATAAAAGAAATTTATGCCTTCCATAT
GTTTGGTTACAGATGGGAAATGGGAATGTTGAAGGACATGAAAGAAAGGATGTTTACACAT
TAAGCATCAGTTCTGAAGCTAGATTGTCTGAGTTTGAATCTTAGCTCTTCCCTTTATTAGC
TCTGTGACCTCGAGCTAGTTACTTAAATGCTCTGATCCTCTATTTCCTGATCAGTGAAACC

ATAGCATGTAGTGTTCAGTACATGTTAAATGTTGTTTTTTATTATGTACAAACATGTGTGG
GCACAGAATTTTAAATCATCTCAACTTTTGAGAAATTTTGAGTTATCAACACCGTTCCCAC

FIGURE 61B

AAGACAGTGGCAAAATTATTGGTGAGAATTAAACAGCTGTTTCTCAGAGGAAGCAATGGAG
GCTTGCTGGGATAAAGGCATTTACTGAGAGGCTGTTACCTAGTGAGAGTGATGAATTAATT
AAAATAGTCGAATCCCTTTCTGACTGTCTCTGAAAGCTTCCGCTTTTATCTTTGAAGAGCA
GAATTGTCACCCCAAGGACATTTATTAATAAAAAGAACAACTGTCCAGTGCAATGAAGGCA
AAGTCATAGGTCTCCCAAGTCTTACCCCATTCCTGTGAAATATCAAGTTCTTGGCTTTTCT
CTGTCATGTAGCCTCAACTTTCTCCGACCGGGTGCATTTCTTTCTCTGGTTTCTAAATTGC
CAGTGGCAAATTTGGATCACTTACTTAATATCTGTTAAATTTTGTGACCCAACAAAGTCTT
TTAGCACTGTGGTGTCAAAAAGAAAAACACCTCCCAGGCATATACATTTTATAGATTCCTG
GAGAATGTTGCTCTCCAGCTCCATCCCCACCCAATGAAATATGATCCAGAGAGTCTTGCAA
AGAGACAAGCCTCATTTTCCACAATTAGCTCTAAAGTGCCTCCAGGAAATGATTTTCTCAG
CTCATCTCTCTGTATTCCCTGTTTTGGATCACAGGGCAATCTGTTTAAATGACTAATTACA
GAAATCATTAAAGGCACCAAGCAAATGTCATCTCTGAATACACACATCCCAAGCTTTACAA
ATCCTGCCTGGCTTGACAGTGATGAGGCCACTTAACAGTCCAGCGCAGGCGGATGTTAAAA
AAAATAAAAAGGTGACCATCTGCGGTTTAGTTTTTTAACTTTCTGATTTCACACTTAACGT
CTGTCATTCTGTTACTGGGCACCTGTTTAAATTCTATTTTAAAATGTTAATGAGTGTTGTT
TAAAATAAAATCAGGAAAGAGAGAAAAAAAAAAAAAAAAAAA

FIGURE 62A

GCGGCGGGAGAGCTGCTGGCTCGCCCGGATCCCGGGAGCTGCCTGGAGGCGGGCCCGGCCC
GGGGAAGGTGAGCGGCTGCGGGACCCAGCCCCTCGCCGGGAGCGGGCACCATGGTGCTGTC
GGTGCCTGTGATCGCGCTGGGCGCCACGCTGGGCACAGCCACCAGCATCCTCGCGTTGTGC
GGGGTCACCTGCCTGTGTCGGCACATGCACCCCAAGAAGGGGCTGCTGCCGCGGGACCAGG
ACCCCGACCTGGAGAAGGCGAAGCCCAGCTTGCTCGGGTCTGCACAACAGTTCAATGTTAA
AAAGTCCACGGAACCTGTTCAGCCCCGTGCCCTCCTCAAGTTCCCAGACATCTATGGACCC
AGGCCAGCTGTGACGGCTCCAGAGGTCATCAACTATGCAGACTATTCACTGAGGTCTACGG
AGGAGCCCACTGCACCTGCCAGCCCCCAACCCCCGAATGACAGTCGCCTCAAGAGGCAGGT
CACAGAGGAGCTGTTCATCCTCCCTCAGAATGGTGTGGTGGAGGATGTCTGTGTCATGGAG
ACCTGGAACCCAGAGAAGGCTGCCAGTTGGAACCAGGCCCCCAAACTCCACTACTGCCTGG
ACTATGACTGTCAGAAGGCAGAATTGTTTGTGACTCGCCTGGAAGCTGTGACCAGCAACCA
CGACGGAGGCTGTGACTGCTACGTCCAAGGGAGTGTGGCCAATAGGACCGGCTCTGTGGAG
GCTCAGACAGCCCTAAAGAAGCGGCAGCTGCACACCACCTGGGAGGAGGGCCTGGTGCTCC
CCCTGGCGGAGGAGGAGCTCCCCACAGCCACCCTGACGCTGACCTTGAGGACCTGCGACCG
CTTCTCCCGTCACAGCGTGGCCGGGGAGCTCCGCCTGGGCCTGGACGGGACATCTGTGCCT
CTAGGGGCTGCCCAGTGGGGCGAGCTGAAGACTTCAGCGAAGGAGCCATCTGCAGGAGCTG
GAGAGGTCCTACTATCCATCAGCTACCTCCCGGCTGCCAACCGCCTCCTGGTGGTGCTGAT
TAAAGCCAAGAACCTCCACTCTAACCAGTCCAAGGAGCTCCTGGGGAAGGATGTCTCTGTC
AAGGTGACCTTGAAGCACCAGGCTCGGAAGCTGAAGAAGAAGCAGACTAAACGAGCTAAGC
ACAAGATCAACCCCGTGTGGAACGAGATGATCATGTTTGAGCTGCCTGACGACCTGCTGCA
GGCCTCCAGTGTGGAGCTGGAAGTGCTGGGCCAGGACGATTCAGGGCAGAGCTGTGCGCTT
GGCCACTGCAGCCTGGGCCTGCACACCTCGGGCTCTGAGCGCAGCCACTGGGAGGAGATGC
TCAAAAACCCTCGCCGGCAGATTGCCATGTGGCACCAGCTGCACCTGTAACCAGCTGCCCA
GCTGCCTCCCTTCTTGGACAGCCCTGACCCGTCCTCTGCAACCTCCTTTCTGTGCCCCTTT
CCTCATTCTGACACCCAGAAGACAGTGACAGATGTGTTTGCAAGGCTGGGATGGCTCTCTC
ATCATACTCTTGTTTCTTAGAAATAAGCAAGACAGAGCAGGAAATGGAATATGCGGGTCAC
ACTGAGGAATGCATTTTGCTCATCTGTGTTATTGAAGGAGGTGCTTATTAAATACAGTTCC
TATGCCTGTTTTATAGGTGGGGTTAGGCCAGATGCAGAGAAAGCTAAATGTGGGAATCATG
GATGCAAAGAAGAATTTGGCTTTTTGAAAAACAAGCATTTCAAAAATGATGAAGGAAGTGA
AAGTATCCTGGATCAACTCCTAGAGTTAGAGATTGCCCAGGTGGAAAGAAACCTTAGCCAG
CGTTCAATCAAGCTCACCATGCAGGGCAGTCACCCGGCAGTTCTCAAACTTTAGCATGTGA
AGAGTCACCAGCAGATTCCTGGGCTCGCCTGGAGACATTCCTAGTCGGTATTCCTGGTCGA
AGCCCAGGAGCCTTCCTTTTTAACAAGCTGATGTAGAGGGTGGAGCACTGTATGTGGAGAA
ATTCCTTCTACAATATTCCACACAGGTTTTTGGCCACAGTCCTTGATGGAGTCCCAAAACC
[illegible]

CTTTTTAATTTAATTATTTTAATTCTTAATACTTTAATTTGTGCATTTCATAAGCCCCCTG
CTCTTGGACTGAATTTTGTGCTTTTTATTGAAGAATTTTATTGTTTTTATCTTAAAATCAG
TTTCTATTATCCTTGGGGAGACCATCCCTAACAAAGTACAGGTGGGATCTCCTGTGAGTCA
TTGGCTGGGTTCTGATTGCTAGATGTCACACCCACCAGCATCACCAAAGTGACTCTGAGAT
AGACCGGTCCCTTCTCAGCGTTCCAGTCACTTCAGGAGGAATTTAGTTATTGACTTAGTCT
ATGACATCTGGCTACATGTAGGTAGAGAAGAAAGACAATTTTAAAAAGGAAATCAGGTCTT
TTGCAACTGTGCCTCCCTCTGTCTGTTTTCACTTGAATGGGTAAATAACCAGCAGCTAGGT
TTTGAATTCCTACCTTGTTATTCTAAACAGATGTCCACATTGTTAATTAAATCTAAATTAT
GAGCCTTGCTGAGTGGATACGGTACTTACACCTGAACCAGGATTCCTGGGTTCTGTTGTTG
ACATTGCCCTTCAGCACCTGTTTGGCCAGCTGTATAAGATAGGACTAATGACTAGGAAGCC
TACCCCAATGAATGATATACTAGATGAAATAGTGTTCAAAACCTGTAGGCACTCTCTGGCT
AAAAACAAACTCTGAGGCCACCAGCAGATCATCTTTAAGCTAAGTTACTATTTTTCACCTT
TTTTTTTAGACGGAGTTTTGCTCTTTGTTGCCCAGGCTGGAGTGCAGTGGCACGATCTCGG
CTCACTGCAACCTCCGCCTCCCAAGTTCAAGCGATTCTCCTGTCTCAGCCTCCTGGGTAGC
TGGGATTACAGGTGCCCACCAACATGCCTGGCTAATTTTTGTACTTTTAGTAGAGATGGGG
TTTCACCATGTTGGCCAGGCTGGTCTTCAACTCCAGATCTCAGGTGATCTACCCTCCTCGG
CCTCCCAAAGTACTGGGATTACAGGCCTGAGCCACCGCGCCCGGCCTATTTTTCACTTTAA
TTTGGCAGCTGAGAATGCCCAAAAAGTGCCAGAAGCATCGTGGCATTTCCAGAACCATGGA
TTCTGCCTTTGGACCCCTCTCTATTAATATTAAAACTCTGGGCCTTCAGATGTCACCCTAA
TCCACTGCCCTAAGACAGAATTTCTGGACAAGATGGGTAAGGGCTTCATTCCTTCAACAAG
TCAAGTCATACTTGGCCTCTCCCTGAGAATCTGAGCAGGAGCCTTATAACCTGTGGTCATT
ATTTTTTCTTTCTGTACAGAAATAGAAAAGCATTAGAAATAACTTCTAACCATCCTCTGAA
AAAACAGAAAAAATATCGAATCCCTCTTTCATGAGAAGTCTTTTGGATAATTGGAAACCTT
CATCACTGAGGTTGGCCAGCCCCTGCCAAGTGTTGTGTAGGCAAAGCACTTGTTAGTGGCT
TCCTATGAAATGTTTTAGAGATCTCTTCACCATACTGGTTTCTTCTCTTTGGTTGGTGTGG
GTAAAAGAAAACAAAACATTTCCTATAAGCTGAAAGCTGACCAGCATTCTCTTCTTGGTAA
CATCTACTACTCCAACCTAGAAAATTTGGATTCTAGACCAAAAATCAGGAAACATGGCTCC
TTATAAATCTGTGCAGCTGCCTTATAGTACCATCAAAGGAATTTCAGGTGGGCTGGGCGGG
GCCCCGATCCCAGAATTATCAACTCCACCCATCATCATTTGGTCATGAAGCATCCTTTCAT
TCTTCTTCTTCTTTTTTTTGGGGGGGGCGGGGCGGGGGAGGGATCTCAAAGTTTTAGTCTT
CCAGAATCCAAATTAAAGGTTGCCCCTGATGGGGGCCAGGTTCCGCCACAGAACATCTTAG
ATGTCAGCCTTGACCTCACTTAGCAGGGATTACAGAAATGAGATACATTTTGAAGGAGAGT
TGTCTGTTATGTTCACTGTATTCTAAGTGCCTGGGATAAAGCTGTCTCATGGGTGCTCCAT
ATATATTCATATATATTTGTTGAGTGAATTAATGAATTAAGAGTGGCTGGCAGAGTAGGCA
GAAAAAGACACTGCAAATGGCATAAAAATTAAAGTCCTAGCTGAGTTCTCAATGGTAAAGG
CATCAGATGTCTTAGCAGTCAAGCTAGAAATTCATGACAATGAGTATTACTATTTGCCTAA

CTTTTCCACTAAAACTCCATCTTAGTGAATTTTAAATTATCCAGAGATGTCAAACTGCCAA
ATAAAAATATTTCAGTAGTCTTTGCATCAGCTTACCTTGTACCAGAAACATTTCCAATTTA

FIGURE 62B

CTATCAAATTATAGTAACTGAGCCTGTGTGAAGTATCTCATCATTTTCGAAAGGAACACCT
TGTGTGATGCCAGTGAGCATTTCTAAAAAGGGTGTGAGGTAGAGGTAAAAATAAGGTGAGA
GACCATTTCAGAATGCACTGTTGCTCAAAAAGGTGATCTGGTTCTTTCTTCAGAGATTTCT
ACGGGGATAGAAAATCGGGAGTCTGCCCTCATTAATCTGTGACTCCACCTCTTGCATCAAA
TCAATATCTATTTGTTGAGCACTTATTGATTAAGACCTTGCATATGTCTGTCCATTTTGAT
TTGAGATACAACTTTTTTGTGTGGGTTGAATGACAAATCACTCCAAACAAAACTGGGCACAG
AGAATCAGCTAGGAGACCAGTTATTCAGGGTCCATTTCTCTTGGATGTAAAGGAGTCCTGG
GTAAAATGTGGCTGTAACCTAAACCAACTAGTCCTTGTGATTTGTTTCTGCCCTCTGTGTT
TCCTGTTGTCAAATGCTAAGTGTGTGTTTTGCAGTCATGAACTAAAGCACAAAAAGATGCA
TGAGACATTGTAGTCATATGTCTGGTGTGACACTTTGGAGCAAAAACCTTGCAGTGGTAAA
TAAAAAATTTCCAACAGGGAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 63A

AGCGGCAGCTCCGCTGAAAACTGCATTCAGCCAGTCCTCCGGACTTCTGGAGCGGGGACAG
GGCGCAGGGCATCAGCAGCCACCAGCAGGACCTGGGAAATAGGGATTCTTCTGCCTCCACT
TCAGGTTTTAGCAGCTTGGTGCTAAATTGCTGTCTCAAAATGCAGAGGATCTAATTTGCAG
AGGAAAACAGCCAAAGAAGGAAGAGGAGGAAAAGGAAAAAAAAAAGGGGTATATTGTGGATG
CTCTACTTTTCTTGGAAATGCAAAAGATTATGCATATTTCTGTCCTCCTTTCTCCTGTTTT
ATGGGGACTGATTTTTGGTGTCTCTTCTAACAGCATACAGATAGGGGGGCTATTTCCTAGG
GGCGCCGATCAAGAATACAGTGCATTTCGAGTAGGGATGGTTCAGTTTTCCACTTCGGAGT
TCAGACTGACACCCCACATCGACAATTTGGAGGTGGCAAACAGCTTCGCAGTCACTAATGC
TTTCTGCTCCCAGTTTTCGAGAGGAGTCTATGCTATTTTTGGATTTTATGACAAGAAGTCT
GTAAATACCATCACATCATTTTGCGGAACACTCCACGTCTCCTTCATCACTCCCAGCTTCC
CAACAGATGGCACACATCCATTTGTCATTCAGATGAGACCCGACCTCAAAGGAGCTCTCCT
TAGCTTGATTGAATACTATCAATGGGACAAGTTTGCATACCTCTATGACAGTGACAGAGGC
TTATCAACACTGCAAGCTGTGCTGGATTCTGCTGCTGAAAAGAAATGGCAAGTGACTGCTA
TCAATGTGGGAAACATTAACAATGACAAGAAAGATGAGATGTACCGATCACTTTTTCAAGA
TCTGGAGTTAAAAAAGGAACGGCGTGTAATTCTGGACTGTGAAAGGGATAAAGTAAACGAC
ATTGTAGACCAGGTTATTACCATTGGAAAACACGTTAAAGGGTACCACTACATCATTGCAA
ATCTGGAATTTACTGATGGAGACCTATTAAAAATCCAGTTTGGAGGTGCAAATGTCTCTGG
ATTTCAGATAGTGGACTATGATGATTCGTTGGTATCTAAATTTATAGAAAGATGGTCAACA
CTGGAAGAAAAAGAATACCCTGGAGCTCACACAACAACAATTAAGTATACTTCTGCTCTGA
CCTATGATGCCGTTCAAGTGATGACTGAAGCCTTCCGCAACCTAAGGAAGCAAAGAATTGA
AATCTCCCGAAGGGGGAATGCAGGAGACTGTCTGGCAAACCCAGCAGTGCCCTGGGGACAA
GGTGTAGAAATAGAAAGGGCCCTCAAACAGGTTCAGGTTGAAGGTCTCTCAGGAAATATAA
AGTTTGACCAGAATGGAAAAAGAATAAACTATACAATTAACATCATGGAGCTCAAAACTAA
TGGGCCCCGGAAGATTGGCTACTGGAGTGAAGTGGACAAAATGGTTGTTACCCTTACTGAG
CTCCCTTCTGGAAATGACACCTCTGGGCTTGAGAATAAGACTGTTGTTGTCACCACAATTT
TGGAATCTCCGTATGTTATGATGAAGAAAAATCATGAAATGCTTGAAGGCAATGAGCGCTA
TGAGGGCTACTGTGTTGACCTGGCTGCAGAAATCGCCAAACATTGTGGGTTCAAGTACAAG
TTGACAATTGTTGGTGATGGCAAGTATGGGGCCAGGGATGCAGACACGAAAATTTGGAATG
GGATGGTTGGAGAACTTGTATATGGGAAAGCTGATATTGCAATTGCTCCATTAACTATTAC
CCTTGTGAGAGAAGAGGTGATTGACTTCTCAAAGCCCTTCATGAGCCTCGGGATATCTATC
ATGATCAAGAAGCCTCAGAAGTCCAAACCAGGAGTGTTTTCCTTTCTTGATCCTTTAGCCT
ATGAGATCTGGATGTGCATTGTTTTTGCCTACATTGGGGTCAGTGTAGTTTTATTCCTGGT
CAGCAGATTTAGCCCCTACGAGTGGCACACTGAGGAGTTTGAAGATGGAAGAGAAACACAA
AGTAGTGAATCAACTAATGAATTTGGGATTTTTAATAGTCTCTGGTTTTCCTTGGGTGCCT
[illegible]ATCCCTCTCTGGGCGCATTGTTGGAGGTGT

GTGGTGGTTCTTTACCCTGATCATAATCTCCTCCTACACGGCTAACTTAGCTGCCTTCCTG
ACTGTAGAGAGGATGGTGTCTCCCATCGAAAGTGCTGAGGATCTTTCTAAGCAAACAGAAA
TTGCTTATGGAACATTAGACTCTGGCTCCACTAAAGAGTTTTTCAGGAGATCTAAAATTGC
AGTGTTTGATAAAATGTGGACCTACATGCGGAGTGCGGAGCCCTCTGTGTTTGTGAGGACT
ACGGCCGAAGGGGTGGCTAGAGTGCGGAAGTCCAAAGGGAAATATGCCTACTTGTTGGAGT
CCACGATGAACGAGTACATTGAGCAAAGGAAGCCTTGCGACACCATGAAAGTTGGTGGAAA
CCTGGATTCCAAAGGCTATGGCATCGCAACACCTAAAGGATCCTCATTAGGAACCCCAGTA
AATCTTGCAGTATTGAAACTCAGTGAGCAAGGCGTCTTAGACAAGCTGAAAAACAAATGGT
GGTACGATAAAGGTGAATGTGGAGCCAAGGACTCTGGAAGTAAGGAAAAGACCAGTGCCCT
CAGTCTGAGCAACGTTGCTGGAGTATTCTACATCCTTGTCGGGGGCCTTGGTTTGGCAATG
CTGGTGGCTTTGATTGAGTTCTGTTACAAGTCAAGGGCCGAGGCGAAACGAATGAAGGTGG
CAAAGAATGCACAGAATATTAACCCATCTTCCTCGCAGAATTCACAGAATTTTGCAACTTA
TAAGGAAGGTTACAACGTATATGGCATCGAAAGTGTTAAAATTTAGGGGATGACCTTGAAT
GATGCCATGAGGAACAAGGCAAGGCTGTCAATTACAGGAAGTACTGGAGAAAATGGACGTG
TTATGACTCCAGAATTTCCCAAAGCAGTGCATGCTGTCCCTTACGTGAGTCCTGGCATGGG
AATGAATGTCAGTGTGACTGATCTCTCGTGATTGATAAGAACCTTTTGAGTGCCTTACACA
ATGGTTTTCTTGTGTGTTTATTGTCAAAGTGGTGAGAGGCATCCAGTATCTTGAAGACTTT
TCTTTCAGCCAAGAATTCTTAAATATGTGGAGTTCATCTTGAATTGTAAGGAATGATTAAT
TAAAACACAACATCTTTTTCTACTCGAGTTACAGACAAAGCGTGGTGGACATGCACAGCTA
ACATGGAAGTACTATAATTTACCTGAAGTCTTTGTACAGACAACAAACCTGTTTCTGCAGC
CACTATTGTTAGTCTCTTGATTCATAATGACTTAAGCACACTTGACATCAACTGCATCAAG
ATGTGACATGTTTTTATAAAAAAAGGAAAAAAAACATTTAAAACTAAAAAATATTTTTAGGT
ATTTTCACAAACAAACTGGCTTTTAAATAAATTTGCTTCCATATTGGTTGAATAAGACAAA
AACAATTAAACTGAGTGGGAAGTGAATAAAAAAAGGCTTTAGGTATCGATTCCATATTTTT
CAAAGCCAAATATGTAAATGCTAAGGAAAGTAAACAAAGAGGAGATTCCAATCTTGTAATT
TAATATTGTTATTAAAACTTTAATGTATCCTATTCTTTAACATTTGGTGTTAATATAAAAT
TACTTGGCAATGCTTGACATTTGAAATAAACATTTTTCTATTGTTTTATTGCAAGTGGTCC
AATTAATTTTGCTTAGCTACAGTTTGGTCATAAATCAAGTGAGTTTAAAGACACTACCAAG
TTGTTAGGTGCCCAGAGAAAATTTCTCCCTTTTAAAAAGGCCAGGTGATTTTTCAAATGTA
ATCTTGCCCCCAAAGTAATATCCGAATATCTTTTTGACATGTCTAAATATATATATATATA
AAGAAATATTTGTTAACACAAAAGCATTTGATCTATGTAGATAAATGCTAATAGATTTAAA
AAGCTAATATTAACAAATACCAGAATACGTGAAGTTCCATTTTTAAAGTGTTTGAGCTTAC
AGAAGAGAAACATTCATTTTAAATGAAGTAAAAAATGCCTTGAAAGTAATTCTTTAGATAG
TTGCCCATTGATTAAATTCCAAAAACTAAATATGTTTTTAGCTTTAAAATTATAAAAGCTG
TCATAAACTTTATATATTATGAATTTTAAAATATGTTTGAGTCTCCTGCAATATAGTTTCA
TCCCATTGACATCAATTAAAAATAACCCTAATATATTATTTTTATATTTATTCCTCAGGTG

AAGAGCCTATATTTATCTAGTTAATGAATTTAAAGGATCTATCTTTCCCTTCATAAAATAC
CTCTTATTTCCATTAAAGCCCCCCAAGTTTAATTAATTTAGGATTTTGAATGATTATTGAC

FIGURE 63B

ATCCAATAGTTATTTTTAATATTTGTATTCTTGTTATTTCTGGAAGAAAGCCTTTGTGTAG
CACTTGGTATTTTGCAAAGTGCTTTTAAAACATTCTTACTTACCGTATTTCATAGAAGGGA
AGGAAAAATGTAAGGTTTAACAGTAAGCACTTGCATTGAACATGGAGGCATGTGGTATCAT
GATATTCTTCACTAAATTTAGCTGTCCCTAATCACAGATCCTAAGGTAATATAATATAATT
TTAGTGCATTTCTCCTCATCAGGAATGCTGGAGGTGCATTTTAAGTTTTAATAATAAGTGC
TAGAATGACCAAATTGCAGACTAATTGTTTCCATATTGTACTTAAAATGAGTTTTTAAAAG
TGAAAAAGAAATGACTATATACAATCAATGCTATTTATTGTACCTCTGGGCCTACTCTTCT
AAAAATTGTAGCTTATCGATTTTTCTCTGTCAAGCTTGAACTAATGTAAATAATTGAAATA
ATGTAAAGTTATATTTTCATGTTTTTATAGATACAACATGACAAGAATACATAATGTAAGA
GTATTTCAACTATGGATAATGTTGATTGGATAATGCACATCTCAGTTACAAGCAGTACTCA
TAGTTTAATATCCATGTAACGGTGCATCAATATATTGCTATATAAATATGTCTGTGTGCAT
ATAAGTGAAAAGTGGTCAAACAAGAGTGATGACAGCTGTCTAAAGGTTTTTTTATTCATTT
TATATAAAAACTGTTATGGAAAGACCAAAATGTTTATGAACTATTCTTATGTAAATTTACA
ATTGTCCTTTACTGTACTTTTTTGTTTACAGTATAGTACCTTATTTTCTGCTGTGTTAAGT
GGGTGTCAAACTCCAAGAAGACATACACTTTCTATAACTTCTATTGAAGATATTGGAATTT
CCAATTTTTCATGTGTACTATGTCAGAAAATGCTTTCGATTTTATTTTTAAATCTAACATC
GGATGGCTTTTCCGGAGTGTTGTAAAAACTTCAATCATACATAAAACATGTTCTTACAAAA
GGCAAAGATCTT

FIGURE 64

GCCAACACTGGCCAAAAGGGATAGCTGTCCAAGGTCTCCCCCAGCACTGAGGAGCTCGCCT
GCTGCCCTCTTGCGCGCGGGAAGCAGCACCAAGTTCACGGCCAACGCCTTGGCACTAGGGT
CCAGAATGGCTACAACAGTCCCTGATGGTTGCCGCAATGGCCTGAAATCCAAGTACTACAG
ACTTTGTGATAAGGCTGAAGCTTGGGGCATCGTCCTAGAAACGGTGGCCACAGCCGGGGTT
GTGACCTCGGTGGCCTTCATGCTCACTCTCCCGATCCTCGTCTGCAAGGTGCAGGACTCCA
ACAGGCGAAAAATGCTGCCTACTCAGTTTCTCTTCCTCCTGGGTGTGTTGGGCATCTTTGG
CCTCACCTTCGCCTTCATCATCGGACTGGACGGGAGCACAGGGCCCACACGCTTCTTCCTC
TTTGGGATCCTCTTTTCCATCTGCTTCTCCTGCCTGCTGGCTCATGCTGTCAGTCTGACCA
AGCTCGTCCGGGGGAGGAAGCCCCTTTCCCTGTTGGTGATTCTGGGTCTGGCCGTGGGCTT
CAGCCTAGTCCAGGATGTTATCGCTATTGAATATATTGTCCTGACCATGAATAGGACCAAC
GTCAATGTCTTTTCTGAGCTTTCCGCTCCTCGTCGCAATGAAGACTTTGTCCTCCTGCTCA
CCTACGTCCTCTTCTTGATGGCGCTGACCTTCCTCATGTCCTCCTTCACCTTCTGTGGTTC
CTTCACGGGCTGGAAGAGACATGGGGCCCACATCTACCTCACGATGCTCCTCTCCATTGCC
ATCTGGGTGGCCTGGATCACCCTGCTCATGCTTCCTGACTTTGACCGCAGGTGGGATGACA
CCATCCTCAGCTCCGCCTTGGCTGCCAATGGCTGGGTGTTCCTGTTGGCTTATGTTAGTCC
CGAGTTTTGGCTGCTCACAAAGCAACGAAACCCCATGGATTATCCTGTTGAGGATGCTTTC
TGTAAACCTCAACTCGTGAAGAAGAGCTATGGTGTGGAGAACAGAGCCTACTCTCAAGAGG
AAATCACTCAAGGTTTTGAAGAGACAGGGGACACGCTCTATGCCCCCTATTCCACACATTT
TCAGCTGCAGAACCAGCCTCCCCAAAAGGAATTCTCCATCCCACGGGCCCACGCTTGGCCG
AGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGGCAGCTAACTCTGTCCTGAAGAGTGGG
ACAAATGCAGCCGGGCGGCAGATCTAGCGGGAGCTCAAAGGGATGTGGGCGAAATCTTGAG
TCTTCTGAGAAAACTGTACAAGACACTACGGGAACAGTTTGCCTCCCTCCCAGCCTCAACC
ACAATTCTTCCATGCTGGGGCTGATGTGGGCTAGTAAGACTCCAGTTCTTAGAGGCGCTGT
AGTATTTTTTTTTTTTTGTCTCATCCTTTAGATACTTCTTTTAAGTGGGAGTCTCAG

FIGURE 65

TCATTCCATTAAACATATTTCTAAATAATAGTAAGTGGTACTAACAAAATAAATAATAATT
TAATAGCCTTAGAAATAAATGACTGTATACTTATACAGGTTGAAAAAAACTCGGTAGGAAT
AAGTTACCTTTTTGTTTACTAATGTTGGTTTCAAAAATACTCAGATTCATTTTAGTTGGCT
GACATCTGGAAGTAGTTAACAACTAACCAGTGGACTTCAACAATCATTTGCTCCCAGGCTT
CCCCCATCATCACCCTCACCACATATCCTGCTAATATCCAACAACAAACAAATATTTAATA
TTGAAATAGCCCATTGCCTGAGAATGAACACAAGCTAAAATACATGCAAGGGTACTTAATG
GAAGCCAAACCATGTTCTATACCTAAGGAGAAAACATGGACATGTAGAATGCTTTTATTCA
TGTATTCAAAATCAGAACAAATCAGTGTATATCACTAGAACATCAGATGGAGGATAACACA
AGAAGTGATACAGATCAGGGTTCACTTCTCTTACCCTCTCTCTGTTAGGACCACATTCCTA
TTTTAGCCAAATGTTTCTGGTACGGGCCATCTTTTCACCATAAATGGCATTATGTTTCAAA
TGGCTAAAAGCTTATTATGGGTATGCTCAAAGGAGTAAAACCCAATTCACAGAGATGTGGC
TTTTCTAAAGAACCAAATTAAGGGAATACTTAGTATCTACAACTAAAGTATTAGTCATTCT
GAGGATTATTTGTTGCTTTAAGTATTAGCTCCCTACTACATCCATAACACCCTACTTACTA
AATTTAATTACACACAACTTTTCAAGAATTCGTATTTTATTTGAAGGGAGGTACCTGTCTA
CTTTATCTACAATAAAAACAAAGGGTATTGGCTTTCTCTAATCCATGCAAACTACAAATTC
CATCGGGAGTCCTACATCACTAACAGTGGTATGAACAAAACTAAGAAAGTACTTCCTATCC
ACATGTGAATGTTTTTAAAAAAGTTTTTGCCATAAAACCTAAGTGTAATTTAGCATACCACA
GTGCTCTGAAGATGGGTCATTGACGATGTACCATTTGTATATAGGTAATACACATGTAAAT
CACTAATTGTTAATTAAATATAAGAAGTAGGTTTTTATCTTTTTANNAAAAAAACAAAAAA
GGTGACTCCCTTTCTCATTCTGTTCTGTTGTATTGCGTCCAAAAGTTGTGTGTAATTTTTT
TAAGGTCCAAGAAATGTAAAGAAATTTGTTGGAATACCTGAATTTCTGTAAAAAAATAAAA
ATAAAAATAAGATTTTGTTACTT

FIGURE 66

GGGAGGCTGTGCCAGGCGAGCCGGAGGGGTGCTCCGCGCTCCCCCGCCCTCCTTCCGGGAG
CGAGGATGCAGACTCTGAAACTGGTGCTGCTGGGCTGAGGCGGAGGCAGGGGAGTTGCAGC
GCGCGAGGCTCCGTGAGTGTGTCTCCTGCGCGCTGAGAGGCGGGGGGAGGCGGAGGACCAG
GAGGAGGAGGAGGAGGAGGAGGAGGGGGAGAATGCCCGGAGCCGCCGCCGCTGCCGCCGCC
GCCGCCGCCGCGATGCTCCCGGCTCAGGAGGCTGCCAAGCTGTACCACACCAACTATGTGC
GGAACTCGCGGGCCATCGGCGTGCTGTGGGCCATCTTCACCATCTGCTTTGCCATCGTCAA
CGTGGTGTGCTTCATCCAGCCCTACTGGATAGGCGACGGCGTGGACACCCCGCAAGCCGGC
TATTTCGGGCTCTTCCACTACTGCATCGGCAACGGCTTCTCCCGGGAGCTGACCTGCAGGG
GCAGCTTCACGGACTTCTCCACGCTGCCCTCGGGCGCCTTCAAAGCCGCCTCCTTCTTTAT
CGGCCTCTCCATGATGCTCATCATTGCCTGCATCATTTGCTTTACCCTCTTCTTCTTCTGC
AACACGGCCACTGTGTACAAGATATGTGCCTGGATGCAGCTCACCTCCGCTGCCTGCCTTG
TGCTTGGCTGTATGATTTTCCCTGATGGCTGGGACTCAGATGAAGTAAAACGGATGTGTGG
AGAAAAGACAGACAAGTACACTCTTGGGGCTTGCTCAGTCCGCTGGGCATACATCCTGGCT
ATTATTGGAATTTTGGATGCCCTGATCCTCTCATTTCTAGCATTTGTGCTTGGTAATCGAC
AAGACAGCTTGATGGCAGAGGAACTGAAGGCAGAAAACAAAGTTCTGCTAAGCCAATATTC
TCTAGAATGAGCACAAAACAAATCGAATAACAGCTAAACAAATCGAATAACAGCTAAACGA
ATCGAATAACAGCTTTTGTACATCAACATCAAGAAGGAATACGCCTGAGAGAGATCAGAGT
ATATAGATGAATATGAACAAGAATGGAACATTCACTTGTCAACGCACTTTCTAAATCTAGA
TCAGCAGAGATGGGAGTGATTTTCTGGAAAGAGATGTGATCATGGATTAAACACCAGCTCA
TTGGAAACTCATTGGATGAGATCAGAAAACGTTCATGAAAAATCATATTCAGGAAATAAGG
AAGAGGAATATAAATGCTCTAGAGTTAACATGTAAAATATATACGTACTGAGGTTTGTAAA
CTGTCCTTTTTAAATCAAACTGAAAACAAAAAGCTTTAACCTTTCAACAGAATTTTTAAAA
AGGCAGTTAGTTCTAAATTATTCCTATCTCAATAGCCAAGAGGCTGATCAAGCGTCATTTA
TTGAGGAAGCATCTTAGAAAATGCCTCTGAATGTTTTCATAGGAGCCGTGACCTTTGGTTC
TTCATCTCTACCATTCATTTACTTCACTGTGTAATTAGTTACAACCACTCAGTTATTAAGA
GACGTAACGCTTCAAACTTTTTACCAAGTCTGTGTTCTGTTTAATCTGTCCATACAAGTTA
TTACTGAGAAAGTGTTTATGCCATATACTATTACTCCATCAAGCTGTATATTACAGGAAGT
ACATCTTTACATCATAGGTTCCCAAGCAACATAGATTTCCCTATCTTTCAGGAAACAGCAT
CAAGGAACTCTGAAAAATATAGAAAAAGTTCATTTTCACCTTGGAAGCTCACGTGTAATAT
TATAGGCTACTATCAAATAAACACTTTTTTTCTAATTCTCCCTAGTATATGCATAGGAATT
TAATATACTTTATAAATAAGTATCTAAAATGTCTCCTACTTTTTTCCTATTTCTTTGCCAT
ACATGTTATCAGAAATCCATGTCTTCTATTTCCCTTACTGATGGGCACTCATTTTTATTTT
TTTAAAAATCATTCCATTAAACATATTTCTAAATAATAGTAAGTGGTACTAACAAAATAAA
TAATAATTTAATAGCCTTAGAAATAAATGACTGTATACTTATACAGGTTGAAAAAAACTCG

AGTTGGCTGACATCTGGAAGTAGTTAACAACTAACCAGTGGACTTCAACAATCATTTGCTC
CCAGGCTTCCCCCATCATCACCCTCACCACATATCCTGCTAATATCCARCAACAAACAAAT
ATTTAATATTGAAATAGCCCATTGCCTGAGAATGAACACAAGCTAAAATACATGCAAGGGT
ACTTAATGGAAGCCAAACCATGTTCTATACCTAAGGAGAAAACATGGACATGTAGAATGCT
TTTATTCATGTATTCAAAATCAGAACAAATCAGTGTATATCACTAGAACATCAGATGGAGG
ATAACACAAGAAGTGATACAGATCAGGGTTCACTTCTCTTACCCTCTCTCTGTTAGGACCA
CATTCCTATTTTAGCCAAATGTTTCTGGTACGGGCCATCTTTTCACCATAAATGGCATTAT
GTTTCAAATGGCTAAAAGCATATTATGGGTATGCTCAAAGGAGTAAAACCCAATTCACAGA
GATGTGGCTTTTCTAAAGAACCAAATTAAGGGAATACTTAGTATCTACAACTAAAGTATTA
GTCATTCTGAGGATTATTTGTTGCTTTAAGTATTAGCTCCCTACTACATCCATAACACCCT
ACTTACTAAATTTAATTACACACAACTTTTCAAGAATTCGTATTTTATTTGAAGGGAGGTA
CCTGTCTACTTTATCTACAATAAAAACAAAGGGTATTGGCTTTCTCTAATCCATGCAAACT
ACAAATTCCATCGGGAGTCCTACATCACTAACAGTGGTATGAACAAAACTAAGAAAGTACT
TCCTATCCACATGTGAATGTTTTAAAAAAGTTTTTGCCATAAAACCTAAGTGTAATTTAGC
ATACCACAGTGCTCTGAAGATGGGTCATTGACGATGTACCATTTGTATATAGGTAATACAC
ATGTAAATCACTAATTGTTAATTAAATATAAGAAGTAGGTTTTTATCTTTTTAAAAAAAAA
ACAAAAAAAGGTGACTCCCTTTCTCATTCTGTTCTGTTGTATTGCGTCCAAAAGTTGTGTGT
AATTTTTTTTAAGGTCCAAGAAATGTAAAGAAATTTGTTGGAATACCTGAATTTCTGTAAAA
AAATAAAAATAAAAATAAGATTTTGTTACTTAAAA

FIGURE 67

GGGAGGCTGTGCCAGGCGAGCCGGAGGGGTGCTCCGCGCTCCCCCGCCCTCCTTCCGGGAG
CGAGGATGCAGACTCTGAAACTGGTGCTGCTGGGCTGAGGCGGAGGCAGGGGAGTTGCAGC
GCGCGAGGCTCCGTGAGTGTGTCTCCTGCGCGCTGAGAGGCGGGGGGAGGCGGAGGACCAG
GAGGAGGAGGAGGAGGAGGAGGAGGGGGAGAATGCCCGGAGCCGCCGCCGCTGCCGCCGCC
GCCGCCGCCGCGATGCTCCCGGCTCAGGAGGCTGCCAAGCTGTACCACACCAACTATGTGC
GGAACTCGCGGGCCATCGGCGTGCTGTGGGCCATCTTCACCATCTGCTTTGCCATCGTCAA
CGTGGTGTGCTTCATCCAGCCCTACTGGATAGGCGACGGCGTGGACACCCCGCAAGCCGGC
TATTTCGGGCTCTTCCACTACTGCATCGGCAACGGCTTCTCCCGGGAGCTGACCTGCAGGG
GCAGCTTCACGGACTTCTCCACGCTGCCCTCGGGCGCCTTCAAAGCCGCCTCCTTCTTTAT
CGGCCTCTCCATGATGCTCATCATTGCCTGCATCATTTGCTTTACCCTCTTCTTCTTCTGC
AACACGGCCACTGTGTACAAGATATGTGCCTGGATGCAGCTCACCTCCGCTGCCTGCCTTG
TGCTTGGCTGTATGATTTTCCCTGATGGCTGGGACTCAGATGAAGTAAAACGGATGTGTGG
AGAAAAGACAGACAAGTACACTCTTGGGGCTTGCTCAGTCCGCTGGGCATACATCCTGGCT
ATTATTGGAATTTTGGATGCCCTGATCCTCTCATTTCTAGCATTTGTGCTTGGTAATCGAC
AAGACAGCTTGATGGCAGAGGAACTGAAGGCAGAAAACAAAGTTCTGCTAAGCCAATATTC
TCTAGAATGAGCACAAAACAAATCGAATAACAGCTAAACAAATCGAATAACAGCTAAACGA
ATCGAATAACAGCTTTTGTACATCAACATCAAGAAGGAATACGCCTGAGAGAGATCAGAGT
ATATAGATGAATATGAACAAGAATGGAACATTCACTTGTCAACGCACTTTCTAAATCTAGA
TCAGCAGAGATGGGAGTGATTTTCTGGAAAGAGATGTGATCATGGATTAAACACCAGCTCA
TTGGAAACTCATTGGATGAGATCAGAAAACGTTCATGAAAAATCATATTCAGGAAATAAGG
AAGAGGAATATAAATGCTCTAGAGTTAACATGTAAAATATATACGTACTGAGGTTTGTAAA
CTGTCCTTTTTAAATCAAACTGAAAACAAAAAGCTTTAACCTTTCAACAGAATTTTTAAAA
AGGCAGTTAGTTCTAAATTATTCCTATCTCAATAGCCAAGAGGCTGATCAAGCGTCATTTA
TTGAGGAAGCATCTTAGAAAATGCCTCTGAATGTTTTCATAGGAGCCGTGACCTTTGGTTC
TTCATCTCTACCATTCATTTACTTCACTGTGTAATTAGTTACAACCACTCAGTTATTAAGA
GACGTAACGCTTCAAACTTTTTACCAAGTCTGTGTTCTGTTTAATCTGTCCATACAAGTTA
TTACTGAGAAAGTGTTTATGCCATATACTATTACTCCATCAAGCTGTATATTACAGGAAGT
ACATCTTTACATCATAGGTTCCCAAGCAACATAGATTTCCCTATCTTTCAGGAAACAGCAT
CAAGGAACTCTGAAAAATATAGAAAAAGTTCATTTTCACCTTGGAAGCTCACGTGTAATAT
TATAGGCTACTATCAAATAAACACTTTTTTTCTAATTCTCCCTAGTATATGCATAGGAATT
TAATATACTTTATAAATAAGTATCTAAAATGTCTCCTACTTTTTTCCTATTTCTTTGCCAT
ACATGTTATCAGAAATCCATGTCTTCTATTTCCCTTACTGATGGGCACTCATTTTTATTTT
TTTAAAAATCATTCCATTAAACATATTTCTAAATAATAGTAAGTGGTACTAACAAAATAAA
TAATAATTTAATAGCCTTAGAAATAAATGACTGTATACTTATACAGGTTGAAAAAAACTCG
GTAGGAATAAGTTACCTTTTTTGTTTACTAATGTTGGTTTCAAAAATACTCAGATTCATTTT

AGTTGGCTGACATCTGGAAGTAGTTAACAACTAACCAGTGGACTTCAACAATCATTTGCTC
CCAGGCTTCCCCCATCATCACCCTCACCACATATCCTGCTAATATCCAACAACAAACAAAT
ATTTAATATTGAAATAGCCCATTGCCTGAGAATGAACACAAGCTAAAATACATGCAAGGGT
ACTTAATGGAAGCCAAACCATGTTCTATACCTAAGGAGAAAACATGGACATGTAGAATGCT
TTTATTCATGTATTCAAAATCAGAACAAATCAGTGTATATCACTAGAACATCAGATGGAGG
ATAACACAAGAAGTGATACAGATCAGGGTTCACTTCTCTTACCCTCTCTCTGTTAGGACCA
CATTCCTATTTTAGCCAAATGTTTCTGGTACGGGCCATCTTTTCACCATAAATGGCATTAT
GTTTCAAATGGCTAAAAGCATATTATGGGTATGCTCAAAGGAGTAAAACCCAATTCACAGA
GATGTGGCTTTTCTAAAGAACCAAATTAAGGGAATACTTAGTATCTACAACTAAAGTATTA
GTCATTCTGAGGATTATTTGTTGCTTTAAGTATTAGCTCCCTACTACATCCATAACACCCT
ACTTACTAAATTTAATTACACACAACTTTTCAAGAATTCGTATTTTATTTGAAGGGAGGTA
CCTGTCTACTTTATCTACAATAAAAACAAAGGGTATTGGCTTTCTCTAATCCATGCAAACT
ACAAATTCCATCGGGAGTCCTACATCACTAACAGTGGTATGAACAAAACTAAGAAAGTACT
TCCTATCCACATGTGAATGTTTTAAAAAAGTTTTTGCCATAAAACCTAAGTGTAATTTAGC
ATACCACAGTGCTCTGAAGATGGGTCATTGACGATGTACCATTTGTATATAGGTAATACAC
ATGTAAATCACTAATTGTTAATTAAATATAAGAAGTAGGTTTTTATCTTTTTAAAAAAAAA
ACAAAAAAGGTGACTCCCTTTCTCATTCTGTTCTGTTGTATTGCGTCCAAAAGTTGTGTGT
AATTTTTTTAAGGTCCAAGAAATGTAAAGAAATTTGTTGGAATACCTGAATTTCTGTAAAA
AAATAAAAATAAAAATAAGATTTTGTTACTTAAAA

FIGURE 68

GGGAGGCTGTGCCAGGCGAGCCGGAGGGGTGCTCCGCGCTCCCCCGCCCTCCTTCCGGGAG
CGAGGATGCAGACTCTGAAACTGGTGCTGCTGGGCTGAGGCGGAGGCAGGGGAGTTGCAGC
GCGCGAGGCTCCGTGAGTGTGTCTCCTGCGCGCTGAGAGGCGGGGGGAGGCGGAGGACCAG
GAGGAGGAGGAGGAGGAGGAGGAGGGGGGAGAATGCCCGGAGCCGCCGCCGCTGCCGCCGCC
GCCGCCGCCGCGATGCTCCCGGCTCAGGAGGCTGCCAAGCTGTACCACACCAACTATGTGC
GGAACTCGCGGGCCATCGGCGTGCTGTGGGCCATCTTCACCATCTGCTTTGCCATCGTCAA
CGTGGTGTGCTTCATCCAGCCCTACTGGATAGGCGACGGCGTGGACACCCCGCAAGCCGGC
TATTTCGGGCTCTTCCACTACTGCATCGGCAACGGCTTCTCCCGGGAGCTGACCTGCAGGG
GCAGCTTCACGGACTTCTCCACGCTGCCCTCGGGCGCCTTCAAAGCCGCCTCCTTCTTTAT
CGGCCTCTCCATGATGCTCATCATTGCCTGCATCATTTGCTTTACCCTCTTCTTCTTCTGC
AACACGGCCACTGTGTACAAGATATGTGCCTGGATGCAGCTCACCTCCGCTGCCTGCCTTG
TGCTTGGCTGTATGATTTTCCCTGATGGCTGGGACTCAGATGAAGTAAAACGGATGTGTGG
AGAAAAGACAGACAAGTACACTCTTGGGGCTTGCTCAGTCCGCTGGGCATACATCCTGGCT
ATTATTGGAATTTTGGATGCCCTGATCCTCTCATTTCTAGCATTTGTGCTTGGTAATCGAC
AAGACAGCTTGATGGCAGAGGAACTGAAGGCAGAAAACAAAGTTCTGCTAAGCCAATATTC
TCTAGAATGAGCACAAAACAAATCGAATAACAGCTAAACAAATCGAATAACAGCTAAACGA
ATCGAATAACAGCTTTTGTACATCAACATCAAGAAGGAATACGCCTGAGAGAGATCAGAGT
ATATAGATGAATATGAACAAGAATGGAACATTCACTTGTCAACGCACTTTCTAAATCTAGA
TCAGCAGAGATGGGAGTGATTTTCTGGAAAGAGATGTGATCATGGATTAAACACCAGCTCA
TTGGAAACTCATTGGATGAGATCAGAAAACGTTCATGAAAAATCATATTCAGGAAATAAGG
AAGAGGAATATAAATGCTCTAGAGTTAACATGTAAAATATATACGTACTGAGGTTTGTAAA
CTGTCCTTTTTAAATCAAACTGAAAACAAAAAGCTTTAACCTTTCAACAGAATTTTTAAAA
AGGCAGTTAGTTCTAAATTATTCCTATCTCAATAGCCAAGAGGCTGATCAAGCGTCATTTA
TTGAGGAAGCATCTTAGAAAATGCCTCTGAATGTTTTCATAGGAGCCGTGACCTTTGGTTC
TTCATCTCTACCATTCATTTACTTCACTGTGTAATTAGTTACAACCACTCAGTTATTAAGA
GACGTAACGCTTCAAACTTTTTACCAAGTCTGTGTTCTGTTTAATCTGTCCATACAAGTTA
TTACTGAGAAAGTGTTTATGCCATATACTATTACTCCATCAAGCTGTATATTACAGGAAGT
ACATCTTTACATCATAGGTTCCCAAGCAACATAGATTTCCCTATCTTTCAGGAAACAGCAT
CAAGGAACTCTGAAAAATATAGAAAAAGTTCATTTTCACCTTGGAAGCTCACGTGTAATAT
TATAGGCTACTATCAAATAAACACTTTTTTTCTAATTCTCCCTAGTATATGCATAGGAATT
TAATATACTTTATAAATAAGTATCTAAAATGTCTCCTACTTTTTTCCTATTTCTTTGCCAT
ACATGTTATCAGAAATCCATGTCTTCTATTTCCCTTACTGATGGGCACTCATTTTTATTTT
TTTAAAAATCATTCCATTAAACATATTTCTAAATAATAGTAAGTGGTACTAACAAAATAAA
TAATAATTTAATAGCCTTAGAAATAAATGACTGTATACTTATACAGGTTGAAAAAAACTCG
[illegible]TTTCAAAAATACTCAGATTCATTTT

AGTTGGCTGACATCTGGAAGTAGTTAACAACTAACCAGTGGACTTCAACAATCATTTGCTC
CCAGGCTTCCCCCATCATCACCCTCACCACATATCCTGCTAATATCCAACAACAAACAAAT
ATTTAATATTGAAATAGCCCATTGCCTGAGAATGAACACAAGCTAAAATACATGCAAGGGT
ACTTAATGGAAGCCAAACCATGTTCTATACCTAAGGAGAAAACATGGACATGTAGAATGCT
TTTATTCATGTATTCAAAATCAGAACAAATCAGTGTATATCACTAGAACATCAGATGGAGG
ATAACACAAGAAGTGATACAGATCAGGGTTCACTTCTCTTACCCTCTCTCTGTTAGGACCA
CATTCCTATTTTAGCCAAATGTTTCTGGTACGGGCCATCTTTTCACCATAAATGGCATTAT
GTTTCAAATGGCTAAAAGCATATTATGGGTATGCTCAAAGGAGTAAAACCCAATTCACAGA
GATGTGGCTTTTCTAAAGAACCAAATTAAGGGAATACTTAGTATCTACAACTAAAGTATTA
GTCATTCTGAGGATTATTTGTTGCTTTAAGTATTAGCTCCCTACTACATCCATAACACCCT
ACTTACTAAATTTAATTACACACAACTTTTCAAGAATTCGTATTTTATTTGAAGGGAGGTA
CCTGTCTACTTTATCTACAATAAAAACAAAGGGTATTGGCTTTCTCTAATCCATGCAAACT
ACAAATTCCATCGGGAGTCCTACATCACTAACAGTGGTATGAACAAAACTAAGAAAGTACT
TCCTATCCACATGTGAATGTTTTAAAAAAGTTTTTGCCATAAAACCTAAGTGTAATTTAGC
ATACCACAGTGCTCTGAAGATGGGTCATTGACGATGTACCATTTGTATATAGGTAATACAC
ATGTAAATCACTAATTGTTAATTAAATATAAGAAGTAGGTTTTTATCTTTTTAAAAAAAAA
ACAAAAAAGGTGACTCCCTTTCTCATTCTGTTCTGTTGTATTGCGTCCAAAAGTTGTGTGT
AATTTTTTTAAGGTCCAAGAAATGTAAAGAAATTTGTTGGAATACCTGAATTTCTGTAAAA
AAATAAAAATAAAAATAAGATTTTGTTACTTAAAA

FIGURE 69

CAGGCCTCTGAGGCTCCCTTGCCGAGGGCCCCGAGCTGCAGGGACAGTGAGCAGTGAGTCC
CTTGGGCATCCCGCTCCTGGGCAGGTCACCAATAGGTCCCCGCAGTTCCCAATGGAACTGT
TCCAGTCCTCCCCGAGGCCTCCACTTCAACCTGTCTGTGTCTGCCCAGGCCTGGAGTTGTG
TGACCCTCCCCACCGCCTGGCCTTCTCCATGGGGGCTGGCCTTTTCTCGGTGGTGGGCACC
CTGCTGCTGCCCGGCCTGGCTGCGCTTGTGCAGGACTGGCGTCTTCTGCAGGGGCTGGGTG
CCCTGATGAGTGGACTCTTGCTGCTCTTTTGGGGGTTCCCGGCCCTGTTCCCCGAGTCTCC
CTGCTGGCTGCTGGCCACAGGTCAGGTAGCTCGAGCCAGGAAGATCCTGTGGCGCTTTGCA
GAAGCCAGTGGCGTGGGCCCCGGGGACAGTTCCTTGGAGGAGAACTCCCTGGCTACAGAGC
TGACCATGCTGTCTGCACGGAGCCCCCAGCCCCGGTACCACTCCCCACTGGGGCTTCTGCG
TACCCGAGTCACCTGGAGAAACGGGCTTATCTTGGGCTTCAGCTCGCTGGTTGGTGGAGGC
ATCAGAGCTAGCTTCCGCCGCAGCCTGGCACCTCAGGTGCCGACCTTCTACCTGCCCTACT
TCCTGGAGGCCGGCCTGGAGGCGGCAGCCTTGGTCTTCCTGCTCCTGACGGCAGATTGCTG
TGGACGCCGCCCCGTGCTGCTGCTGGGCACCATGGTCACAGGCCTGGCATCCCTGCTGCTC
CTCGCTGGGGCCCAGTATCTGCCAGGCTGGACTGTGCTGTTCCTCTCTGTCCTGGGGCTCC
TGGCCTCCCGGGCTGTGTCCGCACTCAGCAGCCTCTTCGCGGCCGAGGTCTTCCCCACGGT
GATCAGGGGGGCCGGGCTGGGCCTGGTGCTGGGGGCCGGGTTCCTGGGCCAGGCAGCCGGC
CCCCTGGACACCCTGCACGGCCGGCAGGGCTTCTTCCTGCAACAAGTCGTCTTCGCCTCCC
TTGCTGTCCTTGCCCTGCTGTGTGTCCTGCTGCTGCCTGAGAGCCGAAGCCGGGGGCTGCC
CCAGTCACTGCAGGACGCCGACCGCCTGCGCCGCTCCCCACTCCTGCGGGGCCGCCCCCGC
CAGGACCACCTGCCTCTGCTGCCGCCCTCCAACTCCTACTGGGCCGGCCACACCCCCGAGC
AGCACTAGTCCTGCCTGGTGGCCCTGGGAGCCAGGATGGGACCAAAGTCAAGGCCTGGGGC
ATGGCTGAGTACCCCAGACGTCTGGTCCAGGGCAGACACATTCCTCTCAGAAGCCCGTGTC
TCAGTGCAGGTGGAGCCGTGGGGACAGCGTGAAGGTGTCTCCAGCCAGGCCCCAGGCACTG
GGAGGCCCTGGGTCTCCCCCCAGCCACACCCAGTAGGTGTGGAGGATAAAGGCTTCTGTGG
AACTTGGCGTCCTCCCTGGTTCTTCCCTCAGGCCACACCTGGGGCCTGGAAGGGTCTCACC
TGTGCCTGCTCCCCACCTCTGGGCTCACCAGCAAGGAGCCAGGGTGTCAGCGGCTGCTCCA
TAGCCCCTGTCCAGGTAGTGGCAGCCCCCGGCCGCCTCCTGCCCCAGCCAAGCAGGGTGGT
ATGGGGGCCTGGCCCTCAGGTCCCTGCCAGGCCAGGCCTCACTGGCCCCATCCCCCTCATG
GTCAGTGTGGCAGTGGGCGCCCGTCCTGCATGAAAGAGACTAGCACG

FIGURE 70

AGCGGGGCGGGTCCCCGAAGCCCAGTACCAGACCTCCTGTGCGAAGAGTAACCTGCTCTTC
GGACTGATAACTGGATGGAATGGCCTTTAGAAAGGGAGCATCTTCAGAATCTGCAGCTGGG
ATCAGATTGTTCTTTTGTTGAATCAACTCAGCTCTGCCGAGCCAGCATCTTGCTAAGGGTT
ACAAAATAAATGCCATTTGAACAGTGCCATCTGTCATGGAAAAACGTGAGACGTTTGTACA
ATGCCGTGTCTAAGGAGCTGGTTGGAGAGTTTTTGCAATTTGTTCAACTTGATAAAGAGGC
CTCTGATCCTTTCAGCCTAAATGAATTACTAGATGAATTATCAAGGAAACAGAAAGAAGAA
TTATGGCAAAGGCTGAAGAATTTATTGACAGATGTGTTGTTAGAAAGCCCAGTGGATGGGT
GGCAGGTAGTGGAAGCCCAGGGTGAAGACAATATGGAAACCGAACATGGCTCAAAAATGAG
AAAAAGCATAGAAATAATTTATGCAATTACATCTGTGATTCTTGCTTCTGTGTCTGTAATA
AATGAAAGTGAGAACTACGAAGCCCTACTGGAATGTGTTATTATATTAAATGGTATTTTAT
ATGCATTACCTGAGTCTGAACGAAAACTACAGAGTTCTATTCAGGATTTGTGTGTTACCTG
GTGGGAGAAAGGCCTGCCTGCCAAGGAAGACACAGGAAAGACAGCCTTTGTCATGTTACTA
AGGAGGAGTCTGGAGACTAAGACAGGTGCAGACGTATGTCGGCTTTGGCGTATCCATCAAG
CTTTATATTGCTTTGATTATGATTTGGAGGAAAGTGGAGAAATTAAAGATATGTTACTTGA
GTGCTTCATAAATATTAATTATATCAAGAAAGAAGAGGGAAGAAGATTTCTTAGTTGTCTC
TTCAACTGGAATATCAACTTCATCAAAATGATCCACGGGACCATTAAAAACCAGTTACAGG
GATTACAAAAGTCTTTGATGGTATACATTGCAGAAATTTATTTCAGAGCTTGGAAAAAGGC
TTCAGGGAAAATACTGGAGGCGATTGAAAATGATTGCATCCAGGACTTCATGTTCCACGGG
ATACACCTTCCGAGGAGGTCTCCAGTGCATTCCAAAGTGCGGGAGGTTCTGAGTTACTTTC
ACCATCAAAAGAAAGTTCGGCAGGGAGTGGAAGAGATGCTTTATAGATTATATAAGCCCAT
CCTTTGGAGAGGATTAAAGGCCAGAAACTCTGAAGTTCGATCAAATGCTGCATTGTTGTTT
GTTGAAGCATTTCCTATTAGGGATCCAAACCTTCATGCTATTGAAATGGATAGTRAAATCC
AGAAACAGTTTGAAGAGCTCTATAGCCTTTTAGAAGATCCTTACCCGATGGTCCGTTCCAC
AGGGATCCTTGGTGTTTGTAAAATAACTTCTAAGTACTGGGAAATGATGCCCCCGACCATT
CTTATTGACCTCCTGAAGAAGGTGACTGGGGAACTGGCATTTGACACGAGCTCAGCTGATG
TTCGTTGTTCTGTCTTTAAGTGTCTGCCAATGATTTTGGACAACAAACTGAGCCACCCATT
GTTAGAGCAGCTCCTTCCAGCTCTCAGATACAGTCTCCACGACAATTCGGAGAAAGTGAGG
GTAGCTTTTGTGGACATGCTGTTGAAGATCAAAGCTGTGAGGGCTGCTAAGTTTTGGAAAA
TATGTCCCATGGAGCACATTCTGGTTCGTCTGGAAACTGATTCTCGACCTGTGTCTCGGCG
CCTGGTGAGCCTCATCTTTAATTCTTTCCTGCCTGTGAATCAGCCGGAGGAGGTCTGGTGC
GAGCGCTGTGTCACCCTGGTGCAGATGAACCACGCCGCTGCCAGGAGGTTCTATCAGTACG
CCCACGAACACACCGCCTGCACCAACATAGCAAAGCTGATTCACGTTATTCGTCATTGCTT
AAATGCCTGTATCCAGAGGGCAGTGAGAGAGCCTCCAGAGGACGAGGAGGAAGAGGACGGA
AGGGAGAAGGAGAATGTGACTGTTCTGGACAAAACACTGTCAGTAAACGATGTTGCATGCA
TGGCAGGTTTACTAGAAATCATTGTGATTCTCTGGAAAAGTATTGACAGAAGTATGGAAAA

TAATAAAGAGGCCAAACTTTACACGATTAACAAGTTTGCCTCTGTGCTTCCAGAGTATCTG
AAAGTATTTAAGGATGATCGCTGCAAGATCCCTTTATTCATGCTAATGTCCTTTATGCCGG
CCTCTGCTGTCCCCCCATTCAGCTGTGGTGTGATTTCCACGCTGAGAAGCCGGGAGGAGGG
CGCTGTGGACAAGAGCTACTGCACTTTGTTGGATTGCCTCTGCTCCTGGGGGCAGGTGGGG
CACATTCTGGAGCTTGTTGACAACTGGCTGCCCACAGAGCATGCCCAGGCCAAGAGCAACA
CAGCTTCTAAAGGTAGGGTGCAGATCCATGACACACGCCCAGTCAAACCTGAATTGGCATT
GGTCTACATTGAGTATCTGCTGACTCATCCAAAGAACCGCGAGTGCTTGCTCTCTGCTCCT
CGGAAGAAACTTAACCATCTTTTGAAAGCCCTTGAAACGTCAAAGGCAGATCTGGAGTCAC
TTCTGCAGACACCGGGTGGGAAGCCTCGTGGCTTCAGTGAAGCAGCTGCCCCGCGAGCCTT
TGGTCTCCACTGTCGCCTGAGCATCCATCTTCAGCACAAGTTCTGCTCAGAAGGAAAGGTG
TATTTGTCCATGTTGGAAGACACTGGCTTTTGGTTAGAAAGCAAAATTTTATCTTTTATTC
AAGATCAAGAAGAAGACTACCTGAAGCTTCATAGGGTCATTTATCAGCAAATTATCCAGAC
CTACCTGACTGTGTGTAAAGATGTTGTTATGGTAGGCCTTGGTGACCATCAGTTTCAGATG
CAACTCTTACAGCGGAGTCTTGGAATCATGCAAACAGTGAAGGGATTTTTTTATGTTTCAT
TACTTCTTGACATTCTGAAAGAGATAACTGGAAGTTCCTTGATTCAGAAAACAGATTCAGA
TGAAGAAGTTGCAATGCTGTTGGACACAGTCCAGAAAGTATTTCAGAAAATGTTGGAATGT
ATTGCACGGAGCTTCAGGAAGCAGCCGGAAGAAGGCCTGCGGCTGCTTTATTCTGTTCAGA
GGCCTCTTCATGAGTTCATTACTGCTGTTCAGTCTCGGCACACAGACACCCCTGTGCACCG
GGGTGTACTTTCTACTCTGATCGCTGGGCCTGTGGTTGAGATAAGTCACCAGCTACGGAAG
GTTTCTGACGTAGAAGAGCTTACCCCTCCAGAGCATCTTTCTGATCTTCCACCATTTTCAA
GGTGTTTAATAGGAATAATAATAAAGTCTTCGAATGTGGTCAGGTCATTTTTGGATGAATT
AAAGGCATGTGTGGCTTCTAATGATATTGAAGGCATTGTGTGCCTCACGGCTGCTGTGCAT
ATTATCCTGGTTATTAATGCAGGTAAACATAAAAGCTCAAAAGTGAGGGAGGTTGCAGCCA
CTGTTCACAGAAAACTAAAGACATTCATGGAAATTACTTTGGAAGAGGATAGCATTGAAAG
ATTTCTCTATGAATCATCATCAAGAACTCTGGGAGAACTTTTGAATTCATAACCAAGCCAA
CATCTCCAGACATGTAAAAATAGGGAAAAGTGATTCAAATTGAAATGCCTGTGTATTTTCC
TATTGTTTTTAATGTTAATAACCCATATAATAGGGAAAGGGTGGGATTTTTTTGTGGGAAT
GTGGGAAGGTGGGGGTTATGGAGGAGATAACTCAAAACTTCTTCAATTTTGCCTAGTGCCT
GCGTAAATAATATATTTAATATAAAGGACTCCAGGTATGAATGGTGTAGAAATCCATGATT
CCAAGAAAAAACACTTTTCTAGCAAACCTGGTTGTTTTTAAAATGACTTTTATATATGTAA
TATTGCTTGGAAACTATGAGTAATAAAGCAATGACAACAAMAAAAAAAAAAAAAAA

FIGURE 71

GCCCTTCAGAATCTGCAGCTGGGATCAGATTGTTCTTTTGTTGAATCAACTCAGCTCTGCC
GAGCCAGCATCTTGCTAAGGGTTACAAAATAAATGCCATTTGAACAGTGCCATCTGTCATG
GAAAAACGTGAGACGTTTGTACAAGCCGTGTCTAAGGAGCTGGTTGGAGAGTTTTTGCAAT
TTGTTCAACTTGATAAAGAGGCCTCTGATCCTTTCAGCCTAAATGAATTACTAGATGAATT
ATCAAGGAAACAGAAAGAAGAATTATGGCAAAGGCTGAAGAATTTATTGACAGATGTGTTG
TTAGAAAGCCCAGTGGATGGGTGGCAGGTAGTGGAAGCCCAGGGTGAAGACAATATGGAAA
CCGAACATGGCTCAAAAATGAGAAAAAGCATAGAAATAATTTATGCAATTACATCTGTGAT
TCTTGCTTCTGTGTCTGTAATAAATGAAAGTGAGAACTACGAAGCCCTACTGGAATGTGTT
ATTATATTAAATGGTATTTTATATGCATTACCTGAGTCTGAACGAAAACTACAGAGTTCTA
TTCAGGATTTGTGTGTTACCTGGTGGGAGAAAGGCCTGCCTGCCAAGGAAGACACAGGAAA
GACAGCCTTTGTCATGTTACTAAGGAGGAGTCTGGAGACTAAGACAGGTGCAGACGTATGT
CGGCTTTGGCGTATCCATCAAGCTTTATATTGCTTTGATTATGATTTGGAGGAAAGTGGAG
AAATTAAAGATATGTTACTTGAGTGCTTCATAAATATTAATTATATCAAGAAAGAAGAGGG
AAGAAGATTTCTTAGTTGTCTCTTCAACTGGAATATCAACTTCATCAAAATGATCCACGGG
ACCATTAAAAACCAGTTACAGGGATTACAAAAGTCTTTGATGGTATACATTGCAGAAATTT
ATTTCAGAGCTTGGAAAAAGGCTTCAGGGAAAATACTGGAGGCGATTGAAAATGATTGCAT
CCAGGACTTCATGTTCCACGGGATACACCTTCCGAGGAGGTCTCCAGTGCATTCCAAAGTG
CGGGAGGTTCTGAGTTACTTTCACCATCAAAAGAAAGTTCGGCAGGGAGTGGAAGAGATGC
TTTATAGATTATATAAGCCCATCCTTTGGAGAGGATTAAAGGCCAGAAACTCTGAAGTTCG
ATCAAATGCTGCATTGTTGTTTGTTGAAGCATTTCCTATTAGGGATCCAAACCTTCATGCT
ATTGAAATGGATAGTGAAATCCAGAAACAGTTTGAAGAGCTCTATAGCCTTTTAGAAGATC
CTTACCCGATGGTCCGTTCCACAGGGATCCTTGGTGTTTGTAAAATAACTTCTAAGTACTG
GGAAATGATGCCCCCGACCATTCTTATTGACCTCCTGAAGAAGGTGACTGGGGAACTGGCA
TTTGACACGAGCTCAGCTGATGTTCGTTGTTCTGTCTTTAAGTGTCTGCCAATGATTTTGG
ACAACAAACTGAGCCACCCATTGTTAGAGCAGCTCCTTCCAGCTCTCAGATACAGTCTCCA
CGACAATTCGGAGAAAGTGAGGGTAGCTTTTGTGGACATGCTGTTGAAGATCAAAGCTGTG
AGGGCTGCTAAGTTTTGGAAAATATGTCCCATGGAGCACATTCTGGTTCGTCTGGAAACTG
ATTCTCGACCTGTGTCTCGGCGCCTGGTGAGCCTCATCTTTAATTCTTTCCTGCCTGTGAA
TCAGCCGGAGGAGGTCTGGTGCGAGCGCTGTGTCACCCTGGTGCAGATGAACCACGCCGCT
GCCAGGAGGTTCTATCAGTACGCCCACGAACACACCGCCTGCACCAACATAGCAAAGCTGA
TTCACGTTATTCGTCATTGCTTAAATGCCTGTATCCAGAGGGCAGTGAGAGAGCCTCCAGA
GGACGAGGAGGAAGAGGACGGAAGGGAGAAGGAGAATGTGACTGTTCTGGACAAAACACTG
TCAGTAAACGATGTTGCATGCATGGCAGGTTTACTAGAAATCATTGTGATTCTCTGGAAAA
GTATTGACAGAAGTATGGAAAATAATAAAGAGGCCAAACTTTACACGATTAACAAGTTTGC
CTCTTGTGCTTTCCAGAGTATCTGAAAGTATTTAAGGATGATCGCTGCAAGATCCCTTTATTC

ATGCTAATGTCCTTTATGCCGGCCTCTGCTGTCCCCCCATTCAGCTGTGGTGTGATTTCCA
CGCTGAGAAGCCGGGAGGAGGGCGCTGTGGACAAGAGCTACTGCACTTTGTTGGATTGCCT
CTGCTCCTGGGGGCAGGTGGGGCACATTCTGGAGCTTGTTGACAACTGGCTGCCCACAGAG
CATGCCCAGGCCAAGAGCAACACAGCTTCTAAAGGTAGGGTGCAGATCCATGACACACGCC
CAGTCAAACCTGAATTGGCATTGGTCTACATTGAGTATCTGCTGACTCATCCAAAGAACCG
CGAGTGCTTGCTCTCTGCTCCTCGGAAGAAACTTAACCATCTTTTGAAAGCCCTTGAAACG
TCAAAGGCAGATCTGGAGTCACTTCTGCAGACACCGGGTGGGAAGCCTCGTGGCTTCAGTG
AAGCAGCTGCCCCGCGAGCCTTTGGTCTCCACTGTCGCCTGAGCATCCATCTTCAGCACAA
GTTCTGCTCAGAAGGAAAGGTGTATTTGTCCATGTTGGAAGACACTGGCTTTTGGTTAGAA
AGCAAAATTTTATCTTTTATTCAAGATCAAGAAGAAGACTACCTGAAGCTTCATAGGGTCA
TTTATCAGCAAATTATCCAGACCTACCTGACTGTGTGTAAAGATGTTGTTATGGTAGGCCT
TGGTGACCATCAGTTTCAGATGCAACTCTTACAGCGGAGTCTTGGAATCATGCAAACAGTG
AAGGGATTTTTTTATGTTTCATTACTTCTTGACATTCTGAAAGAGATAACTGGAAGTTCCT
TGATTCAGAAAACAGATTCAGATGAAGAAGTTGCAATGCTGTTGGACACAGTCCAGAAAGT
ATTTCAGAAAATGTTGGAATGTATTGCACGGAGCTTCAGGAAGCAGCCGGAAGAAGGCCTG
CGGCTGCTTTATTCTGTTCAGAGGCCTCTTCATGAGTTCATTACTGCTGTTCAGTCTCGGC
ACACAGACACCCCTGTGCACCGGGGTGTACTTTCTACTCTGATCGCTGGGCCTGTGGTTGA
GATAAGTCACCAGCTACGGAAGGTTTCTGACGTAGAAGAGCTTACCCCTCCAGAGCATCTT
TCTGATCTTCCACCATTTTCAAGGTGTTTAATAGGAATAATAATAAAGTCTTCGAATGTGG
TCAGGTCATTTTTGGATGAATTAAAGGCATGTGTGGCTTCTAATGATATTGAAGGCATTGT
GTGCCTCACGGCTGCTGTGCATATTATCCTGGTTATTAATGCAGGTAAACATAAAAGCTCA
AAAGTGAGGGAGGTTGCAGCCACTGTTCACAGAAAACTAAAGACATTCATGGAAATTACTT
TGGAAGAGGATAGCATTGAAAGATTTCTCTATGAATCATCATCAAGAACTCTGGGAGAACT
TTTGAATTCATAACCAAGCCAACATCTCCAGACATGAAGGGC

FIGURE 72

TCCAACCAGATTTTCACCGCTATGCCTCCTTTCACTCTGGGAGTCTTCCAGAGGTCTTGCA
CTCGGGAGAGCATGCTCAGGTTTCCCCAGCTCTACAAAATCACCCAGAATGCCAAAGACTT
CAACGTGACCAGTGGTCATGCCACTGCCTGTTGATTTGTTGAAAATATTGTTTACACGTAT
GTTCTTGTTACTGATTGTCAGAAAGCTGGTTTTGAGACTGCAGCTTGGACTAAATTCAGTC
ATCTGGCTGTCTGGGGAAGCATGCTGACCAGTCTGGTGTTCTTTGGCATCTACTCAGCCAT
CTGGTCCACCATTCTCATTGCCCCAAATATGAGAGGACAGAAGAATGGTACCGGTACTGCC
AATGGAGATGGAGGAAGGAGACAGAAAGAAACAGAGCCCAGACCCTAGGGACCACCAGCAT
TTGCAGAATGGATAAACAGCCTTCTTCCTAACAAAGGAAGCACAGCAACTGTGATCCTGAG
CTGTGCACACTTCTGGTTGGGATTATTTCTGGTTTCTACTTCCTGTTTGAAGATGTGGCAT
GGAGAGTGAACAAGCTGCTGCCCACCACCTGGCATCACAGCCCCAGAACTCAGCTATTTCC
ATGGGACCACAGCATCTCATCTCTGGGCTGAGCCAGAAAGACCCCTACTGAAGTCCAGAGG
CACTTTTCTGAAAGGCTCTGCTTTGACCTGAAGTATTTTATCTATCCTCAGTCTCAGGACA
CTGTTGATGGAATTAAGGCCAAGCACATCTGCAAAAAAGACATTGCTGGAGGAGGTGCAAA
GAGCTGGAAACCAAGTCTCCAGTCCTGGGAAAAGCAGTGGTATGGAAAAGCAATGGAAAGA
GCATTTTGAAAATGCCATTCCACTGTTTTCTGGCCTTTATGATTTCTGCTGAGAAATCCAC
TGTTAGTCTGATGGGGTCTCCTTCATAGCACCAATGACCTGAAGAGCCTTGTTGAAGGAAG
ACTCCATCTGATGACTCAGAGCAAGTATTTTTTAGTGTGTTATTGTTATTAGCAGAAAGAG
GGCCATAAAATACATGGGGCAAGCTGAATATATCTTAGGCAAAAGAAGAAAATATTCAAAT
TCTTATGTTATTTTATCTAATTATTTTATCTCTTTTTGTGTGTGACTTATAATGTGTGTAT
TGTATTAATAAAAGTATATAAACATGTAGTTT

FIGURE 73

CATTCCCACCCTTCCTTCTTTAATAAGCAGGAGCGAAAAAGACAAATTCCAAAGAGGATTG
TTCAGTTCAAGGGAATGAAGAATTCAGAATAATTTTGGTAAATGGATTCCAATATCGGGAA
TAAGAATAAGCTGAACAGTTGACCTGCTTTGAAGAAACATACTGTCCATTTGTCTAAAATA
ATCTATAACAACCAAACCAATCAAAATGAATTCAACATTATTTTCCCAGGTTGAAAATCAT
TCAGTCCACTCTAATTTCTCAGAGAAGAATGCCCAGCTTCTGGCTTTTGAAAATGATGATT
GTCATCTGCCCTTGGCCATGATATTTACCTTAGCTCTTGCTTATGGAGCTGTGATCATTCT
TGGTGTCTCTGGAAACCTGGCCTTGATCATAATCATCTTGAAACAAAAGGAGATGAGAAAT
GTTACCAACATCCTGATTGTGAACCTTTCCTTCTCAGACTTGCTTGTTGCCATCATGTGTC
TCCCCTTTACATTTGTCTACACATTAATGGACCACTGGGTCTTTGGTGAGGCGATGTGTAA
GTTGAATCCTTTTGTGCAATGTGTTTCAATCACTGTGTCCATTTTCTCTCTGGTTCTCATT
GCTGTGGAACGACATCAGCTGATAATCAACCCTCGAGGGTGGAGACCAAATAATAGACATG
CTTATGTAGGTATTGCTGTGATTTGGGTCCTTGCTGTGGCTTCTTCTTTGCCTTTCCTGAT
CTACCAAGTAATGACTGATGAGCCGTTCCAAAATGTAACACTTGATGCGTACAAAGACAAA
TACGTGTGCTTTGATCAATTTCCATCGGACTCTCATAGGTTGTCTTATACCACTCTCCTCT
TGGTGCTGCAGTATTTTGGTCCACTTTGTTTTATATTTATTTGCTACTTCAAGATATATAT
ACGCCTAAAAAGGAGAAACAACATGATGGACAAGATGAGAGACAATAAGTACAGGTCCAGT
GAAACCAAAAGAATCAATATCATGCTGCTCTCCATTGTGGTAGCATTTGCAGTCTGCTGGC
TCCCTCTTACCATCTTTAACACTGTGTTTGATTGGAATCATCAGATCATTGCTACCTGCAA
CCACAATCTGTTATTCCTGCTCTGCCACCTCACAGCAATGATATCCACTTGTGTCAACCCC
ATATTTTATGGGTTCCTGAACAAAAACTTCCAGAGAGACTTGCAGTTCTTCTTCAACTTTT
GTGATTTCCGGTCTCGGGATGATGATTATGAAACAATAGCCATGTCCACGATGCACACAGA
TGTTTCCAAAACTTCTTTGAAGCAAGCAAGCCCAGTCGCATTTAAAAAAATCAACAACAAT
GATGATAATGAAAAAATCTGAAACTACTTATAGCCTATGGTCCCGGATGACATCTGTTTAA
AAACAAGCACAACCTGCAACATACTTTGATTACCTGTTCTCCCAAGGAATGGGGTTGAAAT
CATTTGAAAATGACTAAGATTTTCTTGTCTTGCTTTTTACTGCTTTTGTTGTAGTTGTCAT
AATTACATTTGGAACAAAAGGTGTGGGCTTTGGGGTCTTCTGGAAATAGTTTTGACCAGAC
ATCTTTGAAGTGCTTTTTGTGAATTTATGCATATAATATAAAGACTTTTATACTGTACTTA
TTGGAATGAAATTTCTTTAAAGTATTACGATGCGCTGACTTCAGAAGTACCTGCCATCCAA
TACGGTCATTAGATTGGGTCATCTTGATTAGATTAGATTAGATTAGATTGTCAACAGATTG
GGCCATCCTTACTTTATGATAGGCATCATTTTAGTGTGTTACAATAGTAACAGTATGCAAA
AGCAGCATTCAGGAGCCGAAAGATAGTCTTGAAGTCATTCAGAAGTGGTTTGAGGTTTCTG
TTTTTTGGTGGTTTTTGTTTGTTTTTTTTTTTTTTCACCTTAAGGGAGGCTTTCATTTCCT
CCCGACTGATTGTCACTTAAATCAAAATTTAAAAATGAATAAAAAGACATACTTCTCAGCT
GCAAATATTATGGAGAATTGGGCACCCACAGGAATGAAGAGAGAAAGCAGCTCCCCAACTT

TTAGTATTGCTGCAAATAGCTAAATTATATTTATTTGAATTGATGGTCAAGAGATTTTCCA
TTTTTTTTACAGACTGTTCAGTGTTTGTCAAGCTTCTGGTCTAATATGTACTCGAAAGACT
TTCCGCTTACAATTTGTAGAAACACAAATATCGTTTTCCATACAGCAGTGCCTATATAGTG
ACTGATTTTAACTTTCAATGTCCATCTTTCAAAGGAAGTAACACCAAGGTACAATGTTAAA
GGAATATTCACTTTACCTAGCAGGGAAAAATACACAAAAACTGCAGATACTTCATATAGCC
CATTTTAACTTGTATAAACTGTGTGACTTGTGGCGTCTTATAAATAATGCACTGTAAAGAT
TACTGAATAGTTGTGTCATGTTAATGTGCCTAATTTCATGTATCTTGTAATCATGATTGAG
CCTCAGAATCATTTGGAGAAACTATATTTTAAAGAACAAGACATACTTCAATGTATTATAC
AGATAAAGTATTACATGTGTTTGATTTTAAAAGGGCGGACATTTTATTAAAATCAATATTG
TTTTTGCTTTTTTCTGAGGAGTCTCTTTCAGTTTCATTTTTTCTCATCCCATGACTTCCCTC
CGATGGT

FIGURE 74

AGTATGAGGCGCCCACCATGGTGCGATGCGACCGCGGGCTGCAGATGCTGCTGACCACGGC
CGGAGCCTTCGCCGCCTTCTCGCTCATGGCCATCGCCATCGGCACCGACTACTGGCTGTAC
TCCAGCGCGCACATCTGCAACGGCACCAACCTGACCATGGACGACGGGCCCCCGCCCCGCC
GCGCCCGCGGCGACCTCACCCACTCTGGTCTGTGGCGGGTGTGCTGCATCGAAGGGATCTA
TAAAGGGCACTGCTTCCGGATCAATCACTTCCCAGAGGACAATGACTACGACCACGACAGC
TCGGAGTACCTCCTCCGCATCGTGCGAGCCTCCAGCGTCTTCCCCATCCTCAGCACCATCC
TGCTCCTGCTGGGTGGCCTGTGCATCGGTGCTGGCAGGATCTACAGCCGCAAGAACAACAT
CGTCCTCAGTGCCGGCATCCTCTTCGTGGCTGCAGGCCTCAGTAACATCATCGGTATCATC
GTCTACATTTCCAGCAACACAGGTGACCCGAGTGACAAGCGGGACGAAGACAAAAAGAACC
ATTACAACTACGGCTGGTCTTTTTACTTTGGAGCTCTGTCTTTCATTGTGGCTGAGACCGT
GGGCGTCCTGGCTGTAAACATTTACATTGAGAAAAATAAAGAGTTGAGGTTTAAGACCAAA
CGGGAATTCCTTAAGGCGTCTTCCTCTTCTCCTTATGCCAGGATGCCGAGCTACAGGTACC
GGCGACGGCGCTCGAGGTCCAGCTCAAGGTCCACCGAGGCCTCGCCCTCCAGGGACGTGTC
GCCCATGGGCCTGAAGATCACAGGGGCCATCCCCATGGGGGAGCTGTCCATGTACACGCTG
TCCAGGGAGCCCCTCAAGGTGACCACCGCAGCCAGCTACAGCCCCGACCAGGAGGCCAGCT
TCCTGCAGGTGCATGACTTTTTCCAGCAGGACCTGAAGGAAGGTTTCCACGTCAGCATGCT
GAACCGACGGACGACCCCTGTGTGA

FIGURE 75

CGGGTCTGATAGTCCCTACCTGTCAGGACTGGTGTTAGGATGAGATAATGTTTGTGAACTG
TAAACATATATAAACGTGTGCTACTGTGAGAACTGGAACAAAGAAGAGAGGGAGTGAGAGA
AATCAAGGGAGGGCTGGGGCTGGGAAAGAACGAAAAGGGAGTCGCGTATAGAGGAGAGGCG
ACAGTCGCGAGCCACACTTTGCAATGAAACTCTTTAGACTTTCTGCCGGGAGAGCGGCCCA
GACGCGCCAGGTCTGTAGCAGGAGGCCGCGAGGGCGGGTCCCCAGAAGCCTACAGGTGAGT
ATCGGTTCTCCCCTTCCCGGCTTTCGGTCCGGAGGAGGCGGGAGCAGCTTCCCTGTTCTGA
TCCTATCGCGGGCGGCGCAGGGCCGGCTTGGCCTTCCGTGGGACGGGGAGGGGGGCGGGAT
GTGTCACCCAAATACCAGTGGGGACGGTCGGTGGTGGAACCAGCCGGGCAGGTCGGGTAGA
GTATAAGAGCCGGAGGGAGCGGCCGGGGCGCAGACGCCTGCAGACCATCCCAGACGCCGGA
GCCCGAGCCCCGACGAGTCCCCGCGCCTCATCCGCCCGCGTCCGGTCCGCGTTCCTCCGCC
CCACCATGGCTCGGGGCCCCGGCCTCGCGCCGCCACCGCTGCGGCTGCCGCTGCTGCTGCT
GGTGCTGGCGGCGGTGACCGGCCACACGGCCGCGCAGGACAACTGCACGTGTCCCACCAAC
AAGATGACCGTGTGCAGCCCCGACGGCCCCGGCGGCCGCTGCCAGTGCCGCGCGCTGGGCT
CGGGCATGGCGGTCGACTGCTCCACGCTGACCTCCAAGTGTCTGCTGCTCAAGGCGCGCAT
GAGCGCCCCCAAGAACGCCCGCACGCTGGTGCGGCCGAGTGAGCACGCGCTCGTGGACAAC
GATGGCCTCTACGACCCCGACTGCGACCCCGAGGGCCGCTTCAAGGCGCGCCAGTGCAACC
AGACGTCGGTGTGCTGGTGCGTGAACTCGGTGGGCGTGCGCCGCACGGACAAGGGCGACCT
GAGCCTACGCTGCGATGAGCTGGTGCGCACCCACCACATCCTCATTGACCTGCGCCACCGC
CCCACCGCCGGCGCCTTCAACCACTCAGACCTGGACGCCGAGCTGAGGCGGCTCTTCCGCG
AGCGCTATCGGCTGCACCCCAAGTTCGTGGCGGCCGTGCACTACGAGCAGCCCACCATCCA
GATCGAGCTGCGGCAGAACACGTCTCAGAAGGCCGCCGGTGAAGTGGATATCGGCGATGCC
GCCTACTACTTCGAGAGGGACATCAAGGGCGAGTCTCTATTCCAGGGCCGCGGCGGCCTGG
ACTTGCGCGTGCGCGGAGAACCCCTGCAGGTGGAGCGCACGCTCATCTATTACCTGGACGA
GATTCCCCCGAAGTTCTCCATGAAGCGCCTCACCGCCGGCCTCATCGCCGTCATCGTGGTG
GTCGTGGTGGCCCTCGTCGCCGGCATGGCCGTCCTGGTGATCACCAACCGGAGAAAGTCGG
GGAAGTACAAGAAGGTGGAGATCAAGGAACTGGGGGAGTTGAGAAAGGAACCGAGCTTGTA
GGTACCCGGCGGGGCAGGGGATGGGGTGGGGTACCGGATTTCGGTATCGTCCCAGACCCAA
GTGAGTCACGCTTCCTGATTCCTCGGCGCAAAGGAGACGTTTATCCTTTCAAATTCCTGCC
TTCCCCCTCCCTTTTGCGCACACACCAGGTTTAATAGATCCTGGCCTCAGGGTCTCCTTTC
TTTCTCACTTCTGTCTTGAGGGAAGCATTTCTAAAATGTATCCCCTTTCGGTCCAACAACA
GGAAACCTGACTGGGGCAGTGAAGGAAGGGATGGCACAGCGTTATGTGTAAAAAACAAGTA
TCTGTATGACAACCCGGGATCGTTTGCAAGTAACTGAATCCATTGCGACATTGTGAAGGCT
TAAATGAGTTTAGATGGGAAATAGCGTTGTTATCGCCTTGGGTTTAAATTATTTGATGAGT
TCCACTTGTATCATGGCCTACCCGAGGAGAAGAGGAGTTTGTTAACTGGGCCTATGTAGTA
CCCTCATTTTACCATCGTTTTGTATTACTGACCACATATGCTTGTCACTGGGAAAGAAGCCTG

TTTCAGCTGCCTGAACGCAGTTTGGATGTCTTTGAGGACAGACATTGCCCGGAAACTCAGT
CTATTTATTCTTCAGCTTGCCCTTACTGCCACTGATATTGGTAATGTTCTTTTTTGTAAAA
TGTTTGTACATATGTTGTCTTTGATAATGTTGCTGTAATTTTTTAAAATAAAACACGAATT
TAATAAAATATGGGAAAGGCACAAACCAGAAGTCGGCATTTGTGAAAAGTCCCTCCAGATT
TCTATCACTTTGGTCTCTAATTTCCCAAGACTTGTATTTTTTTTTTTATTTCAAATTATAAC
ACTTTTTTTTTCCCCCAGAAGTGGGTGTTTCATGTTGCTACTCTGGTGTGTCCCAAGATATC
CTAACTGGCCAGTGTAAATGCTATTCTTTCTAAATAAGATTATTTGGAAACTTCCTTCAAA
CTGCAGGAGGGCGAGCTCTGAGGGCACGAGAAGCTAAAACTAGCTGCTTTTGATGAAAAAG
AGTGCCAGTCTTTGGTCATCTCTAAACAAGGCTTATCACCAATGGAGACAGAAAACTCTAG
TTCAAGAGCTGTACCTCCTTTGAATCCCAGCCCTACTCGAAATAAGTGGTACTATTTCCAT
TTAGCCTTTGAGCAAATCACTTAACTCAAAGGCGTTGTGGCTCTAAGATTAAACGACTTT

FIGURE 76

GCGGCCGCGGAGCAAGAAGGGCGCCGCGTCGTGCGGCCCGCGCAGCCCCCGGAGCCATGGG
CAAGTGCAGCGGGCGCTGCACGCTGGTCGCCTTCTGCTGCCTGCAGCTGGTGGCTGCGCTG
GAGCGGCAGATCTTTGACTTCCTGGGCTACCAGTGGGCTCCCATCCTAGCCAACTTCCTGC
ACATCATGGCAGTCATCCTGGGCATCTTTGGCACCGTGCAGTACCGCTCCCGGTACCTCAT
CCTGTATGCAGCCTGGCTGGTGCTCTGGGTTGGCTGGAATGCATTTATCATCTGCTTCTAC
TTGGAGGTTGGACAGCTGTCCCAGGACCGGGACTTCATCATGACCTTCAACACATCCCTGC
ACCGCTCCTGGTGGATGGAGAATGGGCCAGGCTGCCTGGTGACACCTGTTCTGAACTCCCG
CCTGGCTCTGGAGGACCACCATGTCATCTCTGTCACTGGCTGCCTGCTTGACTACCCCTAC
ATTGAAGCCCTCAGCAGCGCCCTGCAGATCTTCCTGGCACTGTTCGGCTTCGTGTTCGCCT
GCTACGTGAGCAAAGTGTTCCTGGAGGAGGAGGACAGCTTTGACTTCATCGGCGGCTTTGA
CTCCTACGGATACCAGGCGCCCCAGAAGACGTCGCATTTACAGCTGCAGCCTCTGTACACG
TCGGGGTAGCCTCTGCCCCGCGCCCACCCCGGCGCCTCGCCCTGGGCTGACCGCAGCTGCC
GCGAGCTCGGGCCAAGGCGCAGGCGTGTCCCCCTGGTGGCCCGCGCTCTCACTGCAGCCTG
TGCCCAACCCCGCGTCTGCATCTGGAGATGCGGACTTGGACGTGGACTTGGACTTGGACTT
GGATTTGAGCTTGGCTCTTCGCAGCCCGGACTTCGGAGGAGTGGGGCGGGGCGGGGGAGGG
GCACCACGGGTTTTTTGTTTTTTGTTTGTTTGTTTTTAATCTCAGCCTTGGCGTGAGCTGG
GGCCTTCCTCTCTTCTCCAGCCTCTCCCTTTCACTCTTCGCCCAGCATCCTGCCCCCCTGT
CCAAAAACAGCAGGACATCAGGCCCATCCCATCCCACCACACTCACTCACCAGCTCTGGGG
AAAGCTACTGTGAACTAGGAGCAGGATTCCTGGGTTCTAATCGCAGGTCCATCACTGACTG
TGACGTCTAGCAAAGCCCTTGCCCTCTCTGAGCCTCGGTTTCCGCACCTCAAGTAATTAAT
CCCTTAGCAAATGGACTCTTTTAGACTTCTCATTTAACTCAATTCCCTGAGCTAGACTGGG
ATTAAAATTCTCATTTTGCAGTACATTAAAACTGATGCCCAGAGATGTGATTTGCTTGAGG
CCACACAGCTAGATTTTTGGTGGAAGTGGGCCTTGAACACAGTGTACTTTCTGCAGTTTCT
GACTGTAAAACCCAGTGTCTGCTCTCTGAGTTCCATTTCCAAGCCCCCCTCCATCTTGGAC
CTATGTGGTCTCCACCATATTCACACACCACCACCACCACTTGCCAATGCCTCTCTTAAAG
CAATATACCCATTCGTTCTCTTATTGGGAACTGGATGGATGAAGCCCCAAATTCAGCCCCA
CCCACAGAGAAGCCTTCCTACACTCAGCCTCTGTCCACCCTTGGCAAATCTTTCAAGCTCT
CTCCTCCAGGAAAGTGGGGCCCCAACTCAGTCACTCCACCCCCTTCCAGGTCCCTGAGGCT
GGTTCTACTGTATCCCCATCACCTCCACAACTCCACTCACCCCTGACGGCTCCATCCACCT
CACCAGTTGGAAGGCTTGTGGTTTCAGAGAGGAGCAATGCTGGTCAGCGCTGCCCAGACTC
CAGTGTTTACAGATCACCAGCATTTACAACCAATCCAATGGCCAGAAGCCTCCTCTAACAA
GCCCAGAAGGAGTTCTGAAGGGGCAGATGGGGGTGTGAGTAGTCGGGGAGTCGGGATTGCC
AGCACCCTCACCCTTCCTTGGGGGCAAGTAGAGGTGAGAACACTTTCCCCACCTCCCTCCA
CAGACACTCCTGAGGACGCTGCATCCCACGCACTGCCTGGTGCGTCCATAGAGAGAGGATC
[illegible]

CATGGAGGAACTGGGGGGCACCCAGATGGTGCAGATGGTTTGCACACCTGAGCCTGTCTGT
GGTGACCATTCCGCTCCTCTCCCACTACCCTCCAATCTATCATTCCCTACTCTCTAAGGCC
AAAATATCCTGAGCAAGGCTGGCAACCCCACCCCACCATCCCAAATGCAAGCAGCCAGGCC
CAGGAGTTCCTCTGGCCCCCACAGGCATGGAGCTCCCAGCTGGTGGGTACAGCTTGAGAGG
GGGGCAGCTCCCTCAGGCTAAGCTACTGCCCTTCACTGGGCCAGCCCTGCCTCCAGCCCTC
ACCTCTCTCACCCCAACTGTCCCCCAAGCCCCTTTCTACTCAACGGGTGTAGCCACTGGTG
CTTTGAAGCCTTTTGTTTTTATAAGATGGTTTTTGCAAGGGGACCAGGTTCTCTTTTCACT
GGGACCTTGCAAGGAGGGGAGTGCTCTCCTGGTTTCTGTGCAGGCGGGTTGATTAAAGATG
GTGTTTTCTTCTCT

FIGURE 77

GGCGGGGAGAGGAGAGGAGAGAAGAGCCGCGGGGGGCCCAGCCCGGAGCCAGGATGCCCGC
GCCGCGCGCCCGGGAGCAGCCCCGCGTGCCCGGGGAGCGCCAGCCGCTGCTGCCTCGCGGT
GCGCGGGGCCCTCGACGGTGGCGGCGGGCGGCGGGCGCGGCCGTGCTGCTGGTGGAGATGC
TGGAGCGCGCCGCCTTCTTCGGCGTCACCGCCAACCTCGTGCTGTACCTCAACAGCACCAA
CTTCAACTGGACCGGCGAGCAGGCGACGCGCGCCGCGCTGGTATTCCTGGGCGCCTCCTAC
CTGCTGGCGCCCGTGGGCGGCTGGCTGGCCGACGTGTACCTGGGCCGCTACCGCGCGGTCG
CGCTCAGCCTGCTGCTCTACCTGGCCGCCTCGGGCCTGCTGCCCGCCACCGCCTTCCCCGA
CGGCCGCAGCTCCTTCTGCGGAGAGATGCCCGCGTCGCCGCTGGGACCTGCCTGCCCCTCG
GCCGGCTGCCCGCGCTCCTCGCCCAGCCCCTACTGCGCGCCCGTCCTCTACGCGGGCCTGC
TGCTACTCGGCCTGGCCGCCAGCTCCGTCCGGAGCAACCTCACCTCCTTCGGTGCCGACCA
GGTGATGGATCTCGGCCGCGACGCCACCCGCCGCTTCTTCAACTGGTTTTACTGGAGCATC
AACCTGGGTGCTGTGCTGTCGCTGCTGGTGGTGGCGTTTATTCAGCAGAACATCAGCTTCC
TGCTGGGCTACAGCATCCCTGTGGGCTGTGTGGGCCTGGCATTTTTCATCTTCCTCTTTGC
CACCCCCGTCTTCATCACCAAGCCCCCGATGGGCAGCCAAGTGTCCTCTATGCTTAAGCTC
GCTCTCCAAAACTGCTGCCCCCAGCTGTGGCAACGACACTCGGCCAGAGACCGTCAATGTG
CCCGCGTGCTGGCCGACGAGAGGTCTCCCCAGCCAGGGGCTTCCCCGCAAGAGGACATCGC
CAACTTCCAGGTGCTGGTGAAGATCTTGCCCGTCATGGTGACCCTGGTGCCCTACTGGATG
GTCTACTTCCAGATGCAGTCCACCTATGTCCTGCAGGGTCTTCACCTCCACATCCCAAACA
TTTTCCCAGCCAACCCGGCCAACATCTCTGTGGCCCTGAGAGCCCAGGGCAGCAGCTACAC
GATCCCGGAAGCCTGGCTCCTCCTGGCCAATGTTGTGGTGGTGCTGATTCTGGTCCCTCTG
AAGGACCGCTTGATCGACCCTTTACTGCTGCGGTGCAAGCTGCTTCCCTCTGCTCTGCAGA
AGATGGCGCTGGGGATGTTCTTTGGTTTTACCTCCGTCATTGTGGCAGGAGTCCTGGAGAT
GGAGCGCTTACACTACATCCACCACAACGAGACCGTGTCCCAGCAGATTGGGGAGGTCCTG
TACAACGCGGCACCACTGTCCATCTGGTGGCAGATCCCTCAGTACCTGCTCATTGGGATCA
GTGAGATCTTTGCCAGCATCCCAGGCCTGGAGTTTGCCTACTCAGAGGCCCCGCGCTCCAT
GCAGGGCGCCATCATGGGCATCTTCTTCTGCCTGTCGGGGGTGGGCTCACTGTTGGGCTCC
AGCCTAGTGGCACTGCTGTCCTTGCCCGGGGGCTGGCTGCACTGCCCCAAGGACTTTGGGA
ACATCAACAATTGCCGGATGGACCTCTACTTCTTCCTGCTGGCTGGCATTCAGGCCGTCAC
GGCTCTCCTATTTGTCTGGATCGCTGGACGCTATGAGAGGGCGTCCCAGGGCCCAGCCTCC
CACAGCCGTTTCAGCAGGGACAGGGGCTGAACAGGCCCTATTCCAGCCCCCTTGCTTCAC

FIGURE 78A

GGCCGCGGGCTCTCGCGGGGCGGCGACGCCGCGGGGAGGATGCTGCTTGCCGCGCCCGCGT
CCTCACCGTCCTCCCGGGCCGCCTGCTGGGGCTTTGTTGTGGCCCGGACGCCGCGGGCCAC
CCCCTGAAGTCGCCTGCCGCCGCCGCCGCCGCACCTAGCGGACGGGCGGGCGGGCGCGCGT
GTGCCCAGGAGTGCGCGCCTGTCGCGGTGGTGGGTGCAGGACTGGACCCACGGGCCCATTG
TGCGCCCGCCCGCGGCAGCCAGGACCATGTGGGTGAACCCGGAGGAGGTGTTGCTGGCCAA
CGCGCTGTGGATCACCGAGAGGGCCAACCCATACTTCATCCTGCAGCGGAGGAAGGGCCAC
GCCGGCGATGGAGGCGGCGGCGGCGGACTGGCGGGCCTGCTGGTGGGTACCCTTGATGTTG
TGTTGGACTCCAGCGCCCGGGTCGCTCCTTACCGAATCTTGTACCAGACTCCAGACTCCCT
GGTCTACTGGACCATCGCCTGTGGTGGTTCCAGGAAAGAAATCACTGAACACTGGGAATGG
CTTGAGCAAAATCTCTTGCAGACACTCTCCATCTTTGAAAATGAGAATGATATCACCACAT
TTGTGAGAGGAAAAATACAGGGCATCATTGCAGAATACAACAAAATCAATGATGTAAAGGA
AGATGATGACACGGAGAAGTTTAAAGAAGCCATTGTGAAATTTCATAGGCTGTTTGGGATG
CCAGAGGAAGAGAAACTCGTCAACTATTACTCTTGCAGCTATTGGAAGGGGAAGGTCCCCC
GTCAGGGTTGGATGTACCTCAGCATTAACCACCTTTGCTTTTATTCTTTTCTTATGGGAAG
GGAAGCGAAACTGGTCATCCGGTGGGTAGACATCACTCAGCTTGAGAAGAATGCCACCCTG
CTTCTGCCTGATGTGATCAAAGTGAGCACACGGTCCAGTGAGCATTTCTTCTCTGTATTCC
TCAACATCAACGAGACCTTCAAGTTAATGGAGCAGCTTGCCAACATAGCCATGAGGCAACT
CTTAGACAATGAGGGATTTGAACAAGATCGATCCCTGCCCAAACTCAAAAGGAAATCTCCT
AAAAAAGTGTCTGCTCTAAAACGTGATCTTGATGCCAGGGCAAAGAGTGAGAGATACCGTG
CACTTTTCCGGCTGCCCAAAGATGAAAAATTAGATGGCCACACAGACTGCACTCTCTGGAC
TCCATTTAACAAAATGCACATTTTGGGGCAGATGTTTGTGTCCACAAATTACATCTGTTTT
ACCAGCAAGGAGGAGAACTTATGTAGCCTCATTATCCCGCTCCGTGAGGTGACAATTGTGG
AAAAGGCAGACAGCTCCAGTGTGCTCCCCAGTCCCTTATCCATCAGCACCCGAAACAGGAT
GACCTTCCTATTTGCCAACTTGAAAGATAGAGACTTTCTAGTGCAGAGGATCTCAGATTTC
CTGCAACAGACTACTTCCAAAATATATTCTGACAAGGAGTTTGCAGGAAGTTACAACAGTT
CAGATGATGAGGTGTACTCTCGACCCAGCAGCCTCGTCTCCTCCAGCCCCCAGAGAAGCAC
GAGCTCTGATGCTGATGGAGAGCGCCAGTTTAACCTAAATGGCAACAGCGTCCCCACAGCC
ACACAGACCCTGATGACCATGTATCGGCGGCGGTCTCCCGAGGAGTTCAACCCGAAATTGG
CCAAAGAGTTTCTGAAAGAGCAAGCCTGGAAGATTCACTTTGCTGAGTATGGGCAAGGGAT
CTGCATGTACCGCACAGAGAAAACGCGGGAGCTGGTGTTGAAGGGCATCCCGGAGAGCATG
CGTGGGGAGCTCTGGCTGCTGCTGTCAGGTGCCATCAATGAGAAGGCCACACATCCTGGGT
ACTATGAAGACCTAGTGGAGAAGTCCATGGGGAAGTATAATCTCGCCACGGAGGAGATTGA
GAGGGATTTACACCGCTCCCTTCCAGAACACCCAGCTTTTCAGAATGAAATGGGCATTGCT
GCACTAAGGAGAGTCTTAACAGCTTATGCTTTTCGAAATCCCAACATAGGGTATTGCCAGG
CCATGAATATTGTCACTTCAGTGCTGCTGCTTTATGCCAAAGAGGAGGAAGCTTTCTGGCT

GCTTGTGGCTTTGTGTGAGCGCATGCTCCCAGATTACTACAACACCAGAGTTGTGGGTGCA
CTGGTGGACCAAGGTGTCTTTGAGGAGCTAGCACGAGACTACGTCCCACAGCTGTACGACT
GCATGCAAGACCTGGGCGTGATTTCCACCATCTCCCTGTCTTGGTTCCTCACACTATTTCT
CAGTGTGATGCCTTTTGAGAGTGCAGTTGTGGTTGTTGACTGTTTCTTCTATGAAGGAATT
AAAGTGATATTCCAGTTGGCCCTAGCTGTGCTGGATGCAAATGTGGACAAACTGTTGAACT
GCAAGGATGATGGGGAGGCCATGACCGTTTTGGGAAGGTATTTAGACAGTGTGACCAATAA
AGACAGCACACTGCCTCCCATTCCTCACCTCCACTCCTTGCTCAGCGATGATGTGGAACCT
TACCCTGAGGTAGACATCTTTAGACTCATCAGAACTTCCTACGAGAAATTCGGAACTATCC
GGGCAGATTTGATTGAACAGATGAGATTCAAACAGAGACTGAAAGTGATCCAGACGCTGGA
GGATACTACGAAACGCAACGTGGTACGAACCATTGTGACAGAAACTTCCTTTACCATTGAT
GAGCTGGAAGAACTTTATGCTCTTTTCAAGGCAGAACATCTCACCAGCTGCTACTGGGGCG
GGAGCAGCAACGCGCTGGACCGGCATGACCCCAGCCTGCCCTACCTGGAACAGTATCGCAT
TGACTTCGAGCAGTTCAAGGGAATGTTTGCTCTTCTCTTTCCTTGGGCATGTGGAACTCAC
TCTGACGTTCTGGCCTCCCGCTTGTTCCAGTTATTAGATGAAAATGGAGACTCTTTGATTA
ACTTCCGGGAGTTTGTCTCTGGGCTAAGTGCTGCATGCCATGGGGACCTCACAGAGAAGCT
CAAACTCCTGTACAAAATGCACGTCTTGCCTGAGCCATCCTCTGATCAAGATGAACCAGAT
TCTGCTTTTGAAGCAACTCAGTACTTCTTTGAAGATATTACCCCAGAATGTACACATGTTG
TTGGATTGGATAGCAGAAGCAAACAGGGTGCAGATGATGGCTTTGTTACGGTGAGCCTAAA
GCCAGACAAAGGGAAGAGAGCAAATTCCCAAGAAAATCGTAATTATTTGAGACTGTGGACT
CCAGAAAATAAATCTAAGTCAAAGAATGCAAAGGATTTACCCAAATTAAATCAGGGGCAGT
TCATTGAACTGTGTAAGACAATGTATAACATGTTCAGCGAAGACCCCAATGAGCAGGAGCT
GTACCATGCCACGGCAGCAGTGACCAGCCTCCTGCTGGAGATTGGGGAGGTCGGCAAGTTG
TTCGTGGCCCAGCCTGCAAAGGAGGGCGGGAGCGGAGGCAGTGGGCCGTCCTGCCACCAGG
GCATCCCAGGCGTGCTCTTCCCCAAGAAAGGGCCAGGCCAGCCTTACGTGGTGGAGTCTGT
TGAGCCCCTGCCGGCCAGCCTGGCCCCCGACAGCGAGGAACACTCCCTTGGAGGACAAATG
GAGGACATCAAGCTGGAGGACTCCTCGCCCCGGGACAACGGGGCCTGCTCCTCCATGCTGA
TCTCTGACGACGACACCAAGGACGACAGCTCCATGTCCTCATACTCGGTGCTGAGTGCCGG
CTCCCACGAGGAGGACAAGCTGCACTGCGAGGACATCGGAGAGGACACGGTCCTGGTGCGG
AGCGGCCAGGGCACGGCGGCACTGCCCCGGAGCACCAGCCTGGACCGGGACTGGGCCATCA
CCTTCGAGCAGTTCCTGGCCTCCCTCTTAACTGAGCCTGCCCTGGTCAAGTACTTTGACAA
GCCCGTGTGCATGATGGCCAGGATTACCAGTGCAAAAAACATCCGGATGATGGGCAAGCCC
CTCACCTCGGCCAGTGACTATGAAATCTCGGCCATGTCCGGCTGACACGGGCGCCTTCCCG
GGGGAGTGGGAGGAGAGGGAGGGGAGGGATTTTTTATGTTCTTCTGTGTTGAGTTTTTTCT
TTCTTTCTTTTAAATTAAATATTTATTAGTACCTGGCTTGAAGCCTAGTGTTTTCATAATG
TAATTCAATGAAAACTGTTGGAGAAATATTTAAACACCTCAATGTAGGTACATTACACTCT
TGTTGCGGGGAGGGGATTTACCAGAATACAGTTTATTTCGTGAATTCTAAAAAACAAAAAG

ATGGTTTAAAAAATAGTACTGTTTAATGCTAAGTAAGGCAGCAGTCATTTGTGTATTCAGG
CTTTTTAAATAAAATTAGAGCTGTAAGGAAAATGAAAAGCCACAAATGCAAGACTGTTCTT

FIGURE 78B

AAATGGAAGGCATAGTCAGCGAGGGTAAATCCTATACCACTTTAGGAAGTATTAAAAATAT
TTTTAAGATTTGAAATATATTTCATAGAAGTCCTCTATTCAAAATCATATTCCACAGATGT
TCCCCTTCAAAGGGAAAACATTTGGGGTTCTAAACAGTTATGAAAGTAAGTGATTTTTACA
TGATTCCAGAATAACACTTGTATTGACCAATTTAGACAGATACCAGACCAATTTTGCATTT
AAGAAATTGTTCTGATTATTTACGTCAACTCATTAGAATTCAGTGAAAAGTAACAGTCTTT
TGTCACAGAGAATCTGAAAGTAGCAGCAAAGACAGAGGGCTCATGACAGGTTTTTGCTTTT
GCTTTGCTTTTGTTTTTGAAAGAGTAAAAGTACTGATGCTTCTGATACTGGATGTTTAGCT
TCTTACTGCAAAAACATAAGTAAAACAGTCAACTTTACCATTTCCGTATTCTCCATAGATT
GAAGAAATTTATACCACATATCGCATATGACCATCTTTCCATCAAATCAATGTAGAGATAA
TGTAAACTGAAAAAAAATCTGCAAGATAATGTAACTGAATGTTTTAAAAACAGAACTTGTC
ACTTTATATAAAAGAATAGTATGCTCTATTTCCTGAATGGATGTGGAAATGAAAGCTAGCG
CACCTGCACTTTGAATTCTTGCTTCTTTTTTATTACTGTTATGATTTTGCTTTTTACAGAT
GTTGGACGATTTTTTCTTCTGATTGTTGAATTCATAATCATGGTCTCATTTCCTTTGCTTC
TTTGGAATATTTCTTTCAACACATTCCTTTATTTTATTATACATTGTGTCCTTTTTTTAGC
TATTGCTGCTGTTGTTTTTTATTCTATTTACAGGATGATTTTTAAACTGTCAAATGAAGTA
GTGTTAACCTCAAATAGGCTAAATGTGAACAAATAAAATACAGCAAATACTCAGATACAGC
TTTTTATCTTTGTGCTTGAGTTCCTGCCTAAGGAATAACATTATTCTTTTGACAACTTTTG
CAGGGGAAATTATATCAGGCAACCATTTTGATTAAGTAAATAAATTTTATAGGCAAACATA
TAGAGAGANATACAATTTGTAGTATATCAATGACTATATTTAAAATAAGGANTATAATTGT
TATCAGTTATCTAACTTAAAATGCTTATCCATAATGATCAGTGATATTCAGCTTTTTAAAA
TATGCTTGTTGG

FIGURE 79

MRAAYLFLLFLPAGLLAQGQYDLDPLPPFPDHVQYTHYSDQIDNPDYYDYQEVTPRPSEEQ
FQFQSQQQVQQEVIPAPTPEPGNAELEPTEPGPLDCREEQYPCTRLYSIHRPCKQCLNEVC
FYSLRRVYVINKEICVRTVCAHEELLRADLCRDKFSKCGVMASSGLCQSVAASCARSCGSC

Signal sequence.
amino acids 1-19

N-myristoylation sites.
amino acids 161-166, 167-172

FIGURE 80

MAARPLPVSPARALLLALAGALLAPCEARGVSLWNEGRADEVVSASVRSGDLWIPVKSFDS
KNHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTVILGHCYYH
GHVRGYSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFPAKKLKSVRGSCGSHH
NTPNLAAKNVFPPPSQTWARRHKRETLKATKYVELVIVADNREFQRQGKDLEKVKQRLIEI
ANHVDKFYRPLNIRIVLVGVEVWNDMDKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNAQL
VSGVYFQGTTIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNFGMNHDTLDRG
CSCQMAVEKGGCIMNASTGYPFPMVFSSCSRKDLETSLEKGMGVCLFNLPEVRESFGGQKC
GNRFVEEGEECDCGEPEECMNRCCNATTCTLKPDAVCAHGLCCEDCQLKPAGTACRDSSNS
CDLPEFCTGASPHCPANVYLHDGHSCQDVDGYCYNGICQTHEQQCVTLWGPGAKPAPGICF
ERVNSAGDPYGNCGKVSKSSFAKCEMRDAKCGKIQCQGGASRPVIGTNAVSIETNIPLQQG
GRILCRGTHVYLGDDMPDPGLVLAGTKCADGKICLNRQCQNISVFGVHECAMQCHGRGVCN
NRKNCHCEAHWAPPFCDKFGFGGSTDSGPIRQADNQGLTIGILVTILCLLAAGFVVYLKRK
TLIRLLFTNKKTTIEKLRCVRPSRPPRGFQPCQAHLGHLGKGLMRKPPDSYPPKDNPRRLL
QCQNVDISRPLNGLNVPQPQSTQRVLPPLHRAPRAPSVPARPLPAKPALRQAQGTCKPNPP
QKPLPADPLARTTRLTHALARTPGQWETGLRLAPLRPAPQYPHQVPRSTHTAYIK

Signal sequence.
amino acids 1-28

Transmembrane domain.
amino acids 709-729

N-glycosylation sites.
amino acids 111-114, 149-152, 381-384, 452-455, 651-654

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 206-209, 730-733, 742-745

N-myristoylation sites.
amino acids 88-93, 103-108, 177-182, 180-185, 313-318, 343-348, 377-382, 407-412, 423-428, 424-429, 524-529, 630-635, 635-640, 656-661, 668-673, 694-699, 708-713, 712-717, 878-883

signature.
amino acids 347-356.

Disintegrins proteins.
amino acids 454-504, 410-460, 486-536

FIGURE 81

MASRSMRLLLLLSCLAKTGVLGDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDAELECQSY
GNGAHLASILSLKEASTIAEYISGYQRSQPIWIGLHDPQKRQQWQWIDGAMYLYRSWSGKS
MGGNKHCAEISSNNNFLTWSSNECNKRQHFLCKYRP

Signal sequence.
amino acids 1-22

N-glycosylation site.
amino acids 50-53

N-myristoylation sites.
amino acids 64-69, 125-130

C-type lectin domain signature.
amino acids 129-154

Lectin C-type domain.
amino acids 47-156

FIGURE 82

MATMENKVICALVLVSMLALGTLAEAQTETCTVAPRERQNCGFPGVTPSQCANKGCCFDDT VRGVPWCFYPNTIDVPPEEECEF

Signal sequence.
amino acids 1-24

N-myristoylation sites.
amino acids 45-50, 64-69

Trefoil (P-type) domain.
amino acids 30-71

FIGURE 83

MLPQIPFLLLVSLNLVHGVFYAERYQTPTGIKGPLPNTKTQFFIPYTIKSKGIAVRGEQGT
PGPPGPAGPRGHPGPSGPPGKPGYGSPGLQGEPGLPGPPGPSAVGKPGVPGLPGKPGERGP
YGPKGDVGPAGLPGPRGPPGPPGIPGPAGISVPGKPGQQGPTGAPGPRGFPGEKGAPGVPG
MNGQKGEMGYGAPGRPGERGLPGPQGPTGPSGPPGVGKRGENGVPGQPGIKGDRGFPGEMG
PIGPPGPQGPPGERGPEGIGKPGAAGAPGQPGIPGTKGLPGAPGIAGPPGPPGFGKPGLPG
LKGERGPAGLPGGPGAKGEQGPAGLPGKPGLTGPPGNMGPQGPKGIPGSHGLPGPKGETGP
AGPAGYPGAKGERGSPGSDGKPGYPGKPGLDGPKGNPGLPGPKGDPGVGGPPGLPGPVGPA
GAKGMPGHNGEAGPRGAPGIPGTRGPIGPPGIPGFPGSKGDPGSPGPPGPAGIATKGLNGP
TGPPGPPGPRGPSGEPGLPGPPGPPGPPGQAVMPEGFIKAGQRPSLSGTPLVSANQGVTGM
PVSAFTVILSKAYPAIGTPIPFDKILYNRQQHYDPRTGIFTCQIPGIYYFSYHVHVKGTHV
WVGLYKNGTPVMYTYDEYTKGYLDQASGSAIIDLTENDQVWLQLPNAESNGLYSSEYVHSS
FSGFLVAPM

Signal sequence.
amino acids 1-18

Tyrosine kinase phosphorylation site.
amino acids 116-123

N-myristoylation sites.
amino acids 18-23, 276-281, 317-322, 350-355, 380-385,
415-420, 446-451, 529-534, 548-553, 587-592, 613-618, 661-666

Amidation site.
amino acids 219-222

C1q domain signature.
amino acids 571-601

C1q domain.
amino acids 553-677

Collagen triple helix repeat (20 copies).
amino acids 92-150, 156-214, 223-281, 282-340, 344-403,

404-463, 464-522

FIGURE 84

MPLPPLLLLLLAAPWGRAVPCVSGGLPKPANITFLSINMKNVLQWTPPEGLQGVKVTYTVQ
YFIYGQKKWLNKSECRNINRTYCDLSAETSDYEHQYYAKVKAIWGTKCSKWAESGRFYPFL
ETQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIYSNLKYNVSVLNTKSNRTW
SQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLKDQSSEFKAKIIFWY
VLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILIYGNEFDKRFFVPAEKIVINFITLNI
SDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFCDSEENT
EGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLQEEVSTQGTLLESQAAL
AVLGPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQ
DSEGCEPSEGDGLGEEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQMEN

Signal sequence.
amino acids 1-18

Transmembrane domain.
amino acids 239-259

N-glycosylation sites.
amino acids 31-34, 72-75, 80-83, 171-174, 180-183, 189-192, 304-307, 523-526

Tyrosine kinase phosphorylation sites.
amino acids 385-392, 518-526

N-myristoylation sites.
amino acids 53-58, 106-111, 368-373, 492-497

Phosphatidylinositol 3- and 4-kinases proteins.
amino acids 253-279

Tissue factor.
amino acids 1-278

FIGURE 85

MRAPGRPALRPLPLPPLLLLLLAAPWGRAVPCVSGGLPKPANITFLSINMKNVLQWTPPEG
LQGVKVTYTVQYFIYGQKKWLNKSECRNINRTYCDLSAETSDYEHQYYAKVKAIWGTKCSK
WAESGRFYPFLETQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIYSNLKYNV
SVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLKDQS
SEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILIYGNEFDKRFFVPAE
KIVINFITLNISDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHL
MEIFCDSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLQEEVST
QGTLLESQAALAVLGPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEEEPSTTLVDWDPQTGR
LCIPSLSSFDQDSEGCEPSEGDGLGEEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYV
QMEN

Signal sequence.
amino acids 1-29

Transmembrane domain.
amino acids 240-270

FIGURE 86

MAAARLCLSLLLLSTCVALLLQPLLGAQGAPLEPVYPGDNATPEQMAQYAADLRRYINMLT
RPRYGKRHKEDTLAFSEWGSPHAAVPRELSPLDL

Signal sequence.
amino acids 1-29

N-myristoylation site.
amino acids 80-85

Amidation site.
amino acids 65-68

Pancreatic hormone family signature.
amino acids 49-65

Pancreatic hormone peptide.
amino acids 30-65

FIGURE 87

MGPPSAPPCRLHVPWKEVLLTASLLTFWNPPTTAKLTIESTPFNVAEGKEVLLLAHNLPQN
RIGYSWYKGERVDGNSLIVGYVIGTQQATPGPAYSGRETIYPNASLLIQNVTQNDTGFYTL
QVIKSDLVNEEATGQFHVYPELPKPSISSNNSNPVEDKDAVAFTCEPEVQNTTYLWWVNGQ
SLPVSPRLQLSNGNMTLTLLSVKRNDAGSYECEIQNPASANRSDPVTLNVLYGPDGPTISP
SKANYRPGENLNLSCHAASNPPAQYSWFINGTFQQSTQELFIPNITVNNSGSYMCQAHNSA
TGLNRTTVTMITVSGSAPVLSAVATVGITIGVLARVALI

Signal sequence.
amino acids 1-34

Transmembrane domain.
amino acids 322-342

N-glycosylation sites.
amino acids 104-107, 111-114, 115-119, 152-155, 173-176, 197-200, 224-227, 256-259, 274-277, 288-291, 292-295, 309-312

Tyrosine kinase phosphorylation site.
amino acids 206-213

N-myristoylation sites.
amino acids 85-90, 211-216, 295-300, 307-312, 332-337

Immunoglobulin domain.
amino acids 160-217, 252-301

FIGURE 88

MELRVLLCWASLAAALEETLLNTKLETADLKWVTFPQVDGQWEELSGLDEEQHSVRTYEVC
DVQRAPGQAHWLRTGWVPRRGAVHVYATLRFTMLECLSLPRAGRSCKETFTVFYYESDADT
ATALTPAWMENPYIKVDTVAAEHLTRKRPGAEATGKVNVKTLRLGPLSKAGFYLAFQDQGA
CMALLSLHLFYKKCAQLTVNLTRFPETVPRELVVPVAGSCVVDAVPAPGPSPSLYCREDGQ
WAEQPVTGCSCAPGFEAAEGNTKCRACAQGTFKPLSGEGSCQPCPANSHSNTIGSAVCQCR
VGYFRARTDPRGAPCTTPPSAPRSVVSRLNGSSLHLEWSAPLESGGREDLTYALRCRECRP
GGSCAPCGGDLTFDPGPRDLVEPWVVVRGLRPDFTYTFEVTALNGVSSLATGPVPFEPVNV
TTDREVPPAVSDIRVTRSSPSSLSLAWAVPRAPSGAVLDYEVKYHEKGAEGPSSVRFLKTS
ENRAELRGLKRGASYLVQVRARSEAGYGPFGQEHHSQTQLDESEGWREQLALIAGTAVVGV
VLVLVVIVVAVLCLRKQSNGREAEYSDKHGQYLIGHGTKVYIDPFTYEDPNEAVREFAKEI
DVSYVKIEEVIGAGEFGEVCRGRLKAPGKKESCVAIKTLKGGYTERQRREFLSEASIMGQF
EHPNIIRLEGVVTNSMPVMILTEFMENGALDSFLRLNDGQFTVIQLVGMLRGIASGMRYLA
EMSYVHRDLAARNILVNSNLVCKVSDFGLSRFLEENSSDPTYTSSLGGKIPIRWTAPEAIA
FRKFTSASDAWSYGIVMWEVMSFGERPYWDMSNQDVINAIEQDYRLPPPPDCPTSLHQLML
DCWQKDRNARPRFPQVVSALDKMIRNPASLKIVARENGGASHPLLDQRQPHYSAFGSVGEW
LRAIKMGRYEESFAAAGFGSFELVSQISAEDLLRIGVTLAGHQKKILASVQHMKSQAKPGT
PGGTGGPAPQY

Signal sequence.
amino acids 1-15
Transmembrane domain.
amino acids 539-559
N-glycosylation sites.
amino acids 203-206, 335-338, 426-429, 768-771
Glycosaminoglycan attachment site.
amino acids 280-283
cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 564-567, 639-642, 795-798
Tyrosine kinase phosphorylation sites.
amino acids 508-515, 588-596, 608-614, 729-736
N-myristoylation sites.
amino acids 152-157, 182-187, 264-269, 298-303, 317-322, 374-379, 496-501, 681-686, 699-704, 719-724, 723-728,

951-956, 975-980
Amidation site.
amino acids 637-640
Tyrosine protein kinases specific active-site signature.
amino acids 736-748

FIGURE 89

MAQLFLPLLAALVLAQAPAALADVLEGDSSEDRAFRVRIAGDAPLQGVLGGALTIPCHVHY
LRPPPSRRAVLGSPRVKWTFLSRGREAEVLVARGVRVKVNEAYRFRVALPAYPASLTDVSL
ALSELRPNDSGIYRCEVQHGIDDSSDAVEVKVKGVVFLYREGSARYAFSFSGAQEACARIG
AHIATPEQLYAAYLGGYEQCDAGWLSDQTVRYPIQTPREACYGDMDGFPGVRNYGVVDPDD
LYDVYCYAEDLNGELFLGDPPEKLTLEEARAYCQERGAEIATTGQLYAAWDGGLDHCSPGW
LADGSVRYPIVTPSQRCGGGLPGVKTLFLFPNQTGFPNKHSRFNVYCFRDSAQPSAIPEAS
NPASNPASDGLEAIVTVTETLEELQLPQEATESESRGAIYSIPIMEDGGGGSSTPEDPAEA
PRTLLEFETQSMVPPTGFSEEEGKALEEEEKYEDEEEKEEEEEEEEVEDEALWAWPSELSS
PGPEASLPTEPAAQEKSLSQAPARAVLQPGASPLPDGESEASRPPRVHGPPTETLPTPRER
NLASPSPSTLVEAREVGEATGGPELSGVPRGESEETGSSEGAPSLLPATRAPEGTRELEAP
SEDNSGRTAPAGTSVQAQPVLPTDSASRGGVAVVPASGDCVPSPCHNGGTCLEEEEGVRCL
CLPGYGGDLCDVGLRFCNPGWDAFQGACYKHFSTRRSWEEAETQCRMYGAHLASISTPEEQ
DFINNRYREYQWIGLNDRTIEGDFLWSDGVPLLYENWNPGQPDSYFLSGENCVVMVWHDQG
QWSDVPCNYHLSYTCKMGLVSCGPPPELPLAQVFGRPRLRYEVDTVLRYRCREGLAQRNLP
LIRCQENGRWEAPQISCVPRRPARALHPEEDPEGRQGRLLGRWKALLIPPSSPMPGP

Signal peptide.
amino acids 1-15
N-glycosylation sites.
amino acids 130-133, 337-340
Tyrosine kinase phosphorylation sites.
amino acids 128-135, 451-459
N-myristoylation sites.
amino acids 47-52, 50-55, 133-138, 142-147, 174-179, 183-188, 281-286, 288-293, 297-302, 324-329, 403-408, 414-419, 415-420, 576-581, 586-591, 677-682, 684-689, 720-725, 772-777, 811-816
EGF-like domain cysteine pattern signature.
amino acids 670-682
C-type lectin domain signature.
amino acids 784-809
Immunoglobulins and major histocompatibility complex proteins signature.
amino acids 135-142

Link domain proteins.
amino acids 166-216, 264-314
Calcium-binding EGF-like domain proteins pattern proteins.
amino acids 655-676

FIGURE 90

MAQLFLPLLAALVLAQAPAALADVLEGDSSEDRAFRVRIAGDAPLQGVLGGALTIPCHVHY
LRPPPSRRAVLGSPRVKWTFLSRGREAEVLVARGVRVKVNEAYRFRVALPAYPASLTDVSL
ALSELRPNDSGIYRCEVQHGIDDSSDAVEVKVKGVVFLYREGSARYAFSFSGAQEACARIG
AHIATPEQLYAAYLGGYEQCDAGWLSDQTVRYPIQTPREACYGDMDGFPGVRNYGVVDPDD
LYDVYCYAEDLNGELFLGDPPEKLTLEEARAYCQERGAEIATTGQLYAAWDGGLDHCSPGW
LADGSVRYPIVTPSQRCGGGLPGVKTLFLFPNQTGFPNKHSRFNVYCFRDSAQPSAIPEAS
NPASNPASDGLEAIVTVTETLEELQLPQEATESESRGAIYSIPIMEDGGGGSSTPEDPAEA
PRTLLEFETQSMVPPTGFSEEEGKALEEEEKYEDEEEKEEEEEEEEVEDEALWAWPSELSS
PGPEASLPTEPAAQEESLSQAPARAVLQPGASPLPDGESEASRPPRVHGPPTETLPTPRER
NLASPSPSTLVEAREVGEATGGPELSGVPRGESEETGSSEGAPSLLPATRAPEGTRELEAP
SEDNSGRTAPAGTSVQAQPVLPTDSASRGGVAVVPASGNSAQGSTALSILLLFFPLQLWVT

Signal sequence.
amino acids 1-15

Transmembrane domain.
amino acids 652-670

N-glycosylation sites.
amino acids 130-133, 337-340

Tyrosine kinase phosphorylation sites.
amino acids 128-135, 451-459

N-myristoylation sites.
amino acids 47-52, 50-55, 133-138, 142-147, 174-179, 183-188, 281-286, 288-293, 297-302, 324-329, 403-408, 414-419, 415-420, 576-581, 586-591

Immunoglobulins and major histocompatibility complex.
amino acids 135-141

Extracellular link domain.
amino acids 156-251, 257-353

Immunoglobulin domain.
amino acids 50-139

FIGURE 91

MAVRQWVIALALAALLVVDREVPVAAGKLPFSRMPICEHMVESPTCSQMSNLVCGTDGLTY
TNECQLCLARIKTKQDIQIMKDGKC

Signal sequence.
amino acids 1-20

N-myristoylation site.
amino acids 58-63

Kazal-type serine protease inhibitor domain.
amino acids 37-86

FIGURE 92

MAPLCPSPWLPLLIPAPAPGLTVQLLLSLLLLVPVHPQRLPRMQEDSPLGGGSSGEDDPLG
EEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPEVKPKSEEEGSLKLEDLPTVEAPGDP
QEPQNNAHRDKEGDDQSHWRYGGDPPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPLELL
GFQLPPLPELRLRNNGHSVQLTLPPGLEMALGPGREYRALQLHLHWGAAGRPGSEHTVEGH
RFPAEIHVVHLSTAFARVDEALGRPGGLAVLAAFLEEGPEENSAYEQLLSRLEEIAEEGSE
TQVPGLDISALLPSDFSRYFRYEGSLTTPPCAQGVIWTVFNQTVMLSAKQLHTLSDTLWGP
GDSRLQLNFRATQPLNGRVIEASFPAGVDSSPRAAEPVQLNSCLAAGDILALVFGLLFAVT
SVAFLVQMRRQHRRGTKGGVSYRPAEVAETGA

Signal sequence.
amino acids 1-37

Transmembrane domain.
amino acids 409-429

N-glycosylation site.
amino acids 346-349

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 440-443

N-myristoylation sites.
amino acids 50-55, 51-56, 209-214, 236-241, 310-315, 339-344, 421-426, 442-447

Eukaryotic-type carbonic anhydrases signature.
amino acids 237-253

Eukaryotic-type carbonic anhydrase.
amino acids 141-390

FIGURE 93

MDALQLANSAFAVDLFKQLCEKEPLGNVLFSPICLSTSLSLAQVGAKGDTANEIGQVLHFE
NVKDVPFGFQTVTSDVNKLSSFYSLKLIKRLYVDKSLNLSTEFISSTKRPYAKELETVDFK
DKLEETKGQINNSIKDLTDGHFENILADNSVNDQTKILVVNAAYFVGKWMKKFSESETKEC
PFRVNKTDTKPVQMMNMEATFCMGNIDSINCKIIELPFQNKHLSMFILLPKDVEDESTGLE
KIEKQLNSESLSQWTNPSTMANAKVKLSIPKFKVEKMIDPKACLENLGLKHIFSEDTSDFS
GMSETKGVALSNVIHKVCLEITEDGGDSIEVPGARILQHKDELNADHPFIYIIRHNKTRNI
IFFGKFCSP

Signal sequence.
amino acids 1-42

N-glycosylation sites.
amino acids 99-102, 133-136, 188-191, 361-364

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 173-176

N-myristoylation sites.
amino acids 130-135, 207-212, 306-311, 312-317

Serpin (serine protease inhibitor).
amino acids 3-375

Fibrillarin proteins.
amino acids 150-169

FIGURE 94

MDRHSSYIFIWLQLELCAMAVLLTKGEIRCYCDAAHCVATGYMCKSELSACFSRLLDPQNS
NSPLTHGCLDSLASTTDICQAKQARNHSGTTIPTLECCHEDMCNYRGLHDVLSPPRGEASG
QGNRYQHDGSRNLITKVQELTSSKELWFRAAVIAVPIAGGLILVLLIMLALRMLRSENKRL
QDQRQQMLSRLHYSFHGHHSKKGQVAKLDLECMVPVSGHENCCLTCDKMRQADLSNDKILS
LVHWGMYSGHGKLEFV

Signal sequence.
amino acids 1-26

Transmembrane domain.
amino acids 152-172

N-glycosylation site.
amino acids 87-90

Glycosaminoglycan attachment sites.
amino acids 121-124, 252-255

N-myristoylation sites.
amino acids 68-73, 249-254

FIGURE 95

MRNRRNDTLDSTRTLYSSASRSTDLSYSESDLVNFIQANFKKRECVFFTKDSKATENVCKC
GYAQSQHMEGTQINQSEKWNYKKHTKEFPTDAFGDIQFETLGKKGKYIRLSCDTDAEILYE
LLTQHWHLKTPNLVISVTGGAKNFALKPRMRKIFSRLIYIAQSKGAWILTGGTHYGLTKYI
GEVVRDNTISRSSEENIVAIGIAAWGMVSNRDTLIRNCDAEGYFLAQYLMDDFTRDPLYIL
DNNHTHLLLVDNGCHGHPTVEAKLRNQLEKHISERTIQDSNYGGKIPIVCFAQGGGKETLK
AINTSIKNKIPCVVVEGSGRIADVIASLVEVEDAPTSSAVKEKLVRFLPRTVSRLSEEETE
SWIKWLKEILECSHLLTVIKMEEAGDEIVSNAISYALYKAFSTSEQDKDNWNGQLKLLLEW
NQLDLANDEIFTNDRRWESADLQEVMFTALIKDRPKFVRLFLENGLNLRKFLTHDVLTELF
SNHFSTLVYRNLQIAKNSYNDALLTFVWKLVANFRRGFRKEDRNGRDEMDIELHDVSPITR
HPLQALFIWAILQNKKELSKVIWEQTRGCTLAALGASKLLKTLAKVKNDINAAGESEELAN
EYETRAVELFTECYSSDEDLAEQLLVYSCEAWGGSNCLELAVEATDQHFTAQPGVQNFLSK
QWYGEISRDTKNWKIILCLFIIPLVGCGFVSFRKKPVDKHKKLLWYYVAFFTSPFVVFSWN
VVFYIAFLLLFAYVLLMDFHSVPHPPELVLYSLVFVLFCDEVRQWYVNGVNYFTDLWNVMD
TLGLFYFIAGIVFRLHSSNKSSLYSGRVIFCLDYIIFTLRLIHIFTVSRNLGPKIIMLQRM
LIDVFFFLFLFAVWMVAFGVARQGILRQNEQRWRWIFRSVIYEPYLAMFGQVPSDVDGTTY
DFAHCTFTGNESKPLCVELDEHNLPRFPEWITIPLVCIYMLSTNILLVNLLVAMFGYTVGT
VQENNDQVWKFQRYFLVQEYCSRLNIPFPFIVFAYFYMVVKKCFKCCCKEKNMESSVCCFK
NEDNETLAWEGVMKENYLVKINTKANDTSEEMRHRFRQLDTKLNDLKGLLKELANKIK

Transmembrane domains.
amino acids 681-701, 718-738, 734-754, 757-777, 784-804, 819-839, 853-873, 951-971, 995-1015

N-glycosylation sites.
amino acids 6-9, 75-78, 247-250, 308-311, 812-815, 925-928, 1041-1044, 1063-1066

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 83-86

Tyrosine kinase phosphorylation sites.
amino acids 21-27, 219-226, 279-286

N-myristoylation sites.

577-582, 643-648, 904-909

Amidation site.
amino acids 102-105

FIGURE 96

MRNRRNDTLDSTRTLYSSASRSTDLSYSESDLVNFIQANFKKRECVFFTKDSKATENVCKC
GYAQSQHMEGTQINQSEKWNYKKHTKEFPTDAFGDIQFETLGKKGKYIRLSCDTDAEILYE
LLTQHWHLKTPNLVISVTGGAKNFALKPRMRKIFSRLIYIAQSKGAWILTGGTHYGLMKYI
GEVVRDNTISRSSEENIVAIGIAAWGMVSNRDTLIRNCDAEGYFLAQYLMDDFTRDPLYIL
DNNHTHLLLVDNGCHGHPTVEAKLRNQLEKYISERTIQDSNYGGKIPIVCFAQGGGKETLK
AINTSIKNKIPCVVVEGSGQIADVIASLVEVEDALTSSAVKEKLVRFLPRTVSRLPEEETE
SWIKWLKEILECSHLLTVIKMEEAGDEIVSNAISYALYKAFSTSEQDKDNWNGQLKLLLEW
NQLDLANDEIFTNDRRWESADLQEVMFTALIKDRPKFVRLFLENGLNLRKFLTHDVLTELF
SNHFSTLVYRNLQIAKNSYNDALLTFVWKLVANFRRGFRKEDRNGRDEMDIELHDVSPITR
HPLQALFIWAILQNKKELSKVIWEQTRGCTLAALGASKLLKTLAKVKNDINAAGESEELAN
EYETRAVELFTECYSSDEDLAEQLLVYSCEAWGGSNCLELAVEATDQHFIAQPGVQNFLSK
QWYGEISRDTKNWKIILCLFIIPLVGCGFVSFRKKPVDKHKKLLWYYVAFFTSPFVVFSWN
VVFYIAFLLLFAYVLLMDFHSVPHPPELVLYSLVFVLFCDEVRQWYVNGVNYFTDLWNVMD
TLGLFYFIAGIVFRLHSSNKSSLYSGRVIFCLDYIIFTLRLIHIFTVSRNLGPKIIMLQRM
LIDVFFFLFLFAVWMVAFGVARQGILRQNEQRWRWIFRSVIYEPYLAMFGQVPSDVDGTTY
DFAHCTFTGNESKPLCVELDEHNLPRFPEWITIPLVCIYMLSTNILLVNLLVAMFGYTVGT
VQENNDQVWKFQRYFLVQEYCSRLNIPFPFIVFAYFYMVVKKCFKCCCKEKNMESSVCCFK
NEDNETLAWEGVMKENYLVKINTKANDTSEEMRHRFRQLDTKLNDLKGLLKEIANKIK

Transmembrane domains.
amino acids 681-701, 718-738, 734-754, 757-777, 784-804, 819-839, 853-873, 951-971, 995-1015

N-glycosylation sites.
amino acids 6-9, 75-78, 247-250, 308-311, 812-815, 925-928, 1041-1044, 1063-1066

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 83-86

Tyrosine kinase phosphorylation sites.
amino acids 21-27, 219-226, 279-286

N-myristoylation sites.

577-582, 643-648, 904-909

Amidation site.
amino acids 102-105

FIGURE 97

RRGRGEVLAVERGSGSGGGGTRSGWPAPAAGADKKPSRCGSGREGEGVSEGHKSMTGLYEL
VWRVLHALLCLHRTLTSWLRVRFGTWNWIWRRCCRAASAAVLAPLGFTLRKPPAVGRNRRH
HRHPRGGSCLAAAHHRMRWRADGRSLEKLPVHMGLVITEVEQEPSFSDIASLVVWCMAVGI
SYISVYDHQGIFKRNNSRLMDEILKQQQELLGLDCSKYSPEFANSNDKDDQVLNCHLAVKV
LSPEDGKADIVRAAQDFCQLVAQKQRRPTDLDVDTLASLLSSNGCPDPDLVLKFGPVDSTL
GFLPWHIRLTEIVSLPSHLNISYEDFFSALRQYAACEQRLGK

Transmembrane domain.
amino acids 166-186

N-glycosylation sites.
amino acids 198-201, 325-328

Glycosaminoglycan attachment sites.
amino acids 14-17, 16-19

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 34-37, 270-273

N-myristoylation sites.
amino acids 13-18, 15-20, 17-22, 19-24, 20-25, 47-52, 129-134, 156-161, 215-220

FIGURE 98

MTGLYELVWRVLHALLCLHRTLTSWLRVRFGTWNWIWRRCCRAASAAVLAPLGFTLRKPPA
VGRNRRHHRHPRGGSCLAAAHHRMRWRADGRSLEKLPVHMGLVITEVEQEPSFSDIASLVV
WCMAVGISYISVYDHQGIFKRNNSRLMDEILKQQQELLGLDCSKYSPEFANSNDKDDQVLN
CHLAVKVLSPEDGKADIVRAAQDFCQLVAQKQKRPTDLDVDTLASLLSSNGCPDPDLVLKF
GPVDSTLGFLPWHIRLTEIVSLPSHLNISYEDFFSALRQYAACEQRLGK

Signal sequence.
amino acids 1-23

Transmembrane domain.
amino acids 115-132

N-glycosylation sites.
amino acids 144-147, 271-274

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 216-219

N-myristoylation sites.
amino acids 75-80, 102-107, 161-166

FIGURE 99

MVQRLWVSRLLRHRKAQLLLVNLLTFGLEVCLAAGITYVPPLLLEVGVEEKFMTMVLGIGP
VLGLVCVPLLGSASDHWRGRYGRRRPFIWALSLGILLSLFLIPRAGWLAGLLCPDPRPLEL
ALLILGVGLLDFCGQVCFTPLEALLSDLFRDPDHCRQAYSVYAFMISLGGCLGYLLPAIDW
DTSALAPYLGTQEECLFGLLTLIFLTCVAATLLVAEEAALGPTEPAEGLSAPSLSPHCCPC
RARLAFRNLGALLPRLHQLCCRMPRTLRRLFVAELCSWMALMTFTLFYTDFVGEGLYQGVP
RAEPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFGTRAVYLASVAAFPVAA
GATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGGASSEDSL
MTSFLPGPKPGAPFPNGHVGAGGSGLLPPPPALCGASACDVSVRVVVGEPTEARVVPGRGI
CLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQVVFDKSDLA
KYSA

Signal sequence.
amino acids 1-35

Transmembrane domains.
amino acids 15-35, 52-72, 87-107, 122-142, 160-180, 240-260, 277-297, 321-341, 486-506, 521-541

N-myristoylation sites.
amino acids 27-32, 95-100, 171-176, 299-304, 303-308, 320-325, 419-424, 447-452, 462-467

Amidation site.
amino acids 82-85

Cell attachment sequence.
amino acids 415-417

FIGURE 100

MALRRLGAALLLLPLLAAVEETLMDSTTATAELGWMVHPPSGWEEVSGYDENMNTIRTYQV
CNVFESSQNNWLRTKFIRRRGAHRIHVEMKFSVRDCSSIPSVPGSCKETFNLYYYEADFDS
ATKTFPNWMENPWVKVDTIAADESFSQVDLGGRVMKINTEVRSFGPVSRSGFYLAFQDYGG
CMSLIAVRVFYRKCPRIIQNGAIFQETLSGAESTSLVAARGSCIANAEEVDVPIKLYCNGD
GEWLVPIGRCMCKAGFEAVENGTVCRGCPSGTFKANQGDEACTHCPINSRTTSEGATNCVC
RNGYYRADLDPLDMPCTTIPSAPQAVISSVNETSLMLEWTPPRDSGGREDLVYNIICKSCG
SGRGACTRCGDNVQYAPRQLGLTEPRIYISDLLAHTQYTFEIQAVNGVTDQSPFSPQFASV
NITTNQAAPSAVSIMHQVSRTVDSITLSWSQPDQPNGVILDYELQYYEKELSEYNATAIKS
PTNTVTVQGLKAGAIYVFQVRARTVAGYGRYSGKMYFQTMTEAEYQTSIQEKLPLIIGSSA
AGLVFLIAVVVIAIVCNRRRGFERADSEYTDKLQHYTSGHMTPGMKIYIDPFTYEDPNEAV
REFAKEIDISCVKIEQVIGAGEFGEVCSGHLKLPGKREIFVAIKTLKSGYTEKQRRDFLSE
ASIMGQFDHPNVIHLEGVVTKSTPVMIITEFMENGSLDSFLRQNDGQFTVIQLVGMLRGIA
AGMKYLADMNYVHRDLAARNILVNSNLVCKVSDFGLSRFLEDDTSDPTYTSALGGKIPIRW
TAPEAIQYRKFTSASDVWSYGIVMWEVMSYGERPYWDMTNQDVINAIEQDYRLPPPMDCPS
ALHQLMLDCWQKDRNHRPKFGQIVNTLDKMIRNPNSLKAMAPLSSGINLPLLDRTIPDYTS
FNTVDEWLEAIKMGQYKESFANAGFTSFDVVSQMMMEDILRVGLTLAGHQKKILNSIQVMR
AQMNQIQSVEV

Signal sequence.

amino acids 1-18

Transmembrane domain.

amino acids 542-562

N-glycosylation sites.

amino acids 265-268, 336-339, 428-431, 482-485, 705-708

Glycosaminoglycan attachment site.

amino acids 367-370

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 802-805

Tyrosine kinase phosphorylation sites.

amino acids 374-381, 595-603, 736-743

N-myristoylation sites.

amino acids 182-187, 213-218, 224-229, 271-276, 275-280, 299-304, 366-371, 497-502, 546-551, 706-711, 726-731, 730-735, 875-880, 958-963

Amidation site - amino acids 644-647

Tyrosine protein kinases specific active-site signature - amino acids 743-755.

Receptor tyrosine kinase class V signature 1 - amino acids 182-198

Receptor tyrosine kinase class V signature 2 - amino acids 241-261

Ephrin receptor ligand binding domain - amino acids 20-197

Protein kinase domain - amino acids 622-881

Fibronectin type III domain - amino acids 325-421, 436-520

SAM domain (Sterile alpha motif) - amino acids 912-976

FIGURE 101

MALRRLGAALLLLPLLAAVEETLMDSTTATAELGWMVHPPSGWEEVSGYDENMNTIRTYQV
CNVFESSQNNWLRTKFIRRRGAHRIHVEMKFSVRDCSSIPSVPGSCKETFNLYYYEADFDS
ATKTFPNWMENPWVKVDTIAADESFSQVDLGGRVMKINTEVRSFGPVSRSGFYLAFQDYGG
CMSLIAVRVFYRKCPRIIQNGAIFQETLSGAESTSLVAARGSCIANAEEVDVPIKLYCNGD
GEWLVPIGRCMCKAGFEAVENGTVCRGCPSGTFKANQGDEACTHCPINSRTTSEGATNCVC
RNGYYRADLDPLDMPCTTIPSAPQAVISSVNETSLMLEWTPPRDSGGREDLVYNIICKSCG
SGRGACTRCGDNVQYAPRQLGLTEPRIYISDLLAHTQYTFEIQAVNGVTDQSPFSPQFASV
NITTNQAAPSAVSIMHQVSRTVDSITLSWSQPDQPNGVILDYELQYYEKELSEYNATAIKS
PTNTVTVQGLKAGAIYVFQVRARTVAGYGRYSGKMYFQTMTEAEYQTSIQEKLPLIIGSSA
AGLVFLIAVVVIAIVCNRRGFERADSEYTDKLQHYTSGHMTPGMKIYIDPFTYEDPNEAVR
EFAKEIDISCVKIEQVIGAGEFGEVCSGHLKLPGKREIFVAIKTLKSGYTEKQRRDFLSEA
SIMGQFDHPNVIHLEGVVTKSTPVMIITEFMENGSLDSFLRQNDGQFTVIQLVGMLRGIAA
GMKYLADMNYVHRDLAARNILVNSNLVCKVSDFGLSRFLEDDTSDPTYTSALGGKIPIRWT
APEAIQYRKFTSASDVWSYGIVMWEVMSYGERPYWDMTNQDVINAIEQDYRLPPPMDCPSA
LHQLMLDCWQKDRNHRPKFGQIVNTLDKMIRNPNSLKAMAPLSSGINLPLLDRTIPDYTSF
NTVDEWLEAIKMGQYKESFANAGFTSFDVVSQMMMEDILRVGVTLAGHQKKILNSIQVMRA
QMNQIQSVEV

Signal sequence.
amino acids 1-18
Transmembrane domain.
amino acids 543-563
N-glycosylation sites.
amino acids 265-268, 336-339, 428-431, 482-485, 704-707
Glycosaminoglycan attachment site.
amino acids 367-370
cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 801-804
Tyrosine kinase phosphorylation sites.
amino acids 374-381, 594-602, 735-742
N-myristoylation sites.
amino acids 182-187, 213-218, 224-229, 271-276, 275-280, 299-304, 366-371, 497-502, 546-551, 705-710, 725-730, 729-734, 874-879, 957-962

Amidation site.
amino acids 643-646
Tyrosine protein kinases specific active-site signature.
amino acids 742-754
Receptor tyrosine kinase class V signature 1.
amino acids 182-198
Receptor tyrosine kinase class V signature 2.
amino acids 241-264
Ephrin receptor ligand binding domain.
amino acids 20-197
Protein kinase domain.
amino acids 621-880
Fibronectin type III domains.
amino acids 325-421, 436-520
SAM domain (Sterile alpha motif).
amino acids 911-975
Src homology 2 (SH2) domains.
amino acids 733-769, 787-797, 806-819

FIGURE 102

MNFSTSSSSFAYDREFLRTLPGFLIVAEIVLGLLVWTLIAGTEYFRVPAFGWVMFVAVFYW
VLTVFFLIIYITMTYTRIPQVPWTTVGLCFNGSAFVLYLSAAVVDASSVSPERDSHNFNSW
AASSFFAFLVNICYAGNTYFSFIAWRSRTIQ

Signal sequence.
amino acids 1-41

Transmembrane domains.
amino acids 53-73, 86-106, 122-142

N-glycosylation sites.
amino acids 2-5, 92-95

N-myristoylation site.
amino acids 88-93

Leucine zipper pattern.
amino acids 17-38

FIGURE 103

MNFSTSSSSFAYDREFLRTLPGFLIVAEIVLGLLVWTLIAGTEYFRVPAFGWVMFVAVFYW
VLTVFFLIIYITMTYTRIPQVPWTTVGLCFNGSAFVLYLSAAVVDASSVSPERDSHNFNSW
AASSFFAFLVTICYAGNTYFSFIAWRSRTIQ

Transmembrane domains.
amino acids 19-39, 53-73, 86-106, 122-142

N-glycosylation sites.
amino acids 2-5, 92-95

N-myristoylation site.
amino acids 88-93

Leucine zipper pattern.
amino acids 17-38

FIGURE 104

MPTTVDDVLEHGGEFHFFQKQMFFLLALLSATFAPIYVGIVFLGFTPDHRCRSPGVAELSL
RCGWSPAEELNYTVPGPGPAGEASPRQCRRYEVDWNQSTFDCVDPLASLDTNRSRLPLGPC
RDGWVYETPGSSIVTEFNLVCANSWMLDLFQSSVNVGFFIGSMSIGYIADRFGRKLCLLTT
VLINAAAGVLMAISPTYTWMLIFRLIQGLVSKAGWLIGYILITEFVGRRYRRTVGIFYQVA
YTVGLLVLAGVAYALPHWRWLQFTVALPNFFFLLYYWCIPESPRWLISQNKNAEAMRIIKH
IAKKNGKSLPASLQRLRLEEETGKKLNPSFLDLVRTPQIRKHTMILMYNWFTSSVLYQGLI
MHMGLAGDNIYLDFFYSALVEFPAAFMIILTIDRIGRRYPWAASNMVAGAACLASVFIPGD
LQWLKIIISCLGRMGITMAYEIVCLVNAELYPTFIRNLGVHICSSMCDIGGIITPFLVYRL
TNIWLELPLMVFGVLGLVAGGLVLLLPETKGKALPETIEEAENMQRPRKNKEKMIYLQVQK
LDIPLN

Transmembrane domains.
amino acids 20-40, 174-194, 205-225, 238-258, 269-289, 351-371, 376-396, 412-432, 492-512

N-glycosylation sites.
amino acids 72-75, 97-100, 113-116

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 345-348

Tyrosine kinase phosphorylation site.
amino acids 536-553

N-myristoylation sites.
amino acids 191-196, 254-259, 442-447, 466-471, 504-509

Amidation site.
amino acids 174-177, 229-232, 327-330, 401-404

Sugar (and other) transporter.
amino acids 23-529

FIGURE 105

MPTTVDDVLEHGGEFHFFQKQMFFLLALLSATFAPIYVGIVFLGFTPDHRCRSPGVAELSL
RCGWSPAEELNYTVPGPGPAGEASPRQCRRYEVDWNQSTFDCVDPLASLDTNRSRLPLGPC
RDGWVYETPGSSIVTEFNLVCANSWMLDLFQSSVNVGFFIGSMSIGYIADRFGRKLCLLTT
VLINAAAGVLMAISPTYTWMLIFRLIQGLVSKAGWLIGYILITEFVGRRYRRTVGIFYQVA
YTVGLLVLAGVAYALPHWRWLQFTVSLPNFFFLLYYWCIPESPRWLISQNKNAEAMRIIKH
IAKKNGKSLPASLQRLRLEEETGKKLNPSFLDLVRTPQIRKHTMILMYNWFTSSVLYQGLI
MHMGLAGDNIYLDFFYSALVEFPAAFMIILTIDRIGRRYPWAASNMVAGAACLASVFIPGD
LQWLKIIISCLGRMGITMAYEIVCLVNAELYPTFIRNLGVHICSSMCDIGGIITPFLVYRL
TNIWLELPLMVFGVLGLVAGGLVLLLPETKGKALPETIEEAENMQRPRKNKEKMIYLQVQK
LDIPLN

Transmembrane domains.
amino acids 20-40, 174-194, 205-225, 238-258, 269-289, 351-371, 376-396, 412-432, 492-512

N-glycosylation sites.
amino acids 72-75, 97-100, 113-116

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 345-348

Tyrosine kinase phosphorylation site.
amino acids 536-544

N-myristoylation sites.
amino acids 191-196, 254-259, 442-447, 466-471, 504-509

Amidation sites.
amino acids 174-177, 229-232, 327-330, 401-404

Sugar (and other) transporter.
amino acids 23-529

FIGURE 106

MAPWPELGDAQPNPDKYLEGAAGQQPTAPDKSKETNKTDNTEAPVTKIELLPSYSTATLID
EPTEVDDPWNLPTLQDSGIKWSERDTKGKILCFFQGIGRLILLLGFLYFFVCSLDILSSAF
QLVGGKMAGQFFSNSSIMSNPLLGLVIGVLVTVLVQSSSTSTSIVVSMVSSSLLTVRAAIP
IIMGANIGTSITNTIVALMQVGDRSEFRRAFAGATVHDFFNWLSVLVLLPVEVATHYLEII
TQLIVESFHFKNGEDAPDLLKVITKPFTKLIVQLDKKVISQIAMNDEKAKNKSLVKIWCKT
FTNKTQINVTVPSTANCTSPSLCWTDGIQNWTMKNVTYKENIAKCQHIFVNFHLPDLAVGT
ILLILSLLVLCGCLIMIVKILGSVLKGQVATVIKKTINTDFPFPFAWLTGYLAILVGAGMT
FIVQSSSVFTSALTPLIGIGVITIERAYPLTLGSNIGTTTTAILAALASPGNALRSSLQIA
LCHFFFNISGILLWYPIPFTRLPIRMAKGLGNISAKYRWFAVFYLIIFFFLIPLTVFGLSL
AGWRVLVGVGVPVVFIIILVLCLRLLQSRCPRVLPKKLQNWNFLPLWMRSLKPWDAVVSKF
TGCFQMRCCYCCRVCCRACCLLCGCPKCCRCSKCCEDLEEAQEGQDVPVKAPETFDNITIS
REAQGEVPASDSKTECTAL

Transmembrane domains.
amino acids 96-116, 136-156, 178-198, 219-239, 356-376, 372-392, 406-426, 445-465, 488-508, 523-543, 549-569, 565-585, 592-612

N-glycosylation sites.
amino acids 36-39, 136-139, 295-298, 308-311, 313-316, 321-324, 335-338, 340-343, 495-498, 520-532, 667-670

N-myristoylation sites.
amino acids 23-28, 79-84, 126-131, 131-136, 146-151, 150-155, 187-192, 191-196, 393-398, 423-428, 460-465, 464-469, 519-524, 546-551, 634-639

4Fe-4S ferredoxins, iron-sulfur binding region signature.
amino acids 635-646

Insulin family signature.
amino acids 621-635

amino acids 617-633, 625-641, 613-629

Na+/Pi-cotransporter.
amino acids 118-549

FIGURE 107

MAPWPELGDAQPNPDKYLEGAAGQQPTAPDKSKETNKTDNTEAPVTKIELLPSYSTATLID
EPTEVDDPWNLPTLQDSGIKWSERDTKGKILCFFQGIGRLILLLGFLYFFVCSLDILSSAF
QLVGGKMAGQFFSNSSIMSNPLLGLVIGVLVTVLVQSSSTSTSIVVSMVSSSLLTVRAAIP
IIMGANIGTSITNTIVALMQVGDRSEFRRAFAGATVHDFFNWLSVLVLLPVEVATHYLEII
TQLIVESFHFKNGEDAPDLLKVITKPFTKLIVQLDKKVISQIAMNDEKAKNKSLVKIWCKT
FTNKTQINVTVPSTANCTSPSLCWTDGIQNWTMKNVTYKENIAKCQHIFVNFHLPDLAVGT
ILLILSLLVLCGCLIMIVKILGSVLKGQVATVIKKTINTDFPFPFAWLTGYLAILVGAGMT
FIVQSSSVFTSALTPLIGIGVITIERAYPLTLGSNIGTTTTAILAALASPGNALRSSLQIA
LCHFFFNISGILLWYPIPFTRLPIRMAKGLGNISAKYRWFAVFYLIIFFFLIPLTVFGLSL
AGWRVLVGVGVPVVFIIILVLCLRLLQSRCPRVLPKKLQNWNFLPLWMRSLKPWDAVVSKF
TGCFQMRCCCCCRVCCRACCLLCGCPKCCRCSKCCEDLEEAQEGQDVPVKAPETFDNITIS
REAQGEVPASDSKTECTAL

Transmembrane domains.
amino acids 96-116, 136-156, 178-198, 219-239, 356-376,
372-392, 406-426, 445-465, 488-508, 523-543, 547-567,
563-583, 592-612

N-glycosylation sites.
amino acids 36-39, 136-139, 295-298, 308-311, 313-316,
321-324, 335-338, 340-343, 495-498, 520-523, 667-670

N-myristoylation sites.
amino acids 23-28, 79-84, 126-131, 131-136, 146-151, 150-155,
187-192, 191-196, 393-398, 423-428, 460-465, 464-469,
519-524, 546-551, 634-639

4Fe-4S ferredoxins, iron-sulfur binding region signature.
amino acids 635-645

Insulin family signature.
amino acids 621-635

Na+/Pi-cotransporter.

FIGURE 108

MWSGWWLWPLVAVCTADFFRDEAERIMRDSPVIDGHNDLPWQLLDMFNNRLQDERANLTTL
AGTHTNIPKLRAGFVGGQFWSVYTPCDTQNKDAVRRTLEQMDVVHRMCRMYPETFLYVTSS
AGIRQAFREGKVASLIGVEGGHSIDSSLGVLRALYQLGMRYLTLTHSCNTPWADNWLVDTG
DSEPQSQGLSPFGQRVVKELNRLGVLIDLAHVSVATMKATLQLSRAPVIFSHSSAYSVCAS
RRNVPDDVLRLVKQTDSLVMVNFYNNYISCTNKANLSQVADHLDHIKEVAGARAVGFGGDF
DGVPRVPEGLEDVSKYPDLIAELLRRNWTEAEVKGALADNLLRVFQAVEQASNLTQAPEER
PIPLDQLGGSCRTHYGYSSGASSLHRHWGLLLASLAPLVLCLSLL

Signal sequence.
amino acids 1-16

Transmembrane domain.
amino acids 392-410

N-glycosylation sites.
amino acids 57-60, 279-282, 332-335, 358-361

Tyrosine kinase phosphorylation site.
amino acids 110-118

N-myristoylation sites.
amino acids 63-68, 78-83, 124-129, 139-144, 151-156, 375-380, 395-400

Leucine zipper pattern.
amino acids 390-411

Renal dipeptidase active site.
amino acids 140-162

Renal dipeptidase.
amino acids 1-411

FIGURE 109

MSSCNFTHATFVLIGIPGLEKAHFWVGFPLLSMYVVAMFGNCIVVFIVRTERSLHAPMYLF
LCMLAAIDLALSTSTMPKILALFWFDSREISFEACLTQMFFIHALSAIESTILLAMAFDRY
VAICHPLRHAAVLNNTVTAQIGIVAVVRGSLFFFPLPLLIKRLAFCHSNVLSHSYCVHQDV
MKLAYADTLPNVVYGLTAILLVMGVDVMFISLSYFLIIRTVLQLPSKSERAKAFGTCVSHI
GVVLAFYVPLIGLSVVHRFGNSLHPIVRVVMGDIYLLLPPVINPIIYGAKTKQIRTRVLAM
FKISCDKDLQAVGGK

Transmembrane domains.
amino acids 26-46, 57-77, 96-116, 144-164, 197-217, 238-258, 272-292

N-glycosylation sites.
amino acids 5-8, 136-139

N-myristoylation site.
amino acids 18-23, 238-243, 245-250

7 transmembrane receptor (rhodopsin family).
amino acids 40-131, 212-291

FIGURE 110A

MRILKRFLACIQLLCVCRLDWANGYYRQQRKLVEEIGWSYTGALNQKNWGKKYPTCNSPKQ
SPINIDRDLTQVNVNLKKLKFQGWDKTSLENTFIHNTGKTVEINLTNDYRVSGGVSEMVFK
ASKITFHWGKCNMSSDGSEHSLEGQKFPLEMQIYCFDADRFSSFEEAVKGKGKLRALSILF
EVGTEENLDFKAIIDGVESVSRFGKQAALDPFILLNLLPNSTDKYYIYNGSLTSPPCTDTV
DWIVFKDTVSISESQLAVFCEVLTMQQSGYVMLMDYLQNNFREQQYKFSRQVFSSYTGKEE
IHEAVCSSEPENVQADPENYTSLLVTWERPRVVYDTMIEKFAVLYQQLDGEDQTKHEFLTD
GYQDLGAILNNLLPNMSYVLQIVAICTNGLYGKYSDQLIVDMPTDNPELDLFPELIGTEEI
IKEEEEGKDIEEGAIVNPGRDSATNQIRKKEPQISTTTHYNRIGTKYNEAKTNRSPTRGSE
FSGKGDVPNTSLNSTSQPVTKLATEKDISLTSQTVTELPPHTVEGTSASLNDGSKTVLRSP
HMNLSGTAESLNTVSITEYEEESLLTSFKLDTGAEDSSGSSPATSAIPFISENISQGYIFS
SENPETITYDVLIPESARNASEDSTSSGSEESLKDPSMEGNVWFPSSTDITAQPDVGSGRE
SFLQTNYTEIRVDESEKTTKSFSAGPVMSQGPSVTDLEMPHYSTFAYFPTEVTPHAFTPSS
RQQDLVSTVNVVYSQTTQPVYNGETPLQPSYSSEVFPLVTPLLLDNQILNTTPAASSSDSA
LHATPVFPSVDVSFESILSSYDGAPLLPFSSASFSSELFRHLHTVSQILPQVTSATESDKV
PLHASLPVAGGDLLLEPSLAQYSDVLSTTHAASETLEFGSESGVLYKTLMFSQVEPPSSDA
MMHARSSGPEPSYALSDNEGSQHIFTVSYSSAIPVHDSVGVTYQGSLFSGPSHIPIPKSSL
ITPTASLLQPTHALSGDGEWSGASSDSEFLLPDTDGLTALNISSPVSVAEFTYTTSVFGDD
NKALSKSEIIYGNETELQIPSFNEMVYPSESTVMPNMYDNVNKLNASLQETSVSISSTKGM
FPGSLAHTTTKVFDHEISQVPENNFSVQPTHTVSQASGDTSLKPVLSANSEPASSDPASSE
MLSPSTQLLFYETSASFSTEVLLQPSFQASDVDTLLKTVLPAVPSDPILVETPKVDKISST
MLHLIVSNSASSENMLHSTSVPVFDVSPTSHMHSASLQGLTISYASEKYEPVLLKSESSHQ
VVPSLYSNDELFQTANLEINQAHPPKGRHVFATPVLSIDEPLNTLINKLIHSDEILTSTKS
SVTGKVFAGIPTVASDTFVSTDHSVPIGNGHVAITAVSPHRDGSVTSTKLLFPSKATSELS
HSAKSDAGLVGGGEDGDTDDDGDDDDDRDSDGLSIHKCMSCSSYRESQEKVMNDSDTHENS
LMDQNNPISYSLSENSEEDNRVTSVSSDSQTGMDRSPGKSPSANGLSQKHNDGKEENDIQT
GSALLPLSPESKAWAVLTSDEESGSGQGTSDSLNENETSTDFSFADTNEKDADGILAAGDS
EITPGFPQSPTSSVTSENSEVFHVSEAEASNSSHESRIGLAEGLESEKKAVIPLVIVSALT
FICLVVLVGILIYWRKCFQTAHFYLEDSTSPRVISTPPTPIFPISDDVGAIPIKHFPKHVA
DLHASSGFTEEFETLKEFYQEVQSCTVDLGITADSSNHPDNKHKNRYINIVAYDHSRVKLA
QLAEKDGKLTDYINANYVDGYNRPKAYIAAQGPLKSTAEDFWRMIWEHNVEVIVMITNLVE
KGRRKCDQYWPADGSEEYGNFLVTQKSVQVLAYYTVRNFTLRNTKIKKGSQKGRPSGRVVT
QYHYTQWPDMGVPEYSLPVLTFVRKAAYAKRHAVGPVVVHCSAGVGRTGTYIVLDSMLQQI
QHEGTVNIFGFLKHIRSQRNYLVQTEEQYVFIHDTLVEAILSKETEVLDSHIHAYVNALLI
PGPAGKTKLEKQFQLLSQSNIQQSDYSAALKQCNREKNRTSSIIPVERSRVGISSLSGEGT
DYINASYIMGYYQSNEFIITQHPLLHTIKDFWRMIWDHNAQLVVMIPDGQNMAEDEFVYWP

NKDEPINCESFKVTLMAEEHKCLSNEEKLIIQDFILEATQDDYVLEVRHFQCPKWPNPDSP
ISKTFELISVIKEEAANRDGPMIVHDEHGGVTAGTFCALTTLMHQLEKENSVDVYQVAKMI
NLMRPGVFADIEQYQFLYKVILSLVSTRQEENPSTSLDSNGAALPDGNIAESLESLV

Signal sequence.
amino acids 1-19
Transmembrane domain.
amino acids 1638-1658
N-glycosylation sites.
amino acids 105-108, 134-137, 223-226, 232-235, 324-327, 381-384, 497-500, 501-504, 552-555, 602-605, 629-632, 677-680, 1017-1020, 1050-1053, 1082-1085, 1122-1125, 1456-1459, 1561-1564, 1617-1620, 1868-1871, 2051-2054, 2078-2081
Glycosaminoglycan attachment sites.
amino acids 490-493, 991-994, 1548-1551, 1550-1553, 2070-2073
cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 1877-1880
Tyrosine kinase phosphorylation site.
amino acids 148-156

FIGURE 110B

N-myristoylation sites.
amino acids 139-144, 186-191, 372-377, 471-476, 486-491, 533-538, 555-560, 582-587, 588-593, 638-643, 893-898, 960-965, 1097-1102, 1259-1264, 1385-1390, 1411-1416, 1415-1420, 1549-1554, 1551-1556, 1553-1558, 1579-1584, 1625-1630, 1879-1884, 1935-1940, 2123-2128, 2225-2230, 2226-2231, 2230-2235

Amidation sites.
amino acids 49-52, 1831-1834

Tyrosine specific protein phosphatases active site.
amino acids 1930-1942

Protein-tyrosine phosphatase.
amino acids 1749-1990, 2047-2280

Eukaryotic-type carbonic anhydrase.
amino acids 38-300

Fibronectin type III domain.
amino acids 312-401

MAM domain proteins.
amino acids 1757-1785, 1060-1088, 1813-1853, 2047-2079

FIGURE 111

MCARMAGRTRAAPRGPYGPWLCLLVALALDVVRVDCGQAPLDFVYLHVTAARPAQPTLWTA
KLDRFKGSRHHTTLITCHRAGLTEPDSSSPLELSEFLWVDFVVENSTGGGVAVTRPVTWQL
EYPGQAPEAEKDKMVWEILVSERDIRALIPLAKAEELVNTAPLTGVPQHVPVRLVTVDGGG
ALVEVTEHVGCESANTQVLQVSEACDAVFVAGKESRGARGVRVDFWWRRLRASLRLTVWAP
LLPLRIELTDTTLEQVRGWRVPGPAEGPAEPAAEASDEAERRARGCHLQYQRAGVRFLAPF
AAHPLDGGRRLTHLLGPDWLLDVSHLVAPHARVLDSRVASLEGGRVVVGREPGVTSIEVRS
PLSDSILGEQALAVTDDKVSVLELRVQPVMGISLTLSRGTAHPGEVTATCWAQSALPAPKQ
EVALSLWLSFSDHTVAPAELYDRRDLGLSVSAEEPGAILPAEEQGAQLGVVVSGAGAEGLP
LHVALHPPEPCRRGRHRVPLASGTAWLGLPPASTPAPALPSSPAWSPPATEATMGGKRQVA
GSVGGNTGVRGKFERAEEEARKEETEAREEEEEEEEEMVPAPQHVTELELGMYALLGVFCV
AIFIFLVNGVVFVLRYQRKEPPDSATDPTSPQPHNWVWLGTDQEELSRQLDRQSPGPPKGE
GSCPCESGGGGEAPTLAPGPPGGTTSSSSTLARKEAGGRRKRVEFVTFVPAPPAQSPEEPV
GAPAVQSILVAGEEDIRWVCEDMGLKDPEELRNYMERIRGSS

N-glycosylation site.
amino acids 106-109

Glycosaminoglycan attachment sites.
amino acids 480-483, 678-681

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 314-317

Tyrosine kinase phosphorylation site.
amino acids 758-766

N-myristoylation sites.
amino acids 109-114, 111-116, 193-198, 397-402, 454-459, 472-477, 476-481, 516-521, 550-555, 553-558, 672-677, 693-698, 694-694

Amidation sites.
amino acids 312-315, 543-546, 708-711

FIGURE 112

MATKTELSPTARESKNAQDMQVDETLIPRKGPSLCSARYGIALVLHFCNFTTIAQNVIMNI
TMVAMVNSTSPQSQLNDSSEVLPVDSFGGLSKAPKSLPAKSSILGGQFAIWEKWGPPQERS
RLCSIALSGMLLGCFTAILIGGFISETLGWPFVFYIFGGVGCVCCLLWFVVIYDDPFSYPW
ISTSEKEYIISSLKQQVGSSKQPLPIKAMLRSLPIWSICLGCFSHQWLVSTMVVYIPTYIS
SVYHVNIRDNGLLSALPFIVAWVIGMVGGYLADFLLTKKFRLITVRKIATILGSLPSSALI
VSLPYLNSGYITATALLTLSCGLSTLCQSGIYINVLDIAPRYSSFLMGASRGFSSIAPVIV
PTVSGFLLSQDPEFGWRNVFFLLFAVNLLGLLFYLIFGEADVQEWAKERKLTRL

Transmembrane domains.
amino acids 124-144, 154-174, 255-275, 287-307, 319-339, 350-370, 385-405

N-glycosylation sites.
amino acids 49-52, 60-63, 68-71, 77-80

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 415-418

N-myristoylation sites.
amino acids 106-111, 131-136, 135-140, 143-148, 160-165, 163-168, 255-260, 269-274, 272-277, 297-302, 335-340, 353-358

FIGURE 113

MQVDETLIPRKVPSLCSARYGIALVLHFCNFTTIAQNVIMNITMVAMVNSTSPQSQLNDSS
EVLPVDSFGGLSKAPKSLPTKSSILGGQFAIWEKWGPPQERSRLCSIALSGMLLGCFTAIL
IGGFISETLGWPFVFYIFGGVGCVCCLLWFVVIYDDPVSYPWISTSEKEYIISSLKQQVRS
SKQPLPIKAMLRSLPIWSICLGCFSHQWLVSTMVVYIPTYISSVYHVNIRDNGLLSALPFI
VAWVIGMVGGYLADFLLTKKFRLITVRKIATILGSLPSSALIVSLPYLNSGYITATALLTL
SCGLSTLCQSGIYINVLDIAPRYSSFLMGASRGFSSIAPVIVPTVSGFLLSQDPEFGWRNV
FFLLFAVNLLGLLFYLIFGEADVQEWAKERKLTRL

Transmembrane domains.
amino acids 22-42, 105-125, 137-157, 236-256, 268-288, 300-320, 331-351, 366-386

N-glycosylation sites.
amino acids 30-33, 41-44, 49-52, 58-61

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 396-399

N-myristoylation sites.
amino acids 87-92, 112-117, 116-121, 124-129, 141-146, 144-149, 236-241, 250-255, 253-258, 278-283, 316-321, 334-339

FIGURE 114

MAGAGPKRRALAAPAAREKEEAREKMLAAKSADGSAPAGEGEGVTLQRNITLLNGVAIIVG
TIIGSGIFVTPTGVLKEAGSPGLALVVWAACGVFSIVGALCYAELGTTISKSGGDYAYMLE
VYGSLPAFLKLWIELLIIRPSSQYIVALVFATYLLKPLFPTCPVPEEAAKLVACLCVLLLT
AVNCYSVKAATRVQDAFAAAKLLALALIILLGFVQIGKGDVSNLDPKFSFEGTKLDVGNIV
LALYSGLFAYGGWNYLNFVTEEMINPYRNLPLAIIISLPIVTLVYVLTNLAYFTTLSTEQM
LSSEAVAVDFGNYHLGVMSWIIPVFVGLSCFGSVNGSLFTSSRLFFVGSREGHLPSILSMI
HPQLLTPVPSLVFTCVMTLLYAFSKDIFSVINFFSFFNWLCVALAIIGMIWLRHRKPELER
PIKVNLALPVFFILACLFLIAVSFWKTPVECGIGFTIILSGLPVYFFGVWWKNKPKWLLQG
IFSTTVLCQKLMQVVPQET

Transmembrane domains.
amino acids 48-68, 83-103, 120-140, 137-157, 165-185, 200-220, 236-256, 274-294, 316-336, 364-384, 394-414, 430-450

N-glycosylation sites.
amino acids 49-52, 340-343

Tyrosine kinase phosphorylation site.
amino acids 112-119

N-myristoylation sites.
amino acids 34-39, 61-66, 107-112, 114-119, 125-130, 337-342, 341-346, 353-358, 459-464, 488-493

Amino acid permease.
amino acids 46-481

FIGURE 115

MAGAGPKRRALAAPAAEEKEEAREKMLAAKSADGSAPAGEGEGVTLQRNITLLNGVAIIVG
TIIGSGIFVTPTGVLKEAGSPGLALVVWAACGVFSIVGALCYAELGTTISKSGGDYAYMLE
VYGSLPAFLKLWIELLIIRPSSQYIVALVFATYLLKPLFPTCPVPEEAAKLVACLCVLLLT
AVNCYSVKAATRVQDAFAAAKLLALALIILLGFVQIGKGDVSNLDPNFSFEGTKLDVGNIV
LALYSGLFAYGGWNYLNFVTEEMINPYRNLPLAIIISLPIVTLVYVLTNLAYFTTLSTEQM
LSSEAVAVDFGNYHLGVMSWIIPVFVGLSCFGSVNGSLFTSSRLFFVGSREGHLPSILSMI
HPQLLTPVPSLVFTCVMTLLYAFSKDIFSVINFFSFFNWLCVALAIIGMIWLRHRKPELER
PIKVNLALPVFFILACLFLIAVSFWKTPVECGIGFTIILSGLPVYFFGVWWKNKPKWLLQG
IFSTTVLCQKLMQVVPQET

Transmembrane domains.
amino acids 48-68, 83-103, 120-140, 137-157, 165-185,
199-219, 236-256, 274-294, 316-336, 364-384, 394-414, 430-450

N-glycosylation sites.
amino acids 49-52, 230-233, 340-343

Tyrosine kinase phosphorylation site.
amino acids 112-119

N-myristoylation sites.
amino acids 34-39, 61-66, 107-112, 114-119, 125-130, 337-342,
341-346, 353-358, 459-464, 488-493

Amino acid permeases.
amino acids 46-481

FIGURE 116

MALPTARPLLGSCGTPALGSLLFLLFSLGWVQPSRTLAGETGQEAAPLDGVLANPPNISSL
SPRQLLGFPCAEVSGLSTERVRELAVALAQKNVKLSTEQLRCLAHRLSEPPEDLDALPLDL
LLFLNPDAFSGPQACTRFFSRITKANVDLLPRGAPERQRLLPAALACWGVRGSLLSEADVR
ALGGLACDLPGRFVAESAEVLLPRLVSCPGPLDQDQQEAARAALQGGGPPYGPPSTWSVST
MDALRGLLPVLGQPIIRSIPQGIVAAWRQRSSRDPSWRQPERTILRPRFRREVEKTACPSG
KKAREIDESLIFYKKWELEACVDAALLATQMDRVNAIPFTYEQLDVLKHKLDELYPQGYPE
SVIQHLGYLFLKMSPEDIRKWNVTSLETLKALLEVNKGHEMSPQVATLIDRFVKGRGQLDK
DTLDTLTAFYPGYLCSLSPEELSSVPPSSIWAVRPQDLDTCDPRQLDVLYPKARLAFQNMN
GSEYFVKIQSFLGGAPTEDLKALSQQNVSMDLATFMKLRTDAVLPLTVAEVQKLLGPHVEG
LKAEERHRPVRDWILRQRQDDLDTLGLGLQGGIPNGYLVLDLSVQEALSGTPCLLGPGPVL
TVLALLLASTLA

Signal sequence.
amino acids 1-33

Transmembrane domain.
amino acids 603-621

N-glycosylation sites.
amino acids 57-60, 388-391, 488-491, 515-518

Tyrosine kinase phosphorylation site.
amino acids 353-360, 471-477

N-myristoylation sites.
amino acids 171-176, 174-179, 186-191, 266-271, 501-506, 577-582, 580-585, 581-586

Amidation site.
amino acids 304-307

Leucine zipper pattern.
amino acids 101-122

FIGURE 117

MALPTARPLLGSCGTPALGSLLFLLFSLGWVQPSRTLAGETGQEAAPLDGVLANPPNISSL
SPRQLLGFPCAEVSGLSTERVRELAVALAQKNVKLSTEQLRCLAHRLSEPPEDLDALPLDL
LLFLNPDAFSGPQACTRFFSRITKANVDLLPRGAPERQRLLPAALACWGVRGSLLSEADVR
ALGGLACDLPGRFVAESAEVLLPRLVSCPGPLDQDQQEAARAALQGGGPPYGPPSTWSVST
MDALRGLLPVLGQPIIRSIPQGIVAAWRQRSSRDPSWRQPERTILRPRFRREVEKTACPSG
KKAREIDESLIFYKKWELEACVDAALLATQMDRVNAIPFTYEQLDVLKHKLDELYPQGYPE
SVIQHLGYLFLKMSPEDIRKWNVTSLETLKALLEVNKGHEMSPQVATLIDRFVKGRGQLDK
DTLDTLTAFYPGYLCSLSPEELSSVPPSSIWAVRPQDLDTCDPRQLDVLYPKARLAFQNMN
GSEYFVKIQSFLGGAPTEDLKALSQQNVSMDLATFMKLRTDAVLPLTVAEVQKLLGPHVEG
LKAEERHRPVRDWILRQRQDDLDTLGLGLQGGIPNGYLVLDLSVQEALSGTPCLLGPGPVL
TVLALLLASTLA

Signal sequence.
amino acids 1-33

Transmembrane domain.
amino acids 603-621

N-glycosylation sites.
amino acids 57-60, 388-391, 488-491, 515-518

Tyrosine kinase phosphorylation sites.
amino acids 353-360, 471-477

N-myristoylation sites.
amino acids 171-176, 174-179, 186-191, 266-271, 501-506, 577-582, 580-585, 581-586

Amidation site.
amino acids 304-307

Leucine zipper pattern.
amino acids 101-122

FIGURE 118

MATTVPDGCRNGLKSKYYRLCDKAEAWGIVLETVATAGVVTSVAFMLTLPILVCKVQDSNR
RKMLPTQFLFLLGVLGIFGLTFAFIIGLDGSTGPTRFFLFGILFSICFSCLLAHAVSLTKL
VRGRKPLSLLVILGLAVGFSLVQDVIAIEYIVLTMNRTNVNVFSELSAPRRNEDFVLLLTY
VLFLMALTFLMSSFTFCGSFTGWKRHGAHIYLTMLLSIAIWVANITLLMLPDFDRRWDDTI
LSSALAANGWVFLLAYVSPEFWLLTKQRNPMDYPVEDAFCKPQLVKKSYGVENRAYSQEEI
TQGFEETGDTLYAPYSTHFQLQNQPPQKEFSIPRAHAWPSPYKDYEVKKEGS

Transmembrane domains.
amino acids 30-50, 66-86, 98-118, 121-141, 137-157, 174-194, 212-232, 244-264

N-glycosylation site.
amino acids 158-161

N-myristoylation sites.
amino acids 8-13, 38-43, 80-85, 88-93, 102-107, 136-141, 201-206

Amidation site.
amino acids 124-127

7 transmembrane receptor.
amino acids 27-273

FIGURE 119

MIPFLPMFSLLLLLIVNPINANNHYDKILAHSRIRGRDQGPNVCALQQILGTKKKYFSTCK
NWYKKSICGQKTTVLYECCPGYMRMEGMKGCPAVLPIDHVYGTLGIVGATTTQRYSDASKL
REEIEGKGSFTYFAPSNEAWDNLDSDIRRGLESNVNVELLNALHSHMINKRMLTKDLKNGM
IIPSMYNNLGLFINHYPNGVVTVNCARIIHGNQIATNGVVHVIDRVLTQIGTSIQDFIEAE
DDLSSFRAAAITSDILEALGRDGHFTLFAPTNEAFEKLPRGVLERFMGDKVASEALMKYHI
LNTLQCSESIMGGAVFETLEGNTIEIGCDGDSITVNGIKMVNKKDIVTNNGVIHLIDQVLI
PDSAKQVIELAGKQQTTFTDLVAQLGLASALRPDGEYTLLAPVNNAFSDDTLSMVQRLLKL
ILQNHILKVKVGLNELYNGQILETIGGKQLRVFVYRTAVCIENSCMEKGSKQGRNGAIHIF
REIIKPAEKSLHEKLKQDKRFSTFLSLLEAADLKELLTQPGDWTLFVPTNDAFKGMTSEEK
EILIRDKNALQNIILYHLTPGVFIGKGFEPGVTNILKTTQGSKIFLKEVNDTLLVNELKSK
ESDIMTTNGVIHVVDKLLYPADTPVGNDQLLEILNKLIKYIQIKFVRGSTFKEIPVTVYTT
KIITKVVEPKIKVIEGSLQPIIKTEGPTLTKVKIEGEPEFRLIKEGETITEVIHGEPIIKK
YTKIIDGVPVEITEKETREERIITGPEIKYTRISTGGGETEETLKKLLQEEVTKVTKFIEG
GDGHLFEDEEIKRLLQGDTPVRKLQANKKVQGSRRRLREGRSQ

Signal sequence.
amino acids 1-21

N-glycosylation site.
amino acids 599-602

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 507-510, 731-734

N-myristoylation sites.
amino acids 70-75, 106-111, 109-114, 152-157, 193-198, 214-219, 392-397, 476-481, 570-575, 768-773

Fasciclin domain.
amino acids 94-232, 496-630

FIGURE 120

MIPFLPMFSLLLLLIVNPINANNHYDKILAHSRIRGRDQGPNVCALQQILGTKKKYFSTCK
NWYKKSICGQKTTVLYECCPGYMRMEGMKGCPAVLPIDHVYGTLGIVGATTTQRYSDASKL
REEIEGKGSFTYFAPSNEAWDNLDSDIRRGLESNVNVELLNALHSHMINKRMLTKDLKNGM
IIPSMYNNLGLFINHYPNGVVTVNCARIIHGNQIATNGVVHVIDRVLTQIGTSIQDFIEAE
DDLSSFRAAAITSDILEALGRDGHFTLFAPTNEAFEKLPRGVLERIMGDKVASEALMKYHI
LNTLQCSESIMGGAVFETLEGNTIEIGCDGDSITVNGIKMVNKKDIVTNNGVIHLIDQVLI
PDSAKQVIELAGKQQTTFTDLVAQLGLASALRPDGEYTLLAPVNNAFSDDTLSMDQRLLKL
ILQNHILKVKVGLNELYNGQILETIGGKQLRVFVYRTAVCIENSCMEKGSKQGRNGAIHIF
REIIKPAEKSLHEKLKQDKRFSTFLSLLEAADLKELLTQPGDWTLFVPTNDAFKGMTSEEK
EILIRDKNALQNIILYHLTPGVFIGKGFEPGVTNILKTTQGSKIFLKEVNDTLLVNELKSK
ESDIMTTNGVIHVVDKLLYPADTPVGNDQLLEILNKLIKYIQIKFVRGSTFKEIPVTVYRP
TLTKVKIEGEPEFRLIKEGETITEVIHGEPIIKKYTKIIDGVPVEITEKETREERIITGPE
IKYTRISTGGGETEETLKKLLQEEVTKVTKFIEGGDGHLFEDEEIKRLLQGDTPVRKLQAN
KKVQGSRRRLREGRSQ

Signal sequence.
amino acids 1-21

N-glycosylation site.
amino acids 599-602

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 507-510, 704-707

N-myristoylation sites.
amino acids 70-75, 106-111, 109-114, 152-157, 193-198, 214-219, 392-397, 476-481, 570-575, 741-746

Fasciclin domains.
amino acids 94-232, 234-367, 370-494, 496-630

FIGURE 121

MIPFLPMFSLLLLLIVNPINANNHYDKILAHSRIRGRDQGPNVCALQQILGTKKKYFSTCK
NWYKKSICGQKTTVLYECCPGYMRMEGMKGCPAVLPIDHVYGTLGIVGATTTQRYSDASKL
REEIEGKGSFTYFAPSNEAWDNLDSDIRRGLESNVNVELLNALHSHMINKRMLTKDLKNGM
IIPSMYNNLGLFINHYPNGVVTVNCARIIHGNQIATNGVVHVIDRVLTQIGTSIQDFIEAE
DDLSSFRAAAITSDILEALGRDGHFTLFAPTNEAFEKLPRGVLERIMGDKVASEALMKYHI
LNTLQCSESIMGGAVFETLEGNTIEIGCDGDSITVNGIKMVNKKDIVTNNGVIHLIDQVLI
PDSAKQVIELAGKQQTTFTDLVAQLGLASALRPDGEYTLLAPVNNAFSDDTLSMDQRLLKL
ILQNHILKVKVGLNELYNGQILETIGGKQLRVFVYRTAVCIENSCMEKGSKQGRNGAIHIF
REIIKPAEKSLHEKLKQDKRFSTFLSLLEAADLKELLTQPGDWTLFVPTNDAFKGMTSEEK
EILIRDKNALQNIILYHLTPGVFIGKGFEPGVTNILKTTQGSKIFLKEVNDTLLVNELKSK
ESDIMTTNGVIHVVDKLLYPADTPVGNDQLLEILNKLIKYIQIKFVRGSTFKEIPVTVYKP
IIKKYTKIIDGVPVEITEKETREERIITGPEIKYTRISTGGGETEETLKKLLQEDTPVRKL
QANKKVQGSRRRLREGRSQ

Signal sequence.
amino acids 1-21

N-glycosylation site.
amino acids 599-602

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 507-510, 674-677

N-myristoylation sites.
amino acids 70-75, 106-111, 109-114, 152-157, 193-198, 214-219, 392-397, 476-481, 570-575, 711-716

Fasciclin domains.
amino acids 94-232, 234-367, 370-494, 496-630

FIGURE 122

MIPFLPMFSLLLLLIVNPINANNHYDKILAHSRIRGRDQGPNVCALQQILGTKKKYFSTCK
NWYKKSICGQKTTVLYECCPGYMRMEGMKGCPAVLPIDHVYGTLGIVGATTTQRYSDASKL
REEIEGKGSFTYFAPSNEAWDNLDSDIRRGLESNVNVELLNALHSHMINKRMLTKDLKNGM
IIPSMYNNLGLFINHYPNGVVTVNCARIIHGNQIATNGVVHVIDRVLTQIGTSIQDFIEAE
DDLSSFRAAAITSDILEALGRDGHFTLFAPTNEAFEKLPRGVLERIMGDKVASEALMKYHI
LNTLQCSESIMGGAVFETLEGNTIEIGCDGDSITVNGIKMVNKKDIVTNNGVIHLIDQVLI
PDSAKQVIELAGKQQTTFTDLVAQLGLASALRPDGEYTLLAPVNNAFSDDTLSMDQRLLKL
ILQNHILKVKVGLNELYNGQILETIGGKQLRVFVYRTAVCIENSCMEKGSKQGRNGAIHIF
REIIKPAEKSLHEKLKQDKRFSTFLSLLEAADLKELLTQPGDWTLFVPTNDAFKGMTSEEK
EILIRDKNALQNIILYHLTPGVFIGKGFEPGVTNILKTTQGSKIFLKEVNDTLLVNELKSK
ESDIMTTNGVIHVVDKLLYPADTPVGNDQLLEILNKLIKYIQIKFVRGSTFKEIPVTVYRP
TLTKVKIEGEPEFRLIKEGETITEVIHGEPIIKKYTKIIDGVPVEITEKETREERIITGPE
IKYTRISTGGGETEETLKKLLQEDTPVRKLQANKKVQGSRRRLREGRSQ

Signal sequence.
amino acids 1-21

N-glycosylation site.
amino acids 599-602

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 507-510, 704-707

N-myristoylation sites.
amino acids 70-75, 106-111, 109-114, 152-157, 193-198, 214-219, 392-397, 476-481, 570-575, 741-746

Fasciolin domains.
amino acids 94-232, 234-367, 370-494, 496-630

FIGURE 123

MIPFLPMFSLLLLLIVNPINANNHYDKILAHSRIRGRDQGPNVCALQQILGTKKKYFSTCK
NWYKKSICGQKTTVLYECCPGYMRMEGMKGCPAVLPIDHVYGTLGIVGATTTQRYSDASKL
REEIEGKGSFTYFAPSNEAWDNLDSDIRRGLESNVNVELLNALHSHMINKRMLTKDLKNGM
IIPSMYNNLGLFINHYPNGVVTVNCARIIHGNQIATNGVVHVIDRVLTQIGTSIQDFIEAE
DDLSSFRAAAITSDILEALGRDGHFTLFAPTNEAFEKLPRGVLERIMGDKVASEALMKYHI
LNTLQCSESIMGGAVFETLEGNTIEIGCDGDSITVNGIKMVNKKDIVTNNGVIHLIDQVLI
PDSAKQVIELAGKQQTTFTDLVAQLGLASALRPDGEYTLLAPVNNAFSDDTLSMDQRLLKL
ILQNHILKVKVGLNELYNGQILETIGGKQLRVFVYRTAVCIENSCMEKGSKQGRNGAIHIF
REIIKPAEKSLHEKLKQDKRFSTFLSLLEAADLKELLTQPGDWTLFVPTNDAFKGMTSEEK
EILIRDKNALQNIILYHLTPGVFIGKGFEPGVTNILKTTQGSKIFLKEVNDTLLVNELKSK
ESDIMTTNGVIHVVDKLLYPADTPVGNDQLLEILNKLIKYIQIKFVRGSTFKEIPVTVYKP
IIKKYTKIIDGVPVEITEKETREERIITGPEIKYTRISTGGGETEETLKKLLQEEVTKVTK
FIEGGDGHLFEDEEIKRLLQGDTPVRKLQANKKVQGSRRRLREGRSQ

Signal sequence.
amino acids 1-21

N-glycosylation site.
amino acids 599-602

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 507-510, 674-677

N-myristoylation sites.
amino acids 70-75, 106-111, 109-114, 152-157, 193-198, 214-219, 392-397, 476-481, 570-575, 711-716

Fasciclin domains.
amino acids 94-232, 234-367, 370-494, 496-630

FIGURE 124

MVNYAWAGRSQRKLWWRSVAVLTCKSVVRPGYRGGLQARRSTLLKTCARARATAPGAMKMV
APWTRFYSNSCCLCCHVRTGTILLGVWYLIINAVVLLILLSALADPDQYNFSSSELGGDFE
FMDDANMCIAIAISLLMILICAMATYGAYKQRAAWIIPFFCYQIFDFALNMLVAITVLIYP
NSIQEYIRQLPPNFPYRDDVMSVNPTCLVLIILLFISIILTFKGYLISCVWNCYRYINGRN
SSDVLVYVTSNDTTVLLPPYDDATVNGAAKEPPPPYVSA

Transmembrane domains.
amino acids 83-103, 124-144, 162-182, 205-225

N-glycosylation sites.
amino acids 111-114, 244-227, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 39-42

Tyrosine kinase phosphorylation site.
amino acids 243-251

N-myristoylation site.
amino acids 34-39

FIGURE 125

MPGQGGLQARRSTLLKTCARARATAPGAMKMVAPWTRFYSNSCCLCCHVRTGTILLGVWYL
IINAVVLLILLSALADPDQYNFSSSELGGDFEFMDDANMCIAIAISLLMILICAMATYGAY
KQRAAWIIPFFCYQIFDFALNMLVAITVLIYPNSIQEYIRQLPPNFPYRDDVMSVNPTCLV
LIILLFISIILTFKGYLISCVWNCYRYINGRNSSDVLVYVTSNDTTVLLPPYDDATVNGAA
KEPPPPYVSA

Transmembrane domains.
amino acids 54-74, 95-115, 133-153, 176-196

N-glycosylation sites.
amino acids 82-85, 215-218, 226-229

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 10-13

Tyrosine kinase phosphorylation site.
amino acids 214-222

N-myristoylation site.
amino acids 5-10

FIGURE 126

MGRAGGGGPGRGPPPLLLFLGAALVLASGAVPAREAGSAVEAEELVKGSPAWEPPANDTRE
EAGPPAAGEDEASWTAPGGELAGPEEVLQESAAVTGTAWLEADSPGLGGVTARAGSGDAQA
LPATLQAPHEVLGQSIMPPAIPEATEASGPPSPTPGDKLSPASELPKESPLEVWLNLGGST
PDPQVPELTYPFQGTLEPQPASDIIDIDYFEGLDGEGRGADLGSFPGSPGTSENHPDTEGE
TPSWSLLDLYDDFTPFDESDFYPTTSFYDDLDEEEEEEEDDKDAVGGGDLEDENELLVPTG
KPGLGPGTGQPTSRWHAVPPQHTLGSVPGSSIALRPRPGEPGRDLASSENGTECRSGFVRH
NGSCRSVCDLFPSYCHNGGQCYLVENIGAFCRCNTQDYIWHKGMRCESIITDFQVMCVAVG
SAALVLLLLFMMTVFFAKKLYLLKTENTKLRRTNKFRTPSELHNDNFSLSTIAEGSHPNDD
PSAPHKIQEVLKSCLKEEESFNIQNSMSPKLEGGKGDQADLDVNCLQNNLT

Signal sequence.
amino acids 1-30

Transmembrane domain.
amino acids 423-443

N-glycosylation sites.
amino acids 57-60, 355-358, 367-370, 473-476

N-myristoylation sites.
amino acids 6-11, 29-34, 79-84, 109-114, 116-121, 222-227, 226-231, 230-235, 233-238, 308-313, 314-319, 330-335, 334-339, 368-373, 482-487

FIGURE 127

MLPVYQEVKPNPLQDANICSRVFFWWLNPLFKIGHKRRLEEDDMYSVLPEDRSQHLGEELQ
GFWDKEVLRAENDAQKPSLTRAIIKCYWKSYLVLGIFTLIEESAKVIQPIFLGKIINYFEN
YDPMDSVALNTAYAYATVLTFCTLILAILHHLYFYHVQCAGMRLRVAMCHMIYRKALRLSN
MAMGKITTGQIVNLLSNDVNKFDQVTVFLHFLWAGPLQAIAVTALLWMEIGISCLAGMAVL
IILLPLQSCFGKLFSSLRSKTATFTDARIRTMNEVITGIRIIKMYAWEKSFSNLITNLRKK
EISKILRSSCLRGMNLASFFSASKIIVFVTFTTYVLLGSVITASRVFVAVTLYGAVRLTVT
LFFPSAIERVSEAIVSIRRIQTFLLLDEISQRNRQLPSDGKKMVHVQDFTAFWDKASETPT
LQGLSFTVRPGELLAVVGPVGAGKSSLLSAVLGELAPSHGLVSVHGRIAYVSQQPWVFSGT
LRSNILFGKKYEKERYEKVIKACALKKDLQLLEDGDLTVIGDRGTTLSGGQKARVNLARAV
YQDADIYLLDDPLSAVDAEVSRHLFELCICQILHEKITILVTHQLQYLKAASQILILKDGK
MVQKGTYTEFLKSGIDFGSLLKKDNEESEQPPVPGTPTLRNRTFSESSVWSQQSSRPSLKD
GALESQDTENVPVTLSEENRSEGKVGFQAYKNYFRAGAHWIVFIFLILLNTAAQVAYVLQD
WWLSYWANKQSMLNVTVNGGGNVTEKLDLNWYLGIYSGLTVATVLFGIARSLLVFYVLVNS
SQTLHNKMFESILKAPVLFFDRNPIGRILNRFSKDIGHLDDLLPLTFLDFIQTLLQVVGVV
SVAVAVIPWIAIPLVPLGIIFIFLRRYFLETSRDVKRLESTTRSPVFSHLSSSLQGLWTIR
AYKAEERCQELFDAHQDLHSEAWFLFLTTSRWFAVRLDAICAMFVIIVAFGSLILAKTLDA
GQVGLALSYALTLMGMFQWCVRQSAEVENMMISVERVIEYTDLEKEAPWEYQKRPPPAWPH
EGVIIFDNVNFMYSPGGPLVLKHLTALIKSQEKVGIVGRTGAGKSSLISALFRLSEPEGKI
WIDKILTTEIGLHDLRKKMSIIPQEPVLFTGTMRKNLDPFKEHTDEELWNALQEVQLKETI
EDLPGKMDTELAESGSNFSVGQRQLVCLARAILRKNQILIIDEATANVDPRTDELIQKKIR
EKFAHCTVLTIAHRLNTIIDSDKIMVLDSGRLKEYDEPYVLLQNKESLFYKMVQQLGKAEA
AALTETAKQVYFKRNYPHIGHTDHMVTNTSNGQPSTLTIFETAL

Transmembrane domains:

amino acids 134-154, 233-253, 346-366, 708-728, 767-787, 834-854, 857-877, 948-968, 967-987

N-glycosylation sites.

amino acids 651-654, 690-693, 746-749, 754-757, 792-795, 1176-1179, 1309-1312

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 1115-1118

Tyrosine kinase phosphorylation sites.

amino acids 37-45, 497-504, 1253-1259

N-myristoylation sites.

amino acids 192-197, 318-323, 343-348, 430-435, 448-453,

766-771, 770-775, 779-784, 980-985, 1078-1083, 1174-1179, 1313-1318

Amidation sites.

amino acids 405-410, 495-500

ATP/GTP-binding site motif A (P-loop).

amino acids 445-452, 1075-1082

ABC transporter.

amino acids 438-609, 1068-1250

ABC transporter transmembrane region.

amino acids 92-365, 711-997

FIGURE 128

MSALRRKFGDDYQVVTTSSSGSGLQPQGPGQDPQQQLVPKKKRQRFVDKNGRCNVQHGNLG
SETSRYLSDLFTTLVDLKWRWNLFIFILTYTVAWLFMASMWWVIAYTRGDLNKAHVGNYTP
CVANVYNFPSAFLFFIETEATIGYGYRYITDKCPEGIILFLFQSILGSIVDAFLIGCMFIK
MSQPKKRAETLMFSEHAVISMRDGKLTLMFRVGNLRNSHMVSAQIRCKLLKSRQTPEGEFL
PLDQLELDVGFSTGADQLFLVSPLTICHVIDAKSPFYDLSQRSMQTEQFEIVVILEGIVET
TGMTCQARTSYTEDEVLWGHRFFPVISLEEGFFKVDYSQFHATFEVPTPPYSVKEQEEMLL
MSSPLIAPAITNSKERHNSVECLDGLDDITTKLPSKLQKITGREDFPKKLLRMSSTTSEKA
YSLGDLPMKLQRISSVPGNSEEKLVSKTTKMLSDPMSQSVADLPPKLQKMAGGAARMEGNL
PAKLRKMNSDRFT

Transmembrane domains.
amino acids 83-103, 120-140, 158-178

Glycosaminoglycan attachment site.
amino acids 20-23

N-myristoylation sites.
amino acids 58-63, 61-66, 216-221, 301-306, 486-491

Cell attachment sequence.
amino acids 109-114

Inward rectifier potassium channel.
amino acids 47-393

FIGURE 129

MPLFKNTSVSSLYSGCRLTLLRPEKDGAATRVDAVCTHRPDPKSPGLDRERLYWKLSQLTH
GITELGPYTLDRHSLYVNGFTHQSSMTTTRTPDTSTMHLATSRTPASLSGPTTASPLLVLF
TINFTITNLRYEENMHHPGSRKFNTTERVLQGLLRPVFKNTSVGPLYSGCRLTLLRPKKDG
AATKVDAICTYRPDPKSPGLDREQLYWELSQLTHSITELGPYTLDRDSLYVNGFTQRSSVP
TTSIPGTPTVDLGTSGTPVSKPGPSAASPLLVLFTLNFTITNLRYEENMQHPGSRKFNTTE
RVLQGLLRSLFKSTSVGPLYSGCRLTLLRPEKDGTATGVDAICTHHPDPKSPRLDREQLYW
ELSQLTHNITELGHYALDNDSLFVNGFTHRSSVSTTSTPGTPTVYLGASKTPASIFGPSAA
SHLLILFTLNFTITNLRYEENMWPGSRKFNTTERVLQGLLRPLFKNTSVGPLYSGSRLTLL
RPEKDGEATGVDAICTHRPDPTGPGLDREQLYLELSQLTHSITELGPYTLDRDSLYVNGFT
HRSSVPTTSTGVVSEEPFTLNFTINNLRYMADMGQPGSLKFNITDNVMKHLLSPLFQRSSL
GARYTGCRVIALRSVKNGAETRVDLLCTYLQPLSGPGLPIKQVFHELSQQTHGITRLGPYS
LDKDSLYLNGYNEPGLDEPPTTPKPATTFLPPLSEATTAMGYHLKTLTLNFTISNLQYSPD
MGKGSATFNSTEGVLQHLLRPLFQKSSMGPFYLGCQLISLRPEKDGAATGVDTTCTYHPDP
VGPGLDIQQLYWELSQLTHGVTQLGFYVLDRDSLFINGYAPQNLSIRGEYQINFHIVNWNL
SNPDPTSSEYITLLRDIQDKVTTLYKGSQLHDTFRFCLVTNLTMDSVLVTVKALFSSNLDP
SLVEQVFLDKTLNASFHWLGSTYQLVDIHVTEMESSVYQPTSSSSTQHFYPNFTITNLPYS
QDKAQPGTTNYQRNKRNIEDALNQLFRNSSIKSYFSDCQVSTFRSVPNRHHTGVDSLCNFS
PLARRVDRVAIYEEFLRMTRNGTQLQNFTLDRSSVLVDGYSPNRNEPLTGNSDLPFWAVIF
IGLAGLLGLITCLICGVLVTTRRRKKEGEYNVQQQCPGYYQSHLDLEDLQ

Transmembrane domains.
amino acids 107-127, 418-438, 1093-1113

N-glycosylation sites.
amino acids 6-9, 125-128, 146-149, 162-165, 281-284, 302-305, 374-377, 385-388, 437-440, 457-460, 473-476, 570-573, 591-594, 721-724, 741-744, 836-839, 853-856, 895-898, 928-931, 967-970, 1004-1007, 1058-1061, 1064-1067

Glycosaminoglycan attachment site.
amino acids 644-647

Tyrosine kinase phosphorylation sites.
amino acids 358-365, 1121-1128

N-myristoylation sites.
amino acids 15-20, 171-176, 260-265, 310-315, 327-332, 339-344, 482-487, 583-588, 611-616, 778-783, 782-787, 1106-1111, 1114-1119

SEA domain.
amino acids 112-245, 270-391, 661-684

FIGURE 130

MGNWVVNHWFSVLFLVVWLGLNVFLFVDAFLKYEKADKYYYTRKILGSTLACARASALCLN
FNSTLILLPVCRNLLSFLRGTCSFCSRTLRKQLDHNLTFHKLVAYMICLHTAIHIIAHLFN
FDCYSRSRQATDGSLASILSSLSHDEKKGGSWLNPIQSRNTTVEYVTFTSVAGLTGVIMTI
ALILMVTSATEFIRRSYFEVFWYTHHLFIFYILGLGIHGIGGIVRGQTEESMNESHPRKCA
ESFEMWDDRDSHCRRPKFEGHPPESWKWILAPVILYICERILRFYRSQQKVVITKVVMHPS
KVLELQMNKRGFSMEVGQYIFVNCPSISLLEWHPFTLTSAPEEDFFSIHIRAAGDWTENLI
RAFEQQYSPIPRIEVDGPFGTASEDVFQYEVAVLVGAGIGVTPFASILKSIWYKFQCADHN
LKTKKIYFYWICRETGAFSWFNNLLTSLEQEMEELGKVGFLNYRLFLTGWDSNIVGHAALN
FDKATDIVTGLKQKTSFGRPMWDNEFSTIATSHPKSVVGVFLCGPRTLAKSLRKCCHRYSS
LDPRKVQFYFNKENF

Signal sequence.
amino acids 1-29

Transmembrane domain.
amino acids 44-64, 60-80, 99-119, 168-188, 204-224

N-glycosylation sites.
amino acids 63-66, 97-100, 162-165, 236-239

Tyrosine kinase phosphorylation sites.
amino acids 198-206, 367-373

N-myristoylation sites.
amino acids 47-52, 135-140, 178-183, 225-230, 402-407, 527-532

Ferric reductase like.
amino acids 46-563

FIGURE 131

MGNWVVNHWFSVLFLVVWLGLNVFLFVDAFLKYEKADKYYYTRKILGSTLACARASALCLN
FNSTLILLPVCRNLLSFLRGTCSFCSRTLRKQLDHNLTFHKLVAYMICLHTAIHIIAHLFN
FDCYSRSRQATDGSLASILSSLSHDEKKGGSWLNPIQSRNTTVEYVTFTSIAGLTGVIMTI
ALILMVTSATEFIRRSYFEVFWYTHHLFIFYILGLGIHGIGGIVRGQTEESMNESHPRKCA
ESFEMWDDRDSHCRRPKFEGHPPESWKWILAPVILYICERILRFYRSQQKVVITKVVMHPS
KVLELQMNKRGFSMEVGQYIFVNCPSISLLEWHPFTLTSAPEEDFFSIHIRAAGDWTENLI
RAFEQQYSPIPRIEVDGPFGTASEDVFQYEVAVLVGAGIGVTPFASILKSIWYKFQCADHN
LKTKKIYFYWICRETGAFSWFNNLLTSLEQEMEELGKVGFLNYRLFLTGWDSNIVGHAALN
FDKATDIVTGLKQKTSFGRPMWDNEFSTIATSHPKSVVGVFLCGPRTLAKSLRKCCHRYSS
LDPRKVQFYFNKENF

Transmembrane domains.
amino acids 6-26, 44-64, 60-80, 99-119, 168-188, 204-224

N-glycosylation sites.
amino acids 63-66, 97-100, 162-165, 236-239

Tyrosine kinase phosphorylation sites.
amino acids 198-206, 367-373

N-myristoylation sites.
amino acids 47-52, 135-140, 178-183, 225-230, 402-407, 527-532

Ferric reductase like transmembrane.
amino acids 46-563

FIGURE 132

MGNWVVNHWPSVLFLVVWLGLNVFLFVDAFLKYEKADKYYYTRKILGSWKWILAPVILYIC
ERILRFYRSQQKVVITKVVMHPSKVLELQMNKRGFSMEVGQYIFVNCPSISLLEWHPFTLT
SAPEEDFFSIHIRAAGDWTENLIRAFEQQYSPIPRIEVDGPFGTASEDVFQYEVAVLVGAG
IGVTPFASILKSIWYKFQCADHNLKTKKVGHAALNFDKATDIVTGLKQKTSFGRPMWDNEF
STIATSHPKSVVGVFLCGPRTLAKSLRKCCHRYSSLDPRKVQFYFNKENFO

Transmembrane domains.
amino acids 6-26, 44-64

Tyrosine kinase phosphorylation site.
amino acids 146-152

N-myristoylation sites.
amino acids 181-186, 257-262

Ferric reductase like transmembrane.
amino acids 1-293

FIGURE 133

MGNWVVNHWFSVLFLVVWLGLNVFLFVDAFLKYEKADKYYYTRKILGSTLACARASALCLN
FNSTLILLPVCRNLLSFLRGTCSFCSRTLRKQLDHNLTFHKLVAYMICLHTAIHIIAHLFN
FDCYSRSRQATDGSLASILSSLSHDEKKGGSWLNPIQSRNTTVEYVTFTSIAGLTGVIMTI
ALILMVTSATEFIRRSYFEVFWYTHHLFIFYILGLGIHGIGGIVRGQTEESMNESHPRKCA
ESFEMWDDRDSHCRRPKFEGHPPESWKWILAPVILYICERILRFYRSQQKVVITKVVMHPS
KVLELQMNKRGFSMEVGQYIFVNCPSISLLEWHPFTLTSAPEEDFFSIHIRAAGDWTENLI
RAFEQQYSPIPRIEVDGPFGTASEDVFQYEVAVLVGAGIGVTPFASILKSIWYKFQCADHN
LKTKKVGHAALNFDKATDIVTGLKQKTSFGRPMWDNEFSTIATSHPKSVVGVFLCGPRTLA
KSLRKCCHRYSSLDPRKVQFYFNKENF

Transmembrane domains.
amino acids 6-26, 44-64, 60-80, 99-119, 168-188, 204-224

N-glycosylation sites.
amino acids 63-66, 97-100, 162-165, 236-239

Tyrosine kinase phosphorylation sites.
amino acids 198-206, 367-373

N-myristoylation sites.
amino acids 47-52, 135-140, 178-183, 225-230, 402-407, 478-483

Ferric reductase like transmembrane.
amino acids 46-514

FIGURE 134

MTSGSVFFYILIFGKYFSHGGGQDVKCSLGYFPCGNITKCLPQLLHCNGVDDCGNQADEDN
CGDNNGWSMQFDKYFASYYKMTSQYPFEAETPECLVGSVPVQCLCQGLELDCDETNLRAVP
SVSSNVTAMSLQWNLIRKLPPDCFKNYHDLQKLYLQNNKITSISIYAFRGLNSLTKLYLSH
NRITFLKPGVFEDLHRLEWLIIEDNHLSRISPPTFYGLNSLILLVLMNNVLTRLPDKPLCQ
HMPRLHWLDLEGNHIHNLRNLTFISCSNLTVLVMRKNKINHLNENTFAPLQKLDELDLGSN
KIENLPPLIFKDLKELSQLNLSYNPIQKIQANQFDYLVKLKSLSLEGIEISNIQQRMFRPL
MNLSHIYFKKFQYCGYAPHVRSCKPNTDGISSLENLLASIIQRVFVWVVSAVTCFGNIFVI
CMRPYIRSENKLYAMSIISLCCADCLMGIYLFVIGGFDLKFRGEYNKHAQLWMESTHCQLV
GSLAILSTEVSVLLLTFLTLEKYICIVYPFRCVRPGKCRTITVLILIWITGFIVAFIPLSN
KEFFKNYYGTNGVCFPLHSEDTESIGAQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSA
ITATEIRNQVKKEMILAKRFFFIVFTDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILP
INSALNPILYTLTTRPFKEMIHRFWYNYRQRKSMDSKGQKTYAPSFIWVEMWPLQEMPPEL
MKPDLFTYPCEMSLISQSTRLNSYS

Transmembrane domains.
amino acids 212-232, 404-424, 443-463, 489-509, 526-546, 577-597, 629-649, 665-685

N-glycosylation sites.
amino acids 36-39, 127-130, 264-267, 272-275, 325-328, 368-371

N-myristoylation sites.
amino acids 49-54, 352-357, 586-591, 659-664

Leucine zipper pattern.
amino acids 487-508

LDL-receptor class A.
amino acids 40-59

G-protein coupled receptor.
amino acids 159-189, 413-432, 422-452, 436-466, 478-518, 528-547, 625-644, 673-689

Leucine rich repeats.
amino acids 127-150, 151-174, 175-198, 199-222, 223-246, 248-271, 272-295, 296-319, 320-343, 344-367

7 transmembrane receptor (rhodopsin family proteins).
amino acids 485-681

Low-density lipoprotein receptor domain.
amino acids 25-64

FIGURE 135

MTSGSVFFYILIFGKYFSHGGGQDVKCSLGYFPCGNITKCLPQLLHCNGVDDCGNQADEDN
CGDNNGWSLQFDKYFASYYKMTSQYPFEAETPECLVGSVPVQCLCQGLELDCDETNLRAVP
SVSSNVTAMSLQWNLIRKLPPDCFKNYHDLQKLYLQNNKITSISIYAFRGLNSLTKLYLSH
NRITFLKPGVFEDLHRLEWLIIEDNHLSRISPPTFYGLNSLILLVLMNNVLTRLPDKPLCQ
HMPRLHWLDLEGNHIHNLRNLTFISCSNLTVLVMRKNKINHLNENTFAPLQKLDELDLGSN
KIENLPPLIFKDLKELSQLNLSYNPIQKIQANQFDYLVKLKSLSLEGIEISNIQQRMFRPL
MNLSHIYFKKFQYCGYAPHVRSCKPNTDGISSLENLLASIIQRVFVWVVSAVTCFGNIFVI
CMRPYIRSENKLYAMSIISLCCADCLMGIYLFVIGGFDLKFRGEYNKHAQLWMESTHCQLV
GSLAILSTEVSVLLLTFLTLEKYICIVYPFRCVRPGKCRTITVLILIWITGFIVAFIPLSN
KEFFKNYYGTNGVCFPLHSEDTESIGAQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSA
ITATEIRNQVKKEMILAKRFFFIVFTDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILP
INSALNPILYTLTTRPFKEMIHRFWYNYRQRKSMDSKGQKTYAPSFIWVEMWPLQEMPPEL
MKPDLFTYPCEMSLISQSTRLNSYS

Transmembrane domains.
amino acids 212-232, 404-424, 443-463, 489-509, 526-546, 577-597, 629-649, 665-685

N-glycosylation sites.
amino acids 36-39, 127-130, 264-267, 272-275, 325-328, 368-371

N-myristoylation sites.
amino acids 49-54, 352-357, 586-591, 659-664

Leucine zipper pattern.
amino acids 487-508

Leucine rich repeats.
amino acids 127-150, 151-174, 175-198, 199-222, 223-246, 248-271, 272-295, 296-319, 320-343, 344-367

7 transmembrane receptor (rhodopsin family).
amino acids 59-259

Low-density lipoprotein receptor domain.
amino acids 25-64

FIGURE 136

MGRLASRPLLLALLSLALCRGRVVRVPTATLVRVVGTELVIPCNVSDYDGPSEQNFDWSFS
SLGSSFVELASTWEVGFPAQLYQERLQRGEILLRRTANDAVELHIKNVQPSDQGHYKCSTP
STDATVQGNYEDTVQVKVLADSLHVGPSARPPPSLSLREGEPFELRCTAASASPLHTHLAL
LWEVHRGPARRSVLALTHEGRFHPGLGYEQRYHSGDVRLDTVGSDAYRLSVSRALSADQGS
YRCIVSEWIAEQGNWQEIQEKAVEVATVVIQPTVLRAAVPKNVSVAEGKELDLTCNITTDR
ADDVRPEVTWSFSRMPDSTLPGSRVLARLDRDSLVHSSPHVALSHVDARSYHLLVRDVSKE
NSGYYYCHVSLWAPGHNRSWHKVAEAVSSPAGVGVTWLEPDYQVYLNASKVPGFADDPTEL
ACRVVDTKSGEANVRFTVSWYYRMNRRSDNVVTSELLAVMDGDWTLKYGERSKQRAQDGDF
IFSKEHTDTFNFRIQRTTEEDRGNYYCVVSAWTKQRNNSWVKSKDVFSKPVNIFWALEDSV
LVVKARQPKPFFAAGNTFEMTCKVSSKNIKSPRYSVLIMAEKPVGDLSSPNETKYIISLDQ
DSVVKLENWTDASRVDGVVLEKVQEDEFRYRMYQTQVSDAGLYRCMVTAWSPVRGSLWREA
ATSLSNPIEIDFQTSGPIFNASVHSDTPSVIRGDLIKLFCIITVEGAALDPDDMAFDVSWF
AVHSFGLDKAPVLLSSLDRKGIVTTSRRDWKSDLSLERVSVLEFLLQVHGSEDQDFGNYYC
SVTPWVKSPTGSWQKEAEIHSKPVFITVKMDVLNAFKYPLLIGIGLSTVIGLLSCLIGYCS
SHWCCKKEVQETRRERRRLMSMEMD

Signal sequence.
amino acids 1-21

Transmembrane domain.
amino acids 833-853

N-glycosylation sites.
amino acids 44-47, 286-289, 300-303, 383-386, 413-416, 525-528, 600-603, 618-621, 691-694

Tyrosine kinase phosphorylation site.
amino acids 632-640

N-myristoylation sites.
amino acids 243-248, 398-403, 511-516, 651-656, 753-758, 789-794, 836-841, 844-849

Cell attachment sequence.

Immunoglobulin domain.

amino acids 36-121, 162-249, 292-375, 422-517, 564-657, 704-795

FIGURE 137

MSRSRHLGKIRKRLEDVKSQWVRPARADFSDNESARLATDALLDGGSEAYWRVLSQEGEVD
FLSSVEAQYIQAQAREPPCPPDTLGGAEAGPKGLDSSSLQSGTYFPVASEGSEPALLHSWA
SAEKPYLKEKSSATVYFQTVKHNNIRDLVRRCITRTSQVLVILMDVFTDVEIFCDILEAAN
KRGVFVCVLLDQGGVKLFQEMCDKVQISDSHLKNISIRSVEGEIYCAKSGRKFAGQIREKF
IISDWRFVLSGSYSFTWLCGHVHRNILSKFTGQAVELFDEEFRHLYASSKPVMGLKSPRLV
APVPPGAAPANGRLSSSSGSASDRTSSNPFSGRSAGSHPGTRSVSASSGPCSPAAPHPPPP
PRFQPHQGPWGAPSPQAHLSPRPHDGPPAAVYSNLGAYRPTRLQLEQLGLVPRLTPTWRPF
LQASPHF

Transmembrane domain.
amino acids 154-175

N-glycosylation sites.
amino acids 32-35, 217-220

Tyrosine kinase phosphorylation site.
amino acids 221-228

N-myristoylation sites.
amino acids 45-50, 86-91, 94-99, 112-117, 186-191, 311-316, 341-346

Amidation site.
amino acids 232-235

FIGURE 138

MALPQPPPPAGARDPVTPEHWASGPPSQAQPLLRQEAKEEEEGEETGVQGAWGTGTAEQRR
RGWGEAAESAAAEEGQAEVGGAAAAGSGSPAGGAGGGLGSWRPLLAWLQRRQPQCCPCAAP
LSRSAAHCCHGGTKMAALAYNLGKREINHYFSVRSAKVLALVAVLLLAACHLASRRYRGND
SCEYLLSSGRFLGEKVWQPHSCMMHKYKISEAKNCLVDKHIAFIGDSRIRQLFYSFVKIIN
PQFKEEGNKHENIPFEDKTASVKVDFLWHPEVNGSMKQCIKVWTEDSIAKPHVIVAGAATW
SIKIHNGSSEALSQYKMNITSIAPLLEKLAKTSDVYWVLQDPVYEDLLSENRKMITNEKID
AYNEAAVSILNSSTRNSKSNVKMFSVSKLIAQETIMESLDGLHLPESSRETTAMILMNVYC
NKILKPVDGSCCQPRPPVTLIQKLAACFFTLSIIGYLIFYIIHRNAHRKNKPCTDLESGEE
KKNIINTPVSSLEILLQSFCKLGLIMAYFYMCDRANLFMKENKFYTHSSFFIPIIYILVLG
VFYNENTKETKVLNREQTDEWKGWMQLVILIYHISGASTFLPVYMHIRVLVAAYLFQTGYG
HFSYFWIKGDFGIYRVCQVLFRLNFLVVVLCIVMDRPYQFYYFVPLVTVWFMVIYVTLALW
PQIIQKKANGNCFWHFGLLLKLGFLLLFICFLAYSQGAFEKIFSLWPLSKCFELKGNVYEW
WFRWRLDRYVVFHGMLFAFIYLALQKRQILSEGKGEPLFSNKISNFLLFISVVSFLTYSIW
ASSCKNKAECNELHPSVSVVQILAFILIRNIPGYARSVYSSFFAWFGKISLELFICQYHIW
LAADTRGILVLIPGNPMLNIIVSTFIFVCVAHEISQITNDLAQIIIPKDNSSLLKRLACIA
AFFCGLLILSSIQDKSKH

Transmembrane domains.
amino acids 153-173, 447-467, 495-515, 532-552, 577-597, 624-644, 651-671, 677-697, 693-713, 733-753, 771-791, 808-828, 866-886, 904-924

N-glycosylation sites.
amino acids 182-185, 277-280, 311-314, 323-326, 377-380, 904-907

Tyrosine kinase phosphorylation site.
amino acids 180-187

N-myristoylation sites.
amino acids 47-52, 50-55, 53-58, 81-86, 82-87, 89-94, 93-98, 94-99, 96-101, 97-102, 134-139, 181-186, 312-317, 511-516, 549-554, 746-751

amino acids 144-147

FIGURE 139

MYHCHSGSKPTEKGANEYAYAKWKLCSASAICFIFMIAEVVGGHIAGSLAVVTDAAHLLID
LTSFLLSLFSLWLSSKPPSKRLTFGWHRAEILGALLSILCIWVVTGVLVYLACERLLYPDY
QIQATVMIIVSSCAVAANIVLTVVLHQRCLGHNHKEVQANASVRAAFVHALGDLFQSISVL
ISALIIYFKPEYKIADPICTFIFSILVLASTITILKDFSILLMEGVPKSLNYSGVKELILA
VDGVLSVHSLHIWSLTMNQVILSAHVATAASRDSQVVRREIAKALSKSFTMHSLTIQMESP
VDQDPDCLFCEDPCD

Transmembrane domains.
amino acids 53-73, 91-111, 127-147, 166-186, 202-222

N-glycosylation sites.
amino acids 162-165, 234-237

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 81-84

Tyrosine kinase phosphorylation site.
amino acids 13-20

N-myristoylation sites.
amino acids 7-12, 42-47, 94-99, 228-233

Cation efflux family.
amino acids 78-306

FIGURE 140

MVLSVPVIALGATLGTATSILALCGVTCLCRHMHPKKGLLPRDQDPDLEKAKPSLLGSAQQ
FNVKKSTEPVQPRALLKFPDIYGPRPAVTAPEVINYADYSLRSTEEPTAPASPQPPNDSRL
KRQVTEELFILPQNGVVEDVCVMETWNPEKAASWNQAPKLHYCLDYDCQKAELFVTRLEAV
TSNHDGGCDCYVQGSVANRTGSVEAQTALKKRQLHTTWEEGLVLPLAEEELPTATLTLTLR
TCDRFSRHSVAGELRLGLDGTSVPLGAAQWGELKTSAKEPSAGAGEVLLSISYLPAANRLL
VVLIKAKNLHSNQSKELLGKDVSVKVTLKHQARKLKKKQTKRAKHKINPVWNEMIMFELPD
DLLQASSVELEVLGQDDSGQSCALGHCSLGLHTSGSERSHWEEMLKNPRRQIAMWHQLHL

Signal sequence.
amino acids 1-24

Transmembrane domain.
amino acids 4-24

N-glycosylation sites.
amino acids 118-121, 201-204, 317-320

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 65-68, 342-345

N-myristoylation sites.
amino acids 11-16, 15-20, 189-194, 197-202, 204-209, 261-266, 380-385, 385-390, 396-401, 401-406

C2 domain.
amino acids 304-394

FIGURE 141

MQKIMHISVLLSPVLWGLIFGVSSNSIQIGGLFPRGADQEYSAFRVGMVQFSTSEFRLTPH
IDNLEVANSFAVTNAFCSQFSRGVYAIFGFYDKKSVNTITSFCGTLHVSFITPSFPTDGTH
PFVIQMRPDLKGALLSLIEYYQWDKFAYLYDSDRGLSTLQAVLDSAAEKKWQVTAINVGNI
NNDKKDEMYRSLFQDLELKKERRVILDCERDKVNDIVDQVITIGKHVKGYHYIIANLEFTD
GDLLKIQFGGANVSGFQIVDYDDSLVSKFIERWSTLEEKEYPGAHTTTIKYTSALTYDAVQ
VMTEAFRNLRKQRIEISRRGNAGDCLANPAVPWGQGVEIERALKQVQVEGLSGNIKFDQNG
KRINYTINIMELKTNGPRKIGYWSEVDKMVVTLTELPSGNDTSGLENKTVVVTTILESPYV
MMKKNHEMLEGNERYEGYCVDLAAEIAKHCGFKYKLTIVGDGKYGARDADTKIWNGMVGEL
VYGKADIAIAPLTITLVREEVIDFSKPFMSLGISIMIKKPQKSKPGVFSFLDPLAYEIWMC
IVFAYIGVSVVLFLVSRFSPYEWHTEEFEDGRETQSSESTNEFGIFNSLWFSLGAFMRQGC
DISPRSLSGRIVGGVWWFFTLIIISSYTANLAAFLTVERMVSPIESAEDLSKQTEIAYGTL
DSGSTKEFFRRSKIAVFDKMWTYMRSAEPSVFVRTTAEGVARVRKSKGKYAYLLESTMNEY
IEQRKPCDTMKVGGNLDSKGYGIATPKGSSLGTPVNLAVLKLSEQGVLDKLKNKWWYDKGE
CGAKDSGSKEKTSALSLSNVAGVFYILVGGLGLAMLVALIEFCYKSRAEAKRMKVAKNAQN
INPSSSQNSQNFATYKEGYNVYGIESVKI

Signal sequence.
amino acids 1-21

Transmembrane domains.
amino acids 4-24, 543-563, 624-644, 811-831

N-glycosylation sites.
amino acids 256-259, 370-373, 406-409, 413-416

Tyrosine kinase phosphorylation sites.
amino acids 35-41, 662-667

N-myristoylation sites.
amino acids 17-22, 21-26, 134-140, 181-186, 254-259, 287-292, 355-360, 405-410, 472-477, 593-598, 669-674, 746-751, 760-765, 764-769, 795-800, 823-828

Amidation site.

Ligand-gated ion channel.
amino acids 543-834

Receptor family ligand binding region.
amino acids 56-398.

FIGURE 142

MATTVPDGCRNGLKSKYYRLCDKAEAWGIVLETVATAGVVTSVAFMLTLPILVCKVQDSNR
RKMLPTQFLFLLGVLGIFGLTFAFIIGLDGSTGPTRFFLFGILFSICFSCLLAHAVSLTKL
VRGRKPLSLLVILGLAVGFSLVQDVIAIEYIVLTMNRTNVNVFSELSAPRRNEDFVLLLTY
VLFLMALTFLMSSFTFCGSFTGWKRHGAHIYLTMLLSIAIWVAWITLLMLPDFDRRWDDTI
LSSALAANGWVFLLAYVSPEFWLLTKQRNPMDYPVEDAFCKPQLVKKSYGVENRAYSQEEI
TQGFEETGDTLYAPYSTHFQLQNQPPQKEFSIPRAHAWPSPYKDYEVKKEGS

Transmembrane domains.
amino acids 30-50, 66-86, 98-118, 121-141, 137-157, 174-194, 212-232, 244-264

N-glycosylation site.
amino acids 158-161

N-myristoylation sites.
amino acids 8-13, 38-43, 80-85, 88-93, 102-107, 136-141, 201-206

Amidation site.
amino acids 124-127

7 transmembrane receptor.
amino acids 27-273

FIGURE 143

MLPAQEAAKLYHTNYVRNSRAIGVLWAIFTICFAIVNVVCFIQPYWIGDGVDTPQAGYFGL
FHYCIGNGFSRELTCRGSFTDFSTLPSGAFKAASFFIGLSMMLIIACIICFTLFFFCNTAT
VYKICAWMQLTSAACLVLGCMIFPDGWDSDEVKRMCGEKTDKYTLGACSVRWAYILAIIGI
LDALILSFLAFVLGNRQDSLMAEELKAENKVLLSQYSLE

Transmembrane domains.
amino acids 21-41, 93-113, 127-147, 173-193

N-myristoylation sites.
amino acids 23-28, 67-72, 89-94, 182-187

FIGURE 144

MPGAAAAAAAAAAAMLPAQEAAKLYHTNYVRNSRAIGVLWAIFTICFAIVNVVCFIQPYWI
GDGVDTPQAGYFGLFHYCIGNGFSRELTCRGSFTDFSTLPSGAFKAASFFIGLSMMLIIAC
IICFTLFFFCNTATVYKICAWMQLTSAACLVLGCMIFPDGWDSDEVKRMCGEKTDKYTLGA
CSVRWAYILAIIGILDALILSFLAFVLGNRQDSLMAEELKAENKVLLSQYSLE

Transmembrane domains.
amino acids 1-15, 34-54, 107-127, 141-161, 187-207

N-myristoylation sites.
amino acids 3-8, 37-42, 81-86, 103-108, 196-201

FIGURE 145

MPGAAAAAAAAAAAMLPAQEAAKLYHTNYVRNSRAIGVLWAIFTICFAIVNVVCFIQPYWI
GDGVDTPQAGYFGLFHYCIGNGFSRELTCRGSFTDFSTLPSGAFKAASFFIGLSMMLIIAC
IICFTLFFFCNTATVYKICAWMQLTSAACLVLGCMIFPDGWDSDEVKRMCGEKTDKYTLGA
CSVRWAYILAIIGILDALILSFLAFVLGNRQDSLMAEELKAENKVLLSQYSLE

Transmembrane domains.

amino acids 1-15, 34-54, 107-127, 141-161, 187-207

N-myristoylation sites.

amino acids 3-8, 37-42, 81-86, 103-108, 196-201

FIGURE 146

RPLRLPCRGPRAAGTVSSESLGHPAPGQVTNRSPQFPMELFQSSPRPPLQPVCVCPGLELC
DPPHRLAFSMGAGLFSVVGTLLLPGLAALVQDWRLLQGLGALMSGLLLLFWGFPALFPESP
CWLLATGQVARARKILWRFAEASGVGPGDSSLEENSLATELTMLSARSPQPRYHSPLGLLR
TRVTWRNGLILGFSSLVGGGIRASFRRSLAPQVPTFYLPYFLEAGLEAAALVFLLLTADCC
GRRPVLLLGTMVTGLASLLLLAGAQYLPGWTVLFLSVLGLLASRAVSALSSLFAAEVFPTV
IRGAGLGLVLGAGFLGQAAGPLDTLHGRQGFFLQQVVFASLAVLALLCVLLLPESRSRGLP
QSLQDADRLRRSPLLRGRFRQDHLPLLPPSNSYWAGHTPEQH

Transmembrane domains.
amino acids 67-87, 102-122, 220-240, 249-269, 265-285, 301-321, 338-358

Glycosaminoglycan attachment site.
amino acids 145-148

N-myristoylation sites.
amino acids 14-19, 27-32, 57-62, 101-106, 146-151, 180-185, 191-196, 202-207, 203-208, 228-233, 253-258, 283-288, 308-313, 312-317, 364-369

Amidation site.
amino acids 244-247

Leucine zipper pattern.
amino acids 96-117

Sugar transport proteins.
amino acids 117-126, 241-253, 276-321

FIGURE 147

MPFEQCHLSWKNVRRLYNAVSKELVGEFLQFVQLDKEASDPFSLNELLDELSRKQKEELWQ
RLKNLLTDVLLESPVDGWQVVEAQGEDNMETEHGSKMRKSIEIIYAITSVILASVSVINES
ENYEALLECVIILNGILYALPESERKLQSSIQDLCVTWWEKGLPAKEDTGKTAFVMLLRRS
LETKTGADVCRLWRIHQALYCFDYDLEESGEIKDMLLECFININYIKKEEGRRFLSCLFNW
NINFIKMIHGTIKNQLQGLQKSLMVYIAEIYFRAWKKASGKILEAIENDCIQDFMFHGIHL
PRRSPVHSKVREVLSYFHHQKKVRQGVEEMLYRLYKPILWRGLKARNSEVRSNAALLFVEA
FPIRDPNLHAIEMDSKIQKQFEELYSLLEDPYPMVRSTGILGVCKITSKYWEMMPPTILID
LLKKVTGELAFDTSSADVRCSVFKCLPMILDNKLSHPLLEQLLPALRYSLHDNSEKVRVAF
VDMLLKIKAVRAAKFWKICPMEHILVRLETDSRPVSRRLVSLIFNSFLPVNQPEEVWCERC
VTLVQMNHAAARRFYQYAHEHTACTNIAKLIHVIRHCLNACIQRAVREPPEDEEEEDGREK
ENVTVLDKTLSVNDVACMAGLLEIIVILWKSIDRSMENNKEAKLYTINKFASVLPEYLKVF
KDDRCKIPLFMLMSFMPASAVPPFSCGVISTLRSREEGAVDKSYCTLLDCLCSWGQVGHIL
ELVDNWLPTEHAQAKSNTASKGRVQIHDTRPVKPELALVYIEYLLTHPKNRECLLSAPRKK
LNHLLKALETSKADLESLLQTPGGKPRGFSEAAAPRAFGLHCRLSIHLQHKFCSEGKVYLS
MLEDTGFWLESKILSFIQDQEEDYLKLHRVIYQQIIQTYLTVCKDVVMVGLGDHQFQMQLL
QRSLGIMQTVKGFFYVSLLLDILKEITGSSLIQKTDSDEEVAMLLDTVQKVFQKMLECIAR
SFRKQPEEGLRLLYSVQRPLHEFITAVQSRHTDTPVHRGVLSTLIAGPVVEISHQLRKVSD
VEELTPPEHLSDLPPFSRCLIGIIIKSSNVVRSFLDELKACVASNDIEGIVCLTAAVHIIL
VINAGKHKSSKVREVAATVHRKLKTFMEITLEEDSIERFLYESSSRTLGELLNS

Transmembrane domain.
amino acids 1081-1101
N-glycosylation sites.
amino acids 120-123, 612-615
cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 280-283, 430-433, 1033-1036
Tyrosine kinase phosphorylation sites.
amino acids 99-106, 314-321, 329-337, 385-391, 844-852
N-myristoylation sites.
amino acids 137-142, 189-194, 254-259, 262-267, 698-703, 920-925, 1015-1020
Amidation site.
amino acids 233-236

ATP/GTP-binding site motif A (P-loop).
amino acids 167-174
Leucine zipper pattern.
amino acids 128-149, 135-156, 787-808
Mitochondrial energy transfer proteins signature.
amino acids 982-990

FIGURE 148

MEKRETFVQAVSKELVGEFLQFVQLDKEASDPFSLNELLDELSRKQKEELWQRLKNLLTDV
LLESPVDGWQVVEAQGEDNMETEHGSKMRKSIEIIYAITSVILASVSVINESENYEALLEC
VIILNGILYALPESERKLQSSIQDLCVTWWEKGLPAKEDTGKTAFVMLLRRSLETKTGADV
CRLWRIHQALYCFDYDLEESGEIKDMLLECFININYIKKEEGRRFLSCLFNWNINFIKMIH
GTIKNQLQGLQKSLMVYIAEIYFRAWKKASGKILEAIENDCIQDFMFHGIHLPRRSPVHSK
VREVLSYFHHQKKVRQGVEEMLYRLYKPILWRGLKARNSEVRSNAALLFVEAFPIRDPNLH
AIEMDSEIQKQFEELYSLLEDPYPMVRSTGILGVCKITSKYWEMMPPTILIDLLKKVTGEL
AFDTSSADVRCSVFKCLPMILDNKLSHPLLEQLLPALRYSLHDNSEKVRVAFVDMLLKIKA
VRAAKFWKICPMEHILVRLETDSRPVSRRLVSLIFNSFLPVNQPEEVWCERCVTLVQMNHA
AARRFYQYAHEHTACTNIAKLIHVIRHCLNACIQRAVREPPEDEEEEDGREKENVTVLDKT
LSVNDVACMAGLLEIIVILWKSIDRSMENNKEAKLYTINKFASVLPEYLKVFKDDRCKIPL
FMLMSFMPASAVPPFSCGVISTLRSREEGAVDKSYCTLLDCLCSWGQVGHILELVDNWLPT
EHAQAKSNTASKGRVQIHDTRPVKPELALVYIEYLLTHPKNRECLLSAPRKKLNHLLKALE
TSKADLESLLQTPGGKPRGFSEAAAPRAFGLHCRLSIHLQHKFCSEGKVYLSMLEDTGFWL
ESKILSFIQDQEEDYLKLHRVIYQQIIQTYLTVCKDVVMVGLGDHQFQMQLLQRSLGIMQT
VKGFFYVSLLLDILKEITGSSLIQKTDSDEEVAMLLDTVQKVFQKMLECIARSFRKQPEEG
LRLLYSVQRPLHEFITAVQSRHTDTPVHRGVLSTLIAGPVVEISHQLRKVSDVEELTPPEH
LSDLPPFSRCLIGIIIKSSNVVRSFLDELKACVASNDIEGIVCLTAAVHIILVINAGKHKS
SKVREVAATVHRKLKTFMEITLEEDSIERFLYESSSRTLGELLNS

Transmembrane domain.
amino acids 1072-1092
N-glycosylation sites.
amino acids 111-114, 603-606
cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 3-6, 271-274, 421-424, 1024-1027
Tyrosine kinase phosphorylation site.
amino acids 91-97, 305-312, 320-328, 376-382, 835-843
N-myristoylation sites.
amino acids 128-133, 180-185, 245-250, 253-258, 689-694, 911-916, 1006-1011
Amidation site.
amino acids 224-227

ATP/GTP-binding site motif A (P-loop).
amino acids 158-165
Leucine zipper pattern.
amino acids 119-140, 126-147, 778-799
Mitochondrial energy transfer proteins signature.
amino acids 973-981

FIGURE 149

MNSTLFSQVENHSVHSNFSEKNAQLLAFENDDCHLPLAMIFTLALAYGAVIILGVSGNLAL
IIIILKQKEMRNVTNILIVNLSFSDLLVAIMCLPFTFVYTLMDHWVFGEAMCKLNPFVQCV
SITVSIFSLVLIAVERHQLIINPRGWRPNNRHAYVGIAVIWVLAVASSLPFLIYQVMTDEP
FQNVTLDAYKDKYVCFDQFPSDSHRLSYTTLLLVLQYFGPLCFIFICYFKIYIRLKRRNNM
MDKMRDNKYRSSETKRINIMLLSIVVAFAVCWLPLTIFNTVFDWNHQIIATCNHNLLFLLC
HLTAMISTCVNPIFYGFLNKNFQRDLQFFFNFCDFRSRDDDYETIAMSTMHTDVSKTSLKQ
ASPVAFKKINNNDDNEKI

Transmembrane domains.
amino acids 38-58, 78-98, 117-137, 155-175, 208-228, 260-280, 296-316

N-glycosylation sites.
amino acids 2-5, 11-14, 17-20, 73-76, 81-84, 186-189

Tyrosine kinase phosphorylation sites.
amino acids 247-253, 341-347

N-myristoylation site.
amino acids 54-59

7 transmembrane receptor (rhodopsin family).
amino acids 57-320

FIGURE 150

MVRCDRGLQMLLTTAGAFAAFSLMAIAIGTDYWLYSSAHICNGTNLTMDDGPPPRRARGDL
THSGLWRVCCIEGIYKGHCFRINHFPEDNDYDHDSSEYLLRIVRASSVFPILSTILLLLGG
LCIGAGRIYSRKNNIVLSAGILFVAAGLSNIIGIIVYISSNTGDPSDKRDEDKKNHYNYGW
SFYFGALSFIVAETVGVLAVNIYIEKNKELRFKTKREFLKASSSSPYARMPSYRYRRRRSR
SSSRSTEASPSRDVSPMGLKITGAIPMGELSMYTLSREPLKVTTAASYSPDQEASFLQVHD
FFQQDLKEGFHVSMLNRRTTPV

Signal sequence.
amino acids 1-19

Transmembrane domains.
amino acids 103-123, 140-160, 182-202

N-glycosylation sites.
amino acids 42-45, 45-48

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 240-243, 322-325

N-myristoylation sites.
amino acids 16-21, 43-48, 74-79, 122-127, 262-267

Cell attachment sequence.
amino acids 58-60

PMP-22/EMP/MP20/Claudin family.
amino acids 6-202

FIGURE 151

MARGPGLAPPPLRLPLLLLVLAAVTGHTAAQDNCTCPTNKMTVCSPDGPGGRCQCRALGSG
MAVDCSTLTSKCLLLKARMSAPKNARTLVRPSEHALVDNDGLYDPDCDPEGRFKARQCNQT
SVCWCVNSVGVRRTDKGDLSLRCDELVRTHHILIDLRHRPTAGAFNHSDLDAELRRLFRER
YRLHPKFVAAVHYEQPTIQIELRQNTSQKAAGEVDIGDAAYYFERDIKGESLFQGRGGLDL
RVRGEPLQVERTLIYYLDEIPPKFSMKRLTAGLIAVIVVVVVALVAGMAVLVITNRRKSGK
YKKVEIKELGELRKEPSL

Signal sequence.
amino acids 1-26

Transmembrane domain.
amino acids 274-294

N-glycosylation sites.
amino acids 33-36, 120-123, 168-171, 208-211

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 271-274, 300-303

N-myristoylation sites.
amino acids 59-64, 132-137

Thyroglobulin type-1 repeat.
amino acids 73-145

FIGURE 152

MGKCSGRCTLVAFCCLQLVAALERQIFDFLGYQWAPILANFLHIMAVILGIFGTVQYRSRY
LILYAAWLVLWVGWNAFIICFYLEVGQLSQDRDFIMTFNTSLHRSWWMENGPGCLVTPVLN
SRLALEDHHVISVTGCLLDYPYIEALSSALQIFLALFGFVFACYVSKVFLEEEDSFDFIGG
FDSYGYQAPQKTSHLQLQPLYTSG

Signal sequence.
amino acids 1-21

Transmembrane domains.
amino acids 33-53, 62-82, 147-167

N-glycosylation site.
amino acids 100-103

Tyrosine kinase phosphorylation site.
amino acids 24-32

N-myristoylation sites.
amino acids 50-55, 182-187

FIGURE 153

MPAPRAREQPRVPGERQPLLPRGARGPRRWRRAAGAAVLLVEMLERAAFFGVTANLVLYLN
STNFNWTGEQATRAALVFLGASYLLAPVGGWLADVYLGRYRAVALSLLLYLAASGLLPATA
FPDGRSSFCGEMPASPLGPACPSAGCPRSSPSPYCAFVLYAGLLLLGLAASSVRSNLTSFG
ADQVMDLGRDATRRFFNWFYWSINLGAVLSLLVVAFIQQNISFLLGYSIPVGCVGLAFFIF
LFATPVFITKPPMGSQVSSMLKLALQNCCPQLWQRHSARDRQCARVLADERSPQPGASPQE
DIANFQVLVKILPVMVTLVPYWMVYFQMQSTYVLQGLHLHIPNIFPANPANISVALRAQGS
SYTIPEAWLLLANVVVVLILVPLKDRLIDPLLLRCKLLPSALQKMALGMFFGFTSVIVAGV
LEMERLHYIHHNETVSQQIGEVLYNAAPLSIWWQIPQYLLIGISEIFASIPGLEFAYSEAP
RSMQGAIMGIFFCLSGVGSLLGSSLVALLSLPGGWLHCPKDFGNINNCRMDLYFFLLAGIQ
AVTALLFVWIAGRYERASQGPASHSRFSRDRG

Transmembrane domains.
amino acids 35-55, 73-93, 98-118, 157-177, 199-219, 250-270, 341-361, 369-389, 408-428, 457-477, 499-519, 537-557

N-glycosylation sites.
amino acids 61-64, 66-69, 178-181, 223-226, 356-359, 439-442

Glycosaminoglycan attachment site.
amino acids 503-506

N-myristoylation sites.
amino acids 51-56, 90-95, 116-121, 147-152, 169-174, 209-214, 258-263, 365-370, 414-419, 479-484, 493-498, 497-502, 506-511, 531-536

POT family.
amino acids 101-503

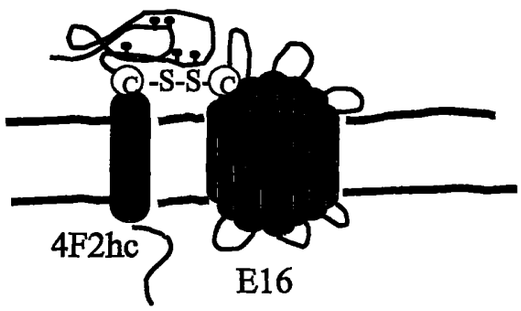
FIGURE 154

MWVNPEEVLLANALWITERANPYFILQRRKGHAGDGGGGGGLAGLLVGTLDVVLDSSARVA
PYRILYQTPDSLVYWTIACGGSRKEITEHWEWLEQNLLQTLSIFENENDITTFVRGKIQGI
IAEYNKINDVKEDDDTEKFKEAIVKFHRLFGMPEEEKLVNYYSCSYWKGKVPRQGWMYLSI
NHLCFYSFLMGREAKLVIRWVDITQLEKNATLLLPDVIKVSTRSSEHFFSVFLNINETFKL
MEQLANIAMRQLLDNEGFEQDRSLPKLKRKSPKKVSALKRDLDARAKSERYRALFRLPKDE
KLDGHTDCTLWTPFNKMHILGQMFVSTNYICFTSKEENLCSLIIPLREVTIVEKADSSSVL
PSPLSISTRNRMTFLFANLKDRDFLVQRISDFLQQTTSKIYSDKEFAGSYNSSDDEVYSRP
SSLVSSSPQRSTSSDADGERQFNLNGNSVPTATQTLMTMYRRRSPEEFNPKLAKEFLKEQA
WKIHFAEYGQGICMYRTEKTRELVLKGIPESMRGELWLLLSGAINEKATHPGYYEDLVEKS
MGKYNLATEEIERDLHRSLPEHPAFQNEMGIAALRRVLTAYAFRNPNIGYCQAMNIVTSVL
LLYAKEEEAFWLLVALCERMLPDYYNTRVVGALVDQGVFEELARDYVPQLYDCMQDLGVIS
TISLSWFLTLFLSVMPFESAVVVVDCFFYEGIKVIFQLALAVLDANVDKLLNCKDDGEAMT
VLGRYLDSVTNKDSTLPPIPHLHSLLSDDVEPYPEVDIFRLIRTSYEKFGTIRADLIEQMR
FKQRLKVIQTLEDTTKRNVVRTIVTETSFTIDELEELYALFKAEHLTSCYWGGSSNALDRH
DPSLPYLEQYRIDFEQFKGMFALLFPWACGTHSDVLASRLFQLLDENGDSLINFREFVSGL
SAACHGDLTEKLKLLYKMHVLPEPSSDQDEPDSAFEATQYFFEDITPECTHVVGLDSRSKQ
GADDGFVTVSLKPDKGKRANSQENRNYLRLWTPENKSKSKNAKDLPKLNQGQFIELCKIMY
NMFSEDPNEQELYHATAAVTSLLLEIGEVGKLFVAQPAKEGGSGGSGPSCHQGIPGVLFPK
KGPGQPYVVESVEPLPASLAPDSEEHSLGGQMEDIKLEDSSPRDNGACSSMLISDDDTKDD
SSMSSYSVLSAGSHEEDKLHCEDIGEDTVLVRSGQGTAALPRSTSLDRDWAITFEQFLASL
LTEPALVKYFDKPVCMMARITSAKNIRMMGKPLTSASDYEISAMSG

Transmembrane domains.
amino acids 35-55, 669-689
N-glycosylation sites.
amino acids 212-215, 239-242, 417-420, 1011-1014
Glycosaminoglycan attachment site.
amino acids 1192-1195
cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 272-275, 277-280, 468-471
N-myristoylation sites.
amino acids 36-41, 37-42, 39-44, 40-45, 44-49, 414-419, 497-502, 515-520, 668-673, 782-787, 845-850, 846-851, 914-919, 977-982, 1078-1083, 1079-1084, 1082-1087, 1144-1149,

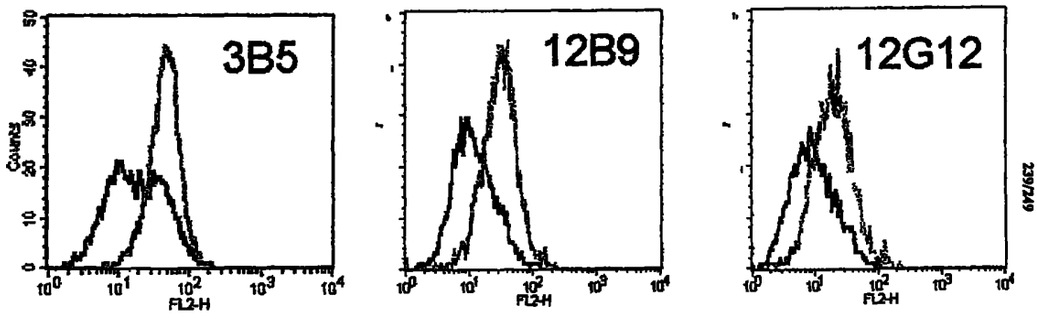
1193-1198
Amidation site.
amino acids 991-994
TBC domain.
amino acids 512-724
S-100/ICaBP type calcium.
amino acids 891-911

Fig 155



448

FIG 156



Green control RNAi

Red: E16 RNAi

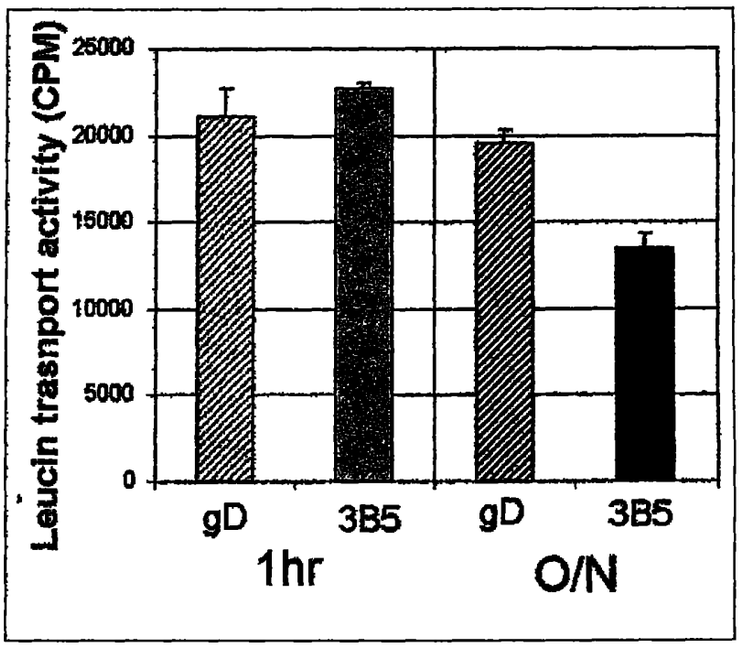
WO 2006/029183 239/249 PCT/US2005/031798

FIG 157



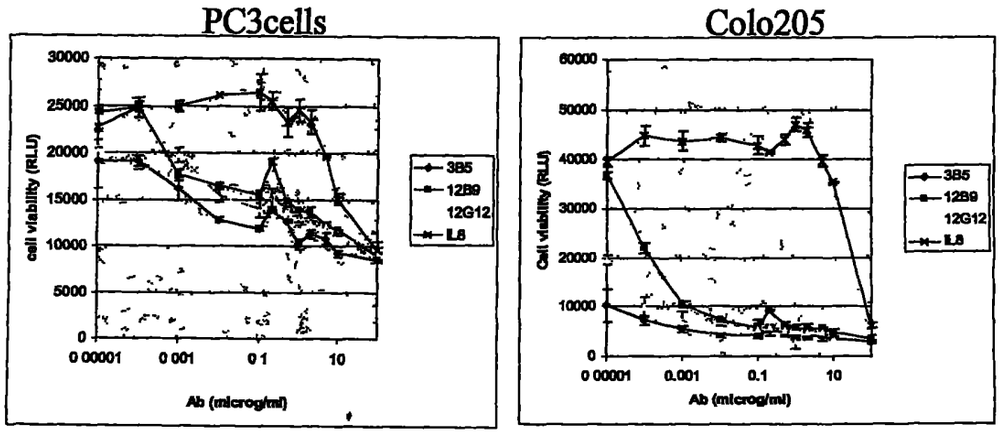
PC3 cells treated with E16 mAb (3B5) for overnight.

FIG 158



WO 2006/029183
241/249
PCT/US2005/031798

Anti-E16-vc-MMAE effect on cancer cell growth in vitro

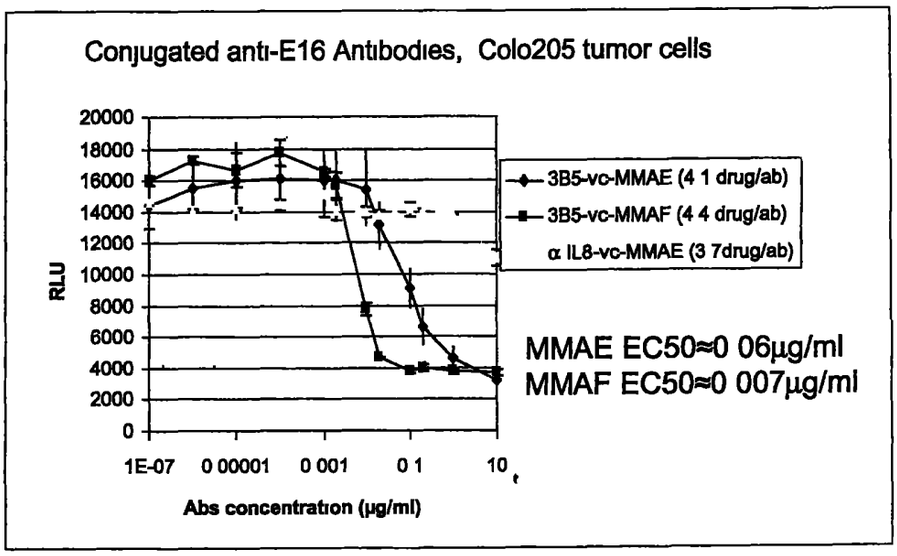


72hrs treatment

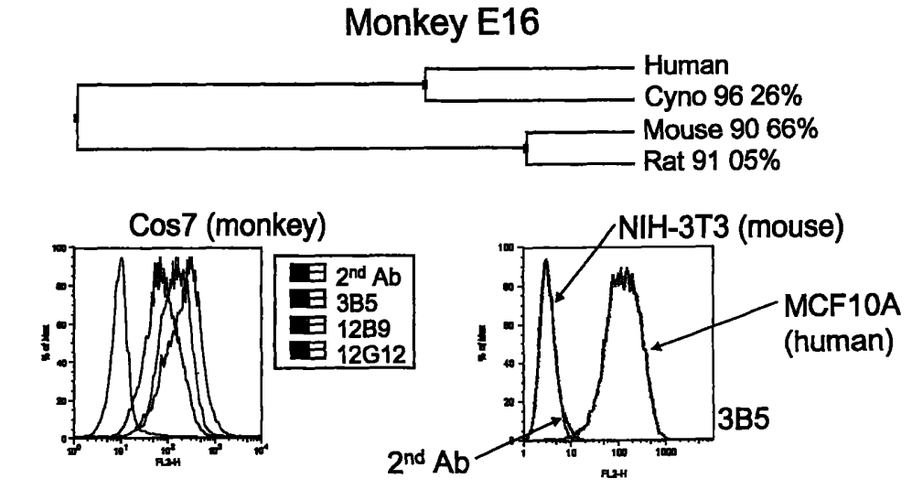
Fig. 159

WO 2006/029183
242/249
PCT/US2005/031798

FIG 160



WO 2006/029183
243/249
PCT/US2005/031798



Anti-E16 antibody do not react with mouse E16, but cross react with monkey E16

FIG 161

WO 2006/029183 244/249 PCT/US2005/031798

Unconjugated anti-E16 Antibodies, PC3 Xenograft

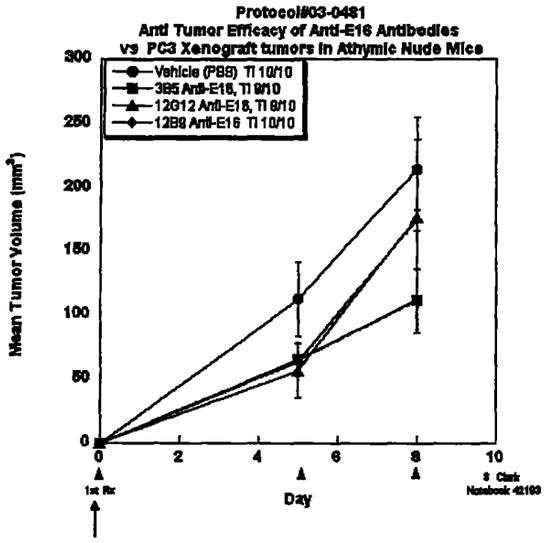


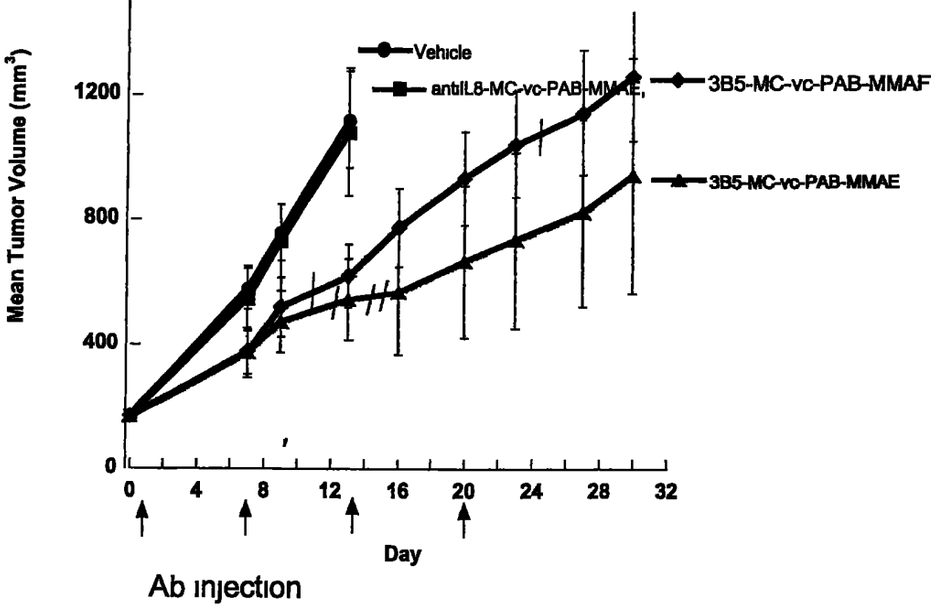
FIG 162

Day0 inject PC3 cells & Ab

WO 2006/029183
245/249
PCT/US2004/031798

FIG 163

Conjugated anti-E16 Antibodies, Colo205 Xenograft



Ab injection

WO 2006/029183
246/249
PCT/US2005/031798

FIG 164

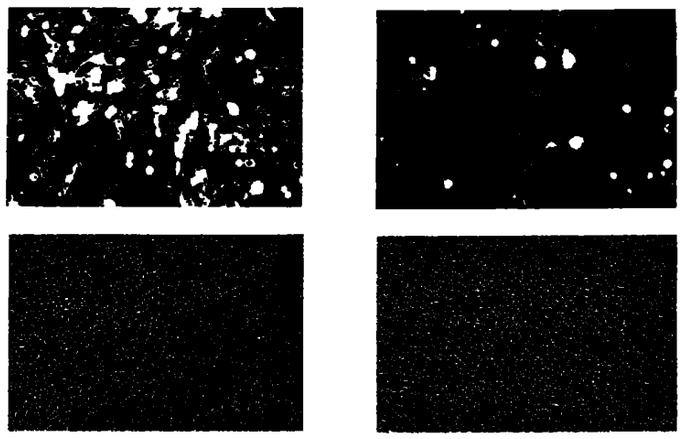


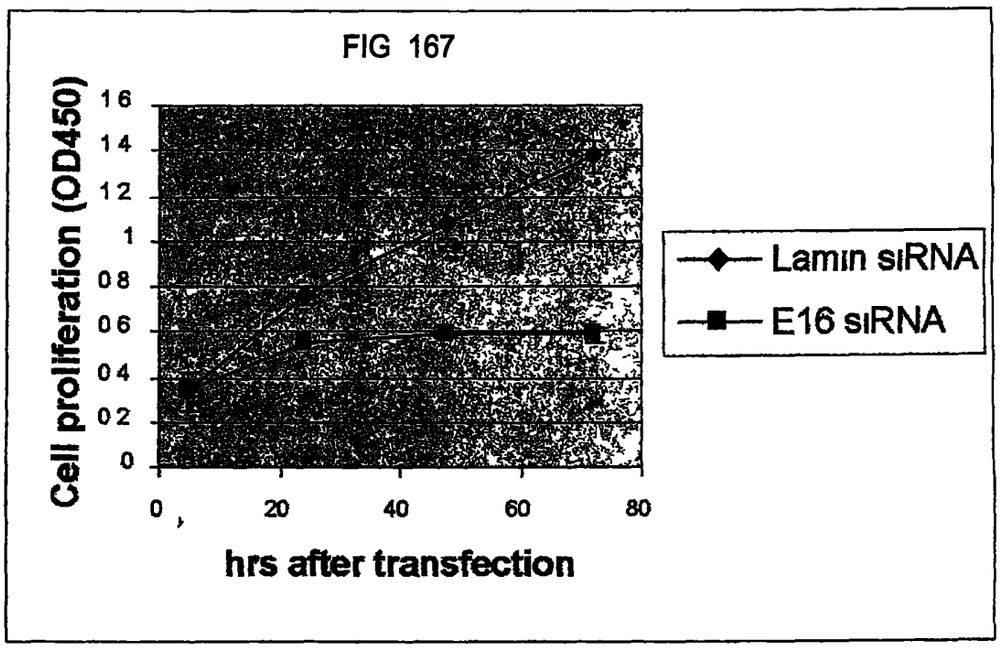
FIG 165



FIG 166

Leucin transport activity
0.6
0.5
0.4
0.3
0.2
0.1
0
Control
E16 RNAi

FIG 167



WO 2006/029183
249/249
PCT/US2005/031798

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2005/031798

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K16/30 A61K47/48 A61K39/395 A61P35/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, Sequence Search, WPI Data, PAJ, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 03/000113 A (GENENTECH, INC) 3 January 2003 (2003-01-03) claims 1-15 page 5, line 20 - line 33 page 6, line 18 - page 9, line 15 figures 1-154; sequences 1-154 page 43, line 10 - page 47, line 5 page 72, line 11 - page 84, line 24 page 87, line 5 - page 109, line 5 examples 12-16 claims 1-15 -/--	1-27, 30-37, 39-46, 51-54, 58-61

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 July 2006	14/08/2006

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Bayer, A

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2005/031798

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HENDERSON MARILYN ET AL: "Microfibril-associated glycoprotein-1 (MAGP-1) is specifically located on the beads of the beaded-filament structure for fibrillin-containing microfibrils as visualized by the rotary shadowing technique" JOURNAL OF HISTOCHEMISTRY AND CYTOCHEMISTRY, vol. 44, no. 12, 1996, pages 1389-1397, XP002379484 ISSN: 0022-1554 abstract page 1390, paragraph 3 - page 1391, paragraph 1 & DATABASE UniProt 1996, retrieved from EBI Database accession no. P55001 abstract	1-3
X	MANNION BRIAN A ET AL: "The light chain of CD98 is identified at E16/TA1 protein" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 273, no. 50, 11 December 1998 (1998-12-11), pages 33127-33129, XP002392000 ISSN: 0021-9258 page 33127, right-hand column, paragraph 4 page 33128, left-hand column, paragraph 2 - paragraph 3 figure 2	1-3, 22-27, 30-32, 39-46
X	EP 1 111 048 A (JAPAN SCIENCE AND TECHNOLOGY CORPORATION; JAPAN SCIENCE AND TECHNOLOGY) 27 June 2001 (2001-06-27) page 4, line 25 - page 5, line 20 page 7, line 33 - line 41 page 8, line 54 - page 9, line 16 page 9, line 25 - line 50 page 19, line 36 - page 23, line 46 page 25, line 29 - page 26, line 36 page 38, line 21 - line 29 claims 20-27,32-40	1-5, 12-15, 22-27, 30-37, 39-46, 51-54, 58-61

-/—

462

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2005/031798

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	DORONINA S O ET AL: "Development of potent monoclonal antibody auristatin conjugates for cancer therapy" NATURE BIOTECHNOLOGY, NATURE PUBLISHING, US, vol. 21, no. 7, July 2003 (2003-07), pages 778-784, XP002280966 ISSN: 1087-0156 cited in the application the whole document	16-21
A	DORONINA SVETLANA ET AL: "Immunoconjugates comprised of drugs with impaired cellular permeability: A new approach to targeted therapy." ABSTRACTS OF PAPERS AMERICAN CHEMICAL SOCIETY, vol. 228, no. Part 1, August 2004 (2004-08), page U908, XP009066098 & MEETING OF THE DIVISION OF CHEMICAL TOXICOLOGY OF THE AMERICAN-CHEMICAL-SOCIETY HELD AT THE 228TH N; PHILADELPHIA, PA, USA; AUGUST 22 -26, 2004 ISSN: 0065-7727 the whole document	20
P,X	US 2005/107595 A1 (CAIRNS BELINDA ET AL) 19 May 2005 (2005-05-19) the whole document	1-27, 30-37, 39-46, 57-61
E	WO 2005/117986 A (GENENTECH, INC; EBENS, JR., ALLEN J; JACOBSON, FREDERIC S; POLAKIS, PA) 15 December 2005 (2005-12-15) page 25, line 17 - page 26, line 6 claims 1,26,30,35,56,64	1,6-11, 16-21, 58-61

Form PCT/ISA/210 (continuation of second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

[redacted] international application No.
PCT/US2005/031798

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.b of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, the international search was carried out on the basis of:

   a. type of material

   - [X] a sequence listing
   - [ ] table(s) related to the sequence listing

   b. format of material

   - [X] on paper
   - [X] in electronic form

   c. time of filing/furnishing

   - [ ] contained in the international application as filed
   - [ ] filed together with the international application in electronic form
   - [X] furnished subsequently to this Authority for the purpose of search

2. [X] In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box II.1

Although claims 58-61 and claims 33-37 as far as in vivo is concerned are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box II.2

Claims Nos.: 28,29,38

The present claims 28, 29 and 38 relate to an extremely large number of possible antibodies defined by amino acid homology to at least one of the CDR(s) of antibodies produced by hybridomas 385.1, 1289.2 and 12612.1. The application, however does not disclose any sequence of said antibodies. Thus, there is no support and disclosure in the sense of Article 6 and 5 PCT. The non-compliance with the substantive provisions is to such an extent, that no search was performed taking into consideration the non-compliance in determining the extent of the search of said claims (PCT Guidelines 9.19 and 9.23).

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure. If the application proceeds into the regional phase before the EPO, the applicant is reminded that a search may be carried out during examination before the EPO (see EPO Guideline C-VI, 8.5), should the problems which led to the Article 17(2) declaration be overcome.

INTERNATIONAL SEARCH REPORT

[redacted] national application No.
PCT/US2005/031798

Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 58-61 and claims 33-37 as far as in vivo is concerned are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.: 28,29,38
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [X] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

1-21, 51-54 (all partially), 22-46, 57-61 (all complete)

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[X] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Inventions 1-74

Claims 1-21, 51-54 (all partially):
Each of the antibodies binding to the polypeptides defined by SEQ-ID NOs: 79-114 and 116-154 and the related subject-matter define one independent invention.

Invention 75

Claims 1-21, 51-54 (all partially); 22-46, 57-61 (all complete):
An isolated antibody binding to the polypeptide defined by SEQ-ID NO: 115, i.e. TAT188 and the related subject-matter.

Invention 76

Claims 47-50, 55, 56 (all complete); 51-54 (all partially):
TAT188 siRNA and the related subject-matter.

INTERNATIONAL SEARCH REPORT

...mation on patent family members

International application No
PCT/US2005/031798

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 03000113	A	03-01-2003	CA	2451239 A1	03-01-2003
			EP	1409549 A2	21-04-2004
			JP	2004537534 T	16-12-2004
			MX	PA03011985 A	26-03-2004
EP 1111048	A	27-06-2001	AU	766306 B2	16-10-2003
			AU	5448799 A	27-03-2000
			CA	2341777 A1	16-03-2000
			DE	69928270 D1	15-12-2005
			WO	0014228 A1	16-03-2000
			JP	2000157286 A	13-06-2000
US 2005107595	A1	19-05-2005	US	2005106644 A1	19-05-2005
WO 2005117986	A	15-12-2005	NONE		

(12) SOLICITUD INTERNACIONAL PUBLICADA BAJ EL TRATADO DE COOPERACIÓN PARA PATENTES (PCT)

(19) Organización Mundial de Propiedad Intelectual
Oficina Internacional

(43) Fecha de Publicación Internacional
Marzo 16, 2006 (16.03.2006)

PCT

(10) Número de Publicación Internacional
WO 2006/029183 A3

(51) Clasificación Internacional de Patente:
C07K 16/30 (2006.01) A61K 39/395 (2006.01)
A61K 47/48 (2006.01) A61P 35/00 (2006.01)

(21) Número de Solicitud Internacional:
PCT/USA2006/031798

(22) Fecha de Presentación Internacional: Septiembre 7, 2005 (07.09.2005)

(25) Idioma de la Presentación: Inglés

(26) Idioma de Publicación: Inglés

(30) Datos de la Prioridad:
10/936.626 Septiembre 8, 2004 (08.09.2004) USA

(71) Solicitante *(Para todos los estados designados salvo* US*)*: GENENTECH, INC. [USA/USA]; MS 49, 1 DNA Way, South San Francisco, California 94080-4990 (USA)

(72) (75) Inventores; e Inventores/Solicitantes *(únicamente para USA)*: CAIRNS, Belinda. [GB/USA]; 129 Paloma Avenue, Burlingame, California 94010 (USA), CHEN, Ruihuan [CN/USA]; 1110 Polynesia Drive, #301, Foster City, CA 94404 (SA). CHUNTHARAPAI, Anan [US/US]; 826 Ellis Drive, Colma, CA 94015 (USA) *(Únicamente USA)*. FRANTZ, Gretchen [USA/USA]; 135 San Benito Way, San Francisco, CA 94127 (USA) *(Únicamente USA)*. HILLAN, Kenneth [GB/USA]; 64 Seward Street, San Francisco, CA 94114 (USA) *(Únicamente USA)*. KOEPPEN, Kartmut [DE/USA]; 1746 Milvia Street, Berkeley, CA 94709 (USA) *(Únicamente USA)*. PHILLIPS, Heidi [USA/USA]; 1511 Walnut Drive, Palo Alto, CA 94303 (USA) *(Únicamente USA)*. POLAKIS, Paul [USA/USA]; 1449 Cortez Avenue, Burlingame, CA 94010 (USA) *(Únicamente USA)*. REED, Chae Janeka [USA/USA]; 2396 Post Street, San Francisco, CA 94115 (USA) *(Únicamente USA)*. SAKANAKA, Chio [JP/USA]; 1481 8th Avenue, San Francisco, CA 94122 (USA) *(Únicamente USA)*. SMITH, Victoria [AU/USA]; 19 Dwight Road, Burlingame, CA 94010 (USA) *(Únicamente USA)*. SPENCER, Susan [USA/USA]; 2395 Paradise Drive, Tiburon, CA 94920 (USA) *(Únicamente USA)*. WILLIAMS, P. Mickey [USA/USA]; 509 Alto Avenue, Half Moon Bay, CA 94019 (USA) *(Únicamente USA)*. WU, Thomas [USA/USA]; 41 Nevada Street, San Francisco, CA 94110 (USA) *(Únicamente USA)*. ZHANG, Zemin [USA/USA]; 876 Taurus Drive, Foster City, CA 94404 (US) *(Únicamente USA)*.

(74) Agentes: CONLEY, Deirdre; MS49 1 DNA Way, South San Francisco, CA 94080-4990 (USA).

(81) Estados Designados (salvo si se indica en contrario, para todas las protecciones nacionales disponibles), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Estados Designados (salvo si se indica en contrario, para todas las protecciones regionales disponibles): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasia (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), Europea (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Publicada:
- *con informe de búsqueda internacional*

(88) *Fecha de publicación del informe de búsqueda internacional:*
Noviembre 16, 2006

Para códigos de dos letras y otras abreviaturas, consulte las "Notas de Guía sobre Códigos y Abreviaturas" que aparece al principio de cada ejemplar regular de la Gaceta de PCT

WO 2006/029183 A3

469

(54) Título : COMPOSICIONES Y MÉTODOS PARA EL DIAGNÓSTICO Y EL TRATAMIENTO DE TUMORES

(57) Extracto: La presente invención se refiere a composiciones de sustancias útiles para el diagnóstico y tratamiento de tumores en mamíferos y a métodos para emplear estas composiciones de sustancias para lo anterior.

COMPOSICIONES Y MÉTODOS PARA EL DIAGNÓSTICO Y TRATAMIENTO DE TUMORES

SOLICITUDES RELACIONADAS

Esta solicitud es una continuación, en parte de, y reivindica la prioridad bajo el 35 U.S.C. §120 de la Solicitud de Estados Unidos Serial No. 10/177,488 presentada el 6/19/2002, la cual reivindica la prioridad bajo el 35 U.S.C. §119(e) de la Solicitud Provisional de Estados Unidos Serial Nos. 60/299,500 presentada el 6/20/2001, 60/301,880 presentada el 6/29/2001 y 60/323,268 presentada el 9/18/2001, cuyas divulgaciones se incorporan a la presente por referencia.

CAMPO DE LA INVENCIÓN

La presente invención se relaciona con composiciones de sustancias útiles para el diagnóstico y tratamiento de tumores en mamíferos y con métodos para utilizar esas composiciones de sustancias para lo anterior.

ANTECEDENTES DE LA INVENCIÓN

Los tumores malignos (cáncer) son la segunda causa de muerte en Estados Unidos, después de las cardiopatías (Boring y colaboradores, *CA Cancel J. Clin.* 43:1 (1993)). El cáncer se caracteriza por un incremento del número de células anormales o neoplásicas, derivadas de tejido normal, las cuales proliferan para formar una masa tumoral, la invasión de tejidos adyacentes por parte de estas células tumorales neoplásicas y la generación de células malignas que eventualmente se propagan a través del sistema sanguíneo o linfático a los nódulos linfáticos regionales y a sitios distantes mediante un proceso denominado metástasis. En un estado canceroso, una célula prolifera bajo condiciones en las cuales las células normales no crecerían. El cáncer se manifiesta en una amplia variedad de formas, y se caracteriza por diferentes grados de invasión y agresividad.

En tentativas por descubrir blancos celulares efectivos para el diagnóstico y tratamiento del cáncer, los investigadores han tratado de identificar polipéptidos transmembrana o de otra manera relacionados con la membrana que se expresan específicamente en la superficie de una o varias clases de células cancerosas cuando se compara(n) con una o varias

células normales no cancerosas. Con frecuencia los citados polipéptidos asociados a la membrana se expresan en forma más abundante en la superficie de las células cancerosas cuando se compara con la superficie de células no cancerosas. La identificación de los citados antígenos de polipéptidos en superficie celular asociadas con un tumor ha dado lugar a la posibilidad de apuntar específicamente a las células de cáncer para la destrucción mediante terapias a base de anticuerpos. En este aspecto, se nota que la terapia a base de anticuerpos ha demostrado ser muy efectiva en el tratamiento de ciertos tipos de cáncer. Por ejemplo, HERCEPTIN® y RITUXAN® (ambos de Genentech Inc., South San Francisco, California) son anticuerpos que se han empleado con éxito para tratar el cáncer de mama y el linfoma no Hodgkin, respectivamente. Más específicamente, HERCEPTIN® es un anticuerpo monoclonal humanizado recombinante derivado de ADN que selectivamente se fija al dominio extracelular del receptor del factor de crecimiento epidérmico humano 2 (HER2) protoncogén. Se observa una sobreexpresión de una proteína HER2 en 25-30% de las neoplasias de mama primarios. RITUXAN® es un anticuerpo monoclonal murino/humano construido genéticamente y dirigido al antígeno CD20 que se encontró en la superficie de linfocitos B normales y malignos. Estos dos anticuerpos se producen de forma recombinante en las células CHO.

En otras tentativas por descubrir otros blancos celulares efectivos para el diagnóstico y tratamiento del cáncer, los investigadores han tratado de identificar (1) polipéptidos no asociados a las membranas que se producen específicamente por un(os) tipo(s) particular(es) de célula(s) cancerosa(s) cuando se compara(n) con uno o varios tipos particulares de célula(s) normal(es) no cancerosa(s), (2) los polipéptidos que se producen por células cancerosas en un nivel de expresión significativamente mayor que el de una o varias célula(s) normal(es) no cancerosa(s) o (3) polipéptidos cuya expresión se limita específicamente a un solo tipo (o a un número muy limitado) de diferente(s) tipo(s) de tejido(s), tanto en los cancerosos y (por ejemplo 1 tejido normal de próstata y un tejido de tumor

de próstata). Los citados polipéptidos pueden permanecer a nivel intracelular o una célula cancerosa los puede secretar. Por otra parte, los citados polipéptidos se pueden expresar, no por una célula cancerosa en sí, sino más bien por las células que producen y/o secretan polipéptidos que tienen un efecto potenciador o intensificador de crecimiento sobre las células cancerosas. Los citados polipéptidos secretados con frecuencia son proteínas que aportan una ventaja de crecimiento a las células cancerosas sobre las células normales e incluyen por ejemplo los factores angiogénicos, los factores de adhesión celular, los factores de crecimiento y similares. Se esperaría que al identificar los antagonistas de dichos polipéptidos no asociados a la membrana estos sirvan como agentes terapéuticos efectivos para el tratamiento de estos tipos de cáncer. Adicionalmente, la identificación del patrón de expresión de los citados polipéptidos sería útil para el diagnóstico y tratamiento de unas neoplasias determinadas en los mamíferos.

La capacidad de modular la expresión del gen en los mamíferos aún es difícil. Tradicionalmente se ha realizado mediante la utilización de herramientas como los vectores virales que van a expresar un polipéptido que no se encuentra presente en el anfitrión. La capacidad de introducir un vector viral en un anfitrión que va a reducir la expresión del gen no se ha perfeccionado. Tradicionalmente la represión de la expresión del gen se ha realizado en los mamíferos en forma de "inactivación", donde al mamífero de prueba, normalmente un ratón, se le extirpa o se le inactiva el gen mediante la técnica de recombinación homóloga. Esta técnica es lenta, dispendiosa y difícil en los ratones, e impráctica en los humanos. Por lo tanto, los solicitantes recurrieron al sistema de vectores que van a expresar un ARN de interferencia (ARNsi) para reducir la expresión del gen.

Los ARNsi han demostrado ser útiles como una herramienta en los estudios para modular la expresión del gen cuando los antagonistas tradicionales como las moléculas pequeñas o los anticuerpos no han tenido éxito. (Shi Y., *Trends in Genetics* 19(1):9-12 (2003)). Los ARNs bicatenarios sintetizados in vitro, que tienen de 21 a 23 nucleótidos de largo

pueden actuar como ARNs de interferencia (ARNsi) y específicamente pueden inhibir la expresión del gen (Fire A., *Trends in Genetics* 391; 806-810 (1999)). Estos ARNsi actúan a través de la mediación de la degradación de sus ARNs objetivo. Aunque tienen menos de 30 nucleótidos de largo, no disparan un mecanismo de defensa celular antiviral. Los citados mecanismos comprenden la producción de interferona, y una suspensión general de la síntesis de proteínas de una célula anfitriona. Prácticamente los ARNs se pueden sintetizar y luego clonar en vectores de ADN. Estos vectores se pueden transinfectar y se pueden poner a expresar el ARNsi a niveles altos y/o en una forma específica para tejidos. El alto nivel de expresión de ARNsi se emplea para "inactivar" o significativamente reducir la cantidad de proteína producida en una célula, y por lo tanto es útil en experimentos donde se cree que una sobreexpresión de una proteína está relacionada con desórdenes como el cáncer. A pesar del progreso en la terapia contra el cáncer en mamíferos, existe una gran necesidad de agentes terapéuticos capaces de inhibir el crecimiento celular neoplásico mediante la reducción de la expresión del gen. Por lo tanto, es un objetivo de la presente invención identificar un sistema que module la expresión del gen.

Mejorar la administración de fármacos y otros agentes a las células blanco, a los tejidos y tumores para alcanzar la máxima eficacia y una toxicidad mínima ha sido el centro de una cantidad considerable de investigación durante muchos años. Aunque se han realizado muchas tentativas por desarrollar unos métodos efectivos para importar unas moléculas biológicamente activas a las células, tanto *in vivo* como *in vitro*, ninguno ha demostrado ser totalmente satisfactorio. La optimización de la asociación del fármaco con su blanco intracelular, minimizando la redistribución intracelular del fármaco, por ejemplo a las células contiguas, con frecuencia es difícil o ineficiente.

La mayoría de los agentes administrados actualmente a un paciente en forma parenteral no tienen un blanco y resultan en una administración sistémica del agente a las células y los tejidos del cuerpo donde no son

necesarios, y con frecuencia son indeseables. Esto puede resultar en unos efectos secundarios adversos y con frecuencia limita la dosis de un fármaco (por ejemplo, agentes quimioterapéuticos (contra el cáncer), citotóxicos, inhibidores de enzimas y fármacos antivirales o antimicrobianos) que se pueden administrar. A modo de comparación, aunque la administración oral de fármacos se considera una forma adecuada y económica de administración, comparte las mismas preocupaciones en relación con la toxicidad no específica para las células sin afectar una vez el fármaco se absorbe en la circulación sistémica. Unas complicaciones adicionales comprenden problemas con la biodisponibilidad oral y la residencia del fármaco en el intestino y por ende las toxicidades en el intestino. Por l tanto, una meta importante es desarrollar unos métodos para agentes con blancos específicos en células y tejidos. Los beneficios del citado tratamiento comprenden evitar unos efectos fisiológicos generales de una administración inapropiada de dichos agentes a otras células y tejidos como las células que no están infectadas. Se pueden impactar los objetivos intracelulares mediante unos métodos, compuestos y formulaciones que permiten la acumulación o la retención de agentes biológicamente activos, es decir metabolitos activos al interior de las células.

Se ha establecido una terapia de anticuerpos monoclonales para el tratamiento dirigido a un blanco para pacientes con cáncer, con desórdenes inmunológicos y angiogénicos.

Existe una variedad de toxinas útiles disponibles. Los péptidos auristatina, auristaina E(AE) y monometilauristatina (MMAE), análogos sintéticos de dolastatina, se conjugaron a: (i) anticuerpos monoclonales quiméricos cBR96 (específico a Lewis Y en carcinomas); (ii) cAC10 que es específico para CD30 en neoplasias hematológicas (Klussman, y colaboradores (2004), Bioconjugate Chemistry 15(4):765-773; Doronina y colaboradores (2003) Nature Biotechnology 21(7):778-784; "Monomethylvalline Compounds Capable of Conjugation to Ligands"; Francisco y colaboradores (2003) Blood 102(4):1458-1465; USA 2004/0018194; (iii) anticuerpos anti-CD20 como rituxan (WO 04/032828)

para el tratamiento de CD20 expresando neoplasias y desórdenes inmunes; (iv) anticuerpos anti-EphB2 2H9 y anti-IL-8 para el tratamiento de cáncer colorrectal (Mao, y colaboradores (2004) Cancer Research 64(3):781-788); (v) anticuerpo E-selectin (Bhaskar y colaboradores (2003) Cancer Res. 63:6387-6394); y (vi) otros anticuerpos anti-CD30 (WO 03/043583). Igualmente se ha conjugado la monometilauristatina (MMAE) a 2H9, un anticuerpo contra EphB2R que es un receptor tirosina cinasa tipo 1 TM con una homología cercana entre ratón y humano, y se sobreexpresa en células cancerosas colorrectales (Mao y colaboradores (2004) Cancer Res. 64:781-788).

Se ha reportado que la monometilauristatina MMAF, una variante de la auristatina E (MMAE) con una fenilalanina en el terminal C, (USA 5767237; USA 6124431), es menos potente que la MMAE, pero es más potente cuando se conjuga a anticuerpos monoclonales (Senter y colaboradores. Proceedings of the American Association for Cancer Research, Volumen 45, Extracto Número 623, presentada el 28 de marzo de 2004). La auristatina F fenileno diamina (AFP); una variante de la fenilalanina de MMAE se enlazó a un anti-CD70 mAb, 1F6, mediante el terminal C de 1F6 a través de un espaciador de fenileno diamina (Law y colaboradores, Proceedings of the American Association for Cancer Research, Volumen 45, Extracto Número 625, presentada el 28 de marzo de 2004).

A pesar del progreso identificado arriba para la terapia contra el cáncer en mamíferos, existe una gran necesidad de unos agentes diagnósticos y terapéuticos adicionales capaces de detectar la presencia de un tumor en un mamífero y para efectivamente inhibir el crecimiento celular neoplásico, respectivamente. Por lo tanto, es un objetivo de la presente invención identificar: (1) polipéptidos asociados a la membrana celular que se expresan en forma más abundante en uno o varios tipo(s) de célula(s) de cáncer cuando se compara(n) con células normales o en otras células cancerosas diferentes, (2) polipéptidos no asociados a las membranas que se producen específicamente por uno o varios tipo(s)

particular(es) de célula(s) de cáncer (o por otras células que producen polipéptidos que tienen un efecto potenciador sobre el crecimiento de la(s) célula(s) cancerosa(s)) cuando se compara(n) con uno o varios tipos particulares de célula(s) normal(es) no cancerosa(s), (3) polipéptidos no asociados a las membranas producidos por las células cancerosas a un nivel de expresión significativamente mayor que el de una o varias célula(s) normal(es) no cancerosa(s), o (4) polipéptidos cuya expresión se limita específicamente a solamente un (o a un número muy limitado de diferentes) tipo(s) de tejido tanto en estado canceroso como no canceroso (por ejemplo, tejido normal de próstata y tejido de tumor de próstata), y para emplear esos polipéptidos, y sus ácidos nucleicos codificados para producir composiciones de sustancias útiles en el tratamiento terapéutico y la detección diagnóstica del cáncer en mamíferos. Igualmente es un objetivo de la presente invención identificar polipéptidos asociados a la membrana celular, secretados o intracelulares cuya expresión se limita a un solo o a un número muy reducido de tejidos, y a emplear esos polipéptidos, y sus ácidos nucleicos codificados para producir composiciones de sustancias útiles en el tratamiento terapéutico y la detección diagnóstica de cáncer en mamíferos.

RESUMEN DE LA INVENCIÓN

A. Incorporaciones

En la presente especificación, los Solicitantes describen por primera vez la identificación de varios polipéptidos celulares (y sus ácidos nucleicos codificados o fragmentos de los mismos) que se expresan a un mayor grado en la superficie de o en uno o varios tipos de célula(s) de cáncer cuando se comparan con la superficie de uno o varios tipos de células normales no cancerosas. Alternativamente, los citados polipéptidos se expresan por células que producen y/o secretan polipéptidos que tienen un efecto de crecimiento potenciador o intensificador sobre las células cancerosas. De nuevo, alternativamente, los citados polipéptidos pueden no estar sobreexpresados por células tumorales cuando se comparan con células normales del mismo tipo de tejido, sino más bien se expresan típicamente

477

tanto por células tumorales y células normales de un solo tipo o de un número muy limitado de tipos de tejido (preferiblemente los tejidos que no son indispensables para la vida, por ejemplo, de próstata, etc.). En la presente se hace referencia a todos los polipéptidos citados anteriormente como polipéptidos Blanco Antigénicos asociados un Tumor (polipéptidos "TAT") y se espera que sirvan como blancos efectivos para la terapia contra el cáncer y el diagnóstico en mamíferos.

Por lo tanto, en una incorporación de la presente invención, la invención dispone una molécula aislada de ácido nucleico que tiene una secuencia nucleótida que codifica un polipéptido blanco antigénico asociado un tumor o un fragmento de la misma (un polipéptido "TAT").

En ciertos aspectos, la molécula aislada de ácido nucleico comprende una secuencia nucleótida que tiene una identidad de secuencia de ácido nucleico por lo menos de alrededor de 80%, alternativamente por lo menos una identidad de secuencia de ácido nucleico de alrededor de 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% con (a) una molécula de ADN que codifica un polipéptido TAT de longitud total teniendo una secuencia de aminoácidos como se divulga en la presente, una secuencia de aminoácidos de polipéptido TAT que carece del péptido de señal como se divulga en la presente, un dominio extracelular de un polipéptido TAT transmembrana, con o sin el péptido de señal, como se divulga en la presente, o cualquier otro fragmento específicamente definido de una secuencia de aminoácidos de polipéptido TAT de longitud total, como se divulga en la presente, o (b) el complemento de la molécula de ADN de (a).

En otros aspectos, la molécula aislada de ácido nucleico comprende una secuencia nucleótida que tiene una identidad de secuencia de ácido nucleico por lo menos de alrededor de 80%, alternativamente por lo menos una identidad de secuencia de ácido nucleico de alrededor de 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% con (a) una molécula de ADN que codifica un polipéptido TAT de longitud total teniendo una secuencia de

aminoácidos como se divulga en la presente, una secuencia de aminoácidos de polipéptido TAT que carece del péptido de señal como se divulga en la presente, una secuencia de codificación de un dominio extracelular de un polipéptido TAT transmembrana, con o sin el péptido de señal, como se divulga en la presente, o cualquier otro fragmento específicamente definido de una secuencia de aminoácidos de polipéptido TAT de longitud total, como se divulga en la presente, o (b) el complemento de la molécula de ADN de (a).

En otros aspectos, la invención se relaciona con una molécula aislada de ácido nucleico que comprende una secuencia nucleótida que tiene una identidad de secuencia de ácido nucleico por lo menos de alrededor de 80%, alternativamente por lo menos alrededor de 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% con (a) una molécula de ADN que codifica el mismo polipéptido maduro codificado por la región de codificación de longitud total de cualquiera de los ADNc de una proteína humana depositados en la ATCC como se divulga en la presente, o (b) el complemento de la molécula de ADN de (a).

Otro aspecto de la invención dispone una molécula aislada de ácido nucleico que comprende una secuencia nucleótida que codifica un polipéptido TAT cuyo dominio de membrana se encuentra eliminado o sino inactivado, o que complementa la citada secuencia nucleótida codificada, donde (los) dominio(s) de transmembrana de lo(s) citado(s) polipéptido(s) se divulga(n) en la presente. Por lo tanto, se contemplan dominios extracelulares solubles de los polipéptidos TAT aquí descritos.

En otros aspectos, la presente invención se relaciona con moléculas aisladas de ácido nucleico que se hibridan a (a) una secuencia nucleótida que codifica un polipéptido TAT teniendo una secuencia de aminoácidos de longitud total como se divulga en la presente, una secuencia de aminoácidos de polipéptido TAT que carece del péptido de señal como se divulga en la presente, un dominio extracelular de un polipéptido TAT transmembrana, con o sin el péptido de señal, como se divulga en la presente o cualquier

otro fragmento específicamente definido de una secuencia de aminoácidos de polipéptido TAT de longitud total como se divulga en la presente, o (b) el complemento de una secuencia nucleótida de (a). Con respecto a lo anterior, una incorporación de la presente invención se relaciona con fragmentos de una secuencia de codificación de polipéptido TAT de longitud total, o el complemento de la misma, como se divulga en la presente, que se pueden emplear como, por ejemplo, sondas de hibridación útiles como, por ejemplo, sondas diagnósticas, sondas oligonucleótidas antisentido, o para codificar fragmentos de un polipéptido TAT de longitud total que opcionalmente puede codificar un polipéptido que comprende un punto de unión para un anticuerpo polipéptido anti TAT, un oligopéptido TAT de unión u otra molécula orgánica pequeña que se fije a un polipéptido TAT. Dichos fragmentos de ácido nucleico tienen normalmente por lo menos alrededor de 5 nucleótidos de longitud, alternativamente por lo menos alrededor de 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, o 1000 nucleótidos de longitud, donde en este contexto el término "alrededor de" significa que la longitud de una secuencia nucleótida de referencia es más o menos el 10% de la citada longitud de referencia. Se nota que unos fragmentos novedosos de una secuencia nucleótida que codifica un polipéptido TAT se puede determinar de forma rutinaria al alinear una secuencia nucleótida que codifica el polipéptido TAT con otras secuencias nucleótidas conocidas empleando cualquier número de programas de alineación de secuencia muy conocidos y determinando cual(es) fragmento(s) de una secuencia de nucleótidos que codifican el polipéptido

TAT son novedosos. Todos estos fragmentos novedosos de secuencias nucleótidas que codifican el polipéptido TAT se contemplan en la presente. Igualmente se contemplan los fragmentos de polipéptidos TAT codificados por estos fragmentos de moléculas de nucleótidos, preferiblemente aquellos fragmentos de polipéptido TAT que comprenden un punto de unión para un anticuerpo anti-TAT, un oligopéptido TAT de unión u otras moléculas orgánicas que se fijan a un polipéptido TAT.

En otra incorporación, la invención dispone polipéptidos TAT aislados codificados por cualquiera de las secuencias de ácido nucleico aislado identificadas arriba.

En determinado aspecto, la invención se relaciona con un polipéptido TAT aislado, que comprende una secuencia de aminoácidos que tiene una identidad de secuencia de aminoácidos por lo menos de alrededor de 80%, alternativamente una identidad de secuencia de aminoácidos por lo menos alrededor de 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100%, con un polipéptido TAT teniendo una secuencia de aminoácidos de longitud total como se divulga en la presente, una secuencia de aminoácidos de polipéptido TAT que carece del péptido de señal como se divulga en la presente, un dominio extracelular de una proteína polipéptido TAT transmembrana, con o sin el péptido de señal, como se divulga en la presente, una secuencia de aminoácidos codificada por cualquiera de las secuencias de ácido nucleico que se divulgan en la presente o cualquier otro fragmento específicamente definido de una secuencia de aminoácidos de polipéptido TAT de longitud total como se divulga en la presente.

En otro aspecto, la invención se relaciona con un polipéptido TAT aislado que comprende una secuencia de aminoácidos que tiene una identidad de secuencia de aminoácidos por lo menos alrededor del 80%, alternativamente una identidad de secuencia de aminoácidos por lo menos alrededor de 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99%, con una secuencia de

aminoácidos codificada por cualquiera de las proteínas humanas ADNc depositadas en la ATCC, como se divulga en la presente.

En un aspecto específico, la invención dispone un polipéptido TAT aislado sin una secuencia de señal del terminal N y/o sen la metionina de inicio y se codifica por una secuencia nucleótida que codifica la citada secuencia de aminoácidos como se describe arriba. Los procesos para producir el mismo también se describen en la presente, donde esos procesos comprenden cultivar una célula hospedera que comprende un vector que comprende la molécula de ácido nucleico codificada adecuada bajo condiciones apropiadas para la expresión del polipéptido TAT y recuperar el polipéptido TAT del cultivo celular.

Otro aspecto de la invención dispone un polipéptido TAT aislado cuyo dominio de membrana se encuentra eliminado o inactivado. Los procesos para producir el mismo también se describen en la presente, donde esos procesos comprenden cultivar una célula hospedera que comprende un vector que comprende la molécula de ácido nucleico codificada adecuada bajo condiciones apropiadas para la expresión del polipéptido TAT y recuperar el polipéptido TAT del cultivo celular.

En otras incorporaciones de la presente invención, la invención dispone vectores que comprenden ADN codificando cualquiera de los polipéptidos descritos en la presente. Las células hospederas que comprenden cualquier citado vector también se disponen. A modo de ejemplo, las células hospederas pueden ser células CHO, células *E. coli* o células de levadura. Adicionalmente se dispone un proceso para producir cualquiera de los polipéptidos descritos en la presente y comprende cultivar células hospederas bajo condiciones apropiadas para la expresión del polipéptido deseado y recuperar el polipéptido deseado del cultivo celular.

En otras incorporaciones, la invención dispone polipéptidos quiméricos aislados que comprenden cualquiera de los polipéptidos TAT aquí descritos fusionados a un polipéptido heterólogo (no-TAT). Ejemplos de dichas moléculas quiméricas comprenden cualquiera de los polipéptidos TAT aquí descritos fusionados a un polipéptido heterólogo, como por

ejemplo, secuencia marcadora de epítopo o una región Fc de una inmunoglobulina.

En otra incorporación, la invención dispone un anticuerpo que se une, preferiblemente específicamente, a cualquiera de los polipéptidos descritos arriba o abajo. Opcionalmente, el anticuerpo es un anticuerpo monoclonal, un fragmento de anticuerpo, un anticuerpo quimérico, un anticuerpo humanizado, un anticuerpo de una sola cadena o un anticuerpo que en forma competitiva inhibe la fijación de un anticuerpo polipéptido anti-TAT a su respectivo epítopo antigénico. Los anticuerpos de la presente invención opcionalmente se pueden conjugar a un agente inhibidor de crecimiento o a un agente citotóxico como una toxina, incluyendo, por ejemplo, un maitansinoide o calicheamicina, un antibiótico, un isótopo radioactivo, una enzima nucleolítica o similares. Los anticuerpos de la presente invención opcionalmente se pueden producir en células CHO o células bacterianas y preferiblemente inducen la muerte de una célula a la cual se fijan. Para efectos diagnósticos, los anticuerpos de la presente invención se pueden marcar en forma detectable, adherido a un soporte sólido o similar.

En otra incorporación de la presente invención que comprende un anticuerpo conjugado a una citotoxina, la toxina se puede seleccionar entre cualquiera de las toxinas conocidas en el arte. En una incorporación, la toxina se enlaza al anticuerpo mediante un enlazante peptídico o peptidomimético. En una incorporación, la toxina se enlaza al anticuerpo mediante un enlazante que comprende valina-citrulina (-vc-). En una incorporación la toxina es MMAE. En una incorporación la toxina es MMAE. En una incorporación, la toxina es una auristatina peptídica.

En otras incorporaciones de la presente invención, la invención dispone vectores que comprenden ADN codificando cualquiera de los anticuerpos aquí descritos. También se dispone una célula hospedera que comprende cualquiera de dichos vectores. A modo de ejemplo, las células hospederas pueden ser células CHO, células *E. coli* o células de levadura. Se dispone además un proceso para producir cualquiera de los anticuerpos

aquí descritos y comprende cultivar células hospederas bajo condiciones apropiadas para la expresión del anticuerpo deseado y recuperar el anticuerpo deseado del cultivo celular.

En otra incorporación, la invención dispone oligopéptidos ("oligopéptidos TAT de unión") que se fijan, preferiblemente específicamente, a cualquiera de los polipéptidos TAT descritos arriba o abajo. Opcionalmente, los oligopéptidos TAT de unión de la presente invención se pueden conjugar a un agente inhibidor de crecimiento o a un agente citotóxico como una toxina, incluyendo, por ejemplo, un maitansinoide o calicheamicina, un antibiótico, un isótopo radioactivo, una enzima nucleolítica o similares. Los oligopéptidos TAT de unión de la presente invención opcionalmente se pueden producir en células CHO o células bacterianas y preferiblemente inducen la muerte de una célula a la cual se fijan. Para efectos diagnósticos, los oligopéptidos TAT de unión de la presente invención se pueden marcar en forma detectable, adherido a un soporte sólido o similar.

En otras incorporaciones de la presente invención, la invención dispone vectores que comprenden ADN codificando cualquiera de los oligopéptidos TAT de unión descrito en la presente. También se disponen las células hospederas que comprenden cualquier dicho vector. A modo de ejemplo, las células hospederas pueden ser células CHO, células *E. coli* o células de levadura. Un proceso para producir cualquiera de los oligopéptidos TAT de unión aquí descritos también se dispone y comprende cultivar células hospederas bajo condiciones apropiadas para la expresión del oligopéptido deseado y recuperar el oligopéptido deseado del cultivo celular.

En otra incorporación, la invención dispone moléculas orgánicas pequeñas ("moléculas orgánicas TAT de unión") que se fijan, preferiblemente específicamente, a cualquiera de los polipéptidos TAT descritos arriba o abajo. Opcionalmente, las moléculas orgánicas TAT de unión de la presente invención se pueden conjugar a un agente inhibidor de crecimiento o a un agente citotóxico como una toxina, incluyendo, por

ejemplo, un maitansinoide o calicheamicina, un antibiótico, un isótopo radioactivo, una enzima nucleolítica o similares. Las moléculas orgánicas TAT de unión de la presente invención preferiblemente inducen la muerte de una célula a la cual se fijan. Para efectos diagnósticos, las moléculas orgánicas TAT de unión de la presente invención se pueden marcar en forma detectable, adherido a un soporte sólido o similar.

En aún otra incorporación, la invención se relaciona con una composición de sustancias que comprende un polipéptido TAT como se describe en la presente, un polipéptido quimérico como se describe en la presente, un anticuerpo anti-TAT como se describe en la presente, un oligopéptido TAT de unión como se describe en la presente, o una molécula orgánica TAT de unión como se describe en la presente, en combinación con un excipiente. Opcionalmente, el excipiente es un excipiente farmacéuticamente aceptable.

En aún otra incorporación, la invención se relaciona con una composición de sustancias que comprende un polipéptido TAT como se describe en la presente, un polipéptido quimérico como se describe en la presente, un anticuerpo anti-TAT como se describe en la presente, un oligopéptido TAT de unión como se describe en la presente, un ARN TAT de interferencia de unión (ARNsi) o una molécula orgánica TAT de unión como se describe en la presente, en combinación con un excipiente. Opcionalmente, el excipiente es un excipiente farmacéuticamente aceptable. Específicamente el polipéptido TAT es TAT188, al cual también se hace referencia en la presente E16 o como TATI88(E16). El anti- TAT188 (al cual también se hace referencia como anti-E16 o los anticuerpos anti-TAT188(E16)) de una incorporación de la invención incluyen sin limitación 3B5, 12B9 y 12G12, como se describe en la presente. El ARNsi de la incorporación incluye siTAT188 (al cual también se hace referencia como siE16 o siTAT188 (E16 o TAT188 ARNsi o E16 ARNsi o TAT188(E16) ARNsi)) como se describe en la presente.

En aún otra incorporación, la invención se relaciona con un artículo de manufactura que comprende un envase y una composición de sustancias

contenidas dentro del envase, donde la composición de las sustancias puede comprender un polipéptido TAT como se describe en la presente, un polipéptido quimérico como se describe en la presente, un anticuerpo anti-TAT como se describe en la presente, un oligopéptido TAT de unión como se describe en la presente, o una molécula orgánica TAT de unión como se describe en la presente. El artículo puede además opcionalmente comprender una etiqueta adherida al envase, o un inserto en el paquete incluido con el envase, que hace referencia al uso de la composición de las sustancias para el tratamiento terapéutico o la detección diagnóstica de un tumor.

Otra incorporación de la presente invención se relaciona con el uso de un polipéptido TAT como se describe en la presente, un polipéptido quimérico como se describe en la presente, un anticuerpo polipéptido anti-TAT como se describe en la presente, un oligopéptido TAT de unión como se describe en la presente, o una molécula orgánica TAT de unión como se describe en la presente, para la preparación de un medicamento útil en el tratamiento de una enfermedad que responde al polipéptido TAT, al polipéptido TAT quimérico, al anticuerpo polipéptido anti-TAT, al oligopéptido TAT de unión o a la molécula orgánica TAT de unión.

B. Incorporaciones Adicionales

Otra incorporación de la presente invención se relaciona con un método para inhibir el crecimiento de una célula que expresa un polipéptido TAT, donde el método comprende poner una célula en contacto con un anticuerpo, un oligopéptido o una molécula orgánica pequeña que se fija al polipéptido TAT, y donde la unión del anticuerpo, del oligopéptido o de la molécula orgánica al polipéptido TAT causa la inhibición del crecimiento de una célula que expresa el polipéptido TAT. En unas incorporaciones preferidas, una célula es una célula de cáncer y la fijación del anticuerpo, del oligopéptido o de la molécula orgánica al polipéptido TAT causa la muerte de una célula que expresa el polipéptido TAT. Opcionalmente, el anticuerpo es un anticuerpo monoclonal, un fragmento de anticuerpo, un

anticuerpo quimérico, un anticuerpo humanizado o un anticuerpo de una sola cadena. Los anticuerpos, los oligopéptidos TAT de unión y las moléculas orgánicas TAT de unión empleados en los métodos de la presente invención opcionalmente se pueden conjugar a un agente inhibidor de crecimiento o a un agente citotóxico como una toxina, incluyendo, por ejemplo, un maitansinoide o calicheamicina, un antibiótico, un isótopo radioactivo, una enzima nucleolítica o similares. Los anticuerpos y los oligopéptidos TAT de unión empleados en los métodos de la presente invención opcionalmente se pueden producir en células CHO o en células bacterianas.

Aún otra incorporación de la presente invención se relaciona con un método para tratar terapéuticamente a un mamífero que tiene un tumor canceroso que comprende células que expresan un polipéptido TAT, donde el método comprende administrar al mamífero una cantidad terapéuticamente efectiva de un anticuerpo, un oligopéptido o una molécula orgánica pequeña que se fija al polipéptido TAT, resultando de esta manera en el tratamiento terapéutico efectivo del tumor. Opcionalmente, el anticuerpo es un anticuerpo monoclonal, un fragmento de anticuerpo, un anticuerpo quimérico, un anticuerpo humanizado, o un anticuerpo de una sola cadena. Los anticuerpos, los oligopéptidos TAT de unión y las moléculas orgánicas TAT de unión empleados en los métodos de la presente invención opcionalmente se pueden conjugar a un agente inhibidor de crecimiento o a un agente citotóxico como una toxina, incluyendo, por ejemplo, un maitansinoide o calicheamicina, un antibiótico, un isótopo radioactivo, una enzima nucleolítica o similares. Los anticuerpos y oligopéptidos empleados en los métodos de la presente invención opcionalmente se pueden producir en células CHO o en células bacterianas.

Aún otra incorporación de la presente invención se relaciona con un método para determinar la presencia de un polipéptido TAT en una muestra que se sospecha contiene el polipéptido TAT, donde el método comprende exponer la muestra a un anticuerpo, a un oligopéptido o a una molécula orgánica pequeña que se fija al polipéptido TAT y determinar la fijación del

anticuerpo, del oligopéptido o de la molécula orgánica al polipéptido TAT en la muestra, donde la presencia de la citada fijación indica la presencia del polipéptido TAT en la muestra. Opcionalmente, la muestra puede contener células (que pueden ser células cancerosas) que se sospecha expresan el polipéptido TAT. El anticuerpo, el oligopéptido TAT de unión o la molécula orgánica TAT de unión empleados en el método opcionalmente se pueden marcar en forma detectable, adherido a un soporte sólido o similar.

Otra incorporación de la presente invención se relaciona con un método para diagnosticar la presencia de un tumor en un mamífero, donde el método comprende detectar el nivel de expresión de un gen que codifica un polipéptido TAT (a) en una muestra de prueba de células de tejido obtenida del citado mamífero y (b) en una muestra de control de células que se sabe son normales y no son cancerosas del mismo tejido de origen o tipo de tejido, donde un nivel de expresión mayor del polipéptido TAT en la muestra de prueba, cuando se compara con la muestra de control, indica la presencia de un tumor en el mamífero del cual se tomó la muestra de prueba.

Otra incorporación de la presente invención se relaciona con un método para diagnosticar la presencia de un tumor en un mamífero, donde el método comprende (a) poner en contacto una muestra de prueba que comprende células de tejido obtenidas del mamífero con un anticuerpo, con un oligopéptido o con una molécula orgánica pequeña que se fija con un polipéptido TAT y (b) detectar la formación de un complejo entre el anticuerpo, el oligopéptido o la molécula orgánica pequeña y el polipéptido TAT en la muestra de prueba, donde la formación de un complejo indica la presencia de un tumor en el mamífero. Opcionalmente, el anticuerpo, el oligopéptido TAT de unión o la molécula orgánica TAT de unión empleados se marcan en forma detectable, se adhieren a un soporte sólido o similar, y/o la muestra de prueba de las células de tejido se obtiene de un individuo que se sospecha tiene un tumor canceroso.

Aún otra incorporación de la presente invención se relaciona con un método para tratar o prevenir un desorden de proliferación celular asociado

con la expresión o actividad alterada, preferiblemente aumentada, de un polipéptido TAT, el método que comprende administrar a un sujeto que necesite el citado tratamiento una cantidad efectiva de un antagonista de un polipéptido TAT. Preferiblemente, el desorden de proliferación celular es cáncer y el antagonista del polipéptido TAT es un anticuerpo polipéptido anti-TAT, un oligopéptido TAT de unión, una molécula orgánica TAT de unión o un oligonucleótido antisentido. El tratamiento efectivo o la prevención del desorden de proliferación celular puede ser el resultado de una muerte directa o de la inhibición del crecimiento de las células que expresan un polipéptido TAT o de antagonizar la actividad potenciadora del crecimiento celular de un polipéptido TAT.

Aún otra incorporación de la presente invención se relaciona con un método para fijar un anticuerpo, un oligopéptido o una molécula orgánica pequeña a una célula que expresa un polipéptido TAT, donde el método comprende poner en contacto una célula que expresa un polipéptido TAT con el citado anticuerpo, oligopéptido o molécula orgánica pequeña bajo condiciones que son apropiadas para la fijación del anticuerpo, del oligopéptido o de la molécula orgánica pequeña al citado polipéptido TAT y permitir que se fijen entre ellos.

Otras incorporaciones de la presente invención se refieren al uso de (a) un polipéptido TAT, (b) un ácido nucleico codificando un polipéptido TAT o un vector o célula hospedera que comprende ese ácido nucleico, (c) un anticuerpo polipéptido anti-TAT, (d) un oligopéptido TAT-de unión o (e) una molécula orgánica pequeña TAT-de unión en la preparación de un medicamento útil para (i) el tratamiento terapéutico o la detección diagnóstica de un cáncer o un tumor o (ii) el tratamiento terapéutico o la prevención de un desorden de proliferación celular.

Otra incorporación de la presente invención se relaciona con un método para inhibir el crecimiento de una célula cancerosa, donde el crecimiento de la citada célula cancerosa depende por lo menos en parte del (los) efecto(s) potenciador(es) de crecimiento de un polipéptido TAT (donde el polipéptido TAT se puede expresar bien sea por una célula

cancerosa en sí o una célula que produce polipéptido(s) que tiene(n) un efecto potenciador de crecimiento sobre las células cancerosas), donde el método comprende poner en contacto el polipéptido TAT con un anticuerpo, un oligopéptido o una molécula orgánica pequeña que se fija al polipéptido TAT, en esta forma antagonizando la actividad potenciadora de crecimiento del polipéptido TAT y a la vez inhibir el crecimiento de una célula cancerosa. Preferiblemente el crecimiento de una célula cancerosa está totalmente inhibido. Aún más preferiblemente, la fijación del anticuerpo, del oligopéptido o de la molécula orgánica pequeña al polipéptido TAT induce la muerte de una célula cancerosa. Opcionalmente, el anticuerpo es un anticuerpo monoclonal, un fragmento de anticuerpo, un anticuerpo quimérico, un anticuerpo humanizado o un anticuerpo de una sola cadena. Los anticuerpos, los oligopéptidos TAT de unión y las moléculas orgánicas TAT de unión empleados en los métodos de la presente invención opcionalmente se pueden conjugar a un agente inhibidor de crecimiento o a un agente citotóxico como una toxina, incluyendo, por ejemplo, un maitansinoide o calicheamicina, un antibiótico, un isótopo radioactivo, una enzima nucleolítica o similares. Los anticuerpos y los oligopéptidos TAT de unión empleados en los métodos de la presente invención opcionalmente se pueden producir en células CHO o células bacterianas.

Aún otra incorporación de la presente invención se relaciona con un método para tratar terapéuticamente un tumor en un mamífero, donde el crecimiento del citado tumor depende por lo menos en parte del (los) efecto(s) potenciador(es) de crecimiento de un polipéptido TAT, donde el método comprende administrar al mamífero una cantidad terapéuticamente efectiva de un anticuerpo, un oligopéptido o una molécula orgánica pequeña que se fija al polipéptido TAT, en esta forma antagonizando la actividad potenciadora de crecimiento del citado polipéptido TAT y resultando en el tratamiento terapéutico efectivo del tumor. Opcionalmente, el anticuerpo es un anticuerpo monoclonal, un fragmento de anticuerpo, un anticuerpo quimérico, un anticuerpo humanizado o un anticuerpo de una sola cadena.

Los anticuerpos, los oligopéptidos TAT de unión y las moléculas orgánicas TAT de unión empleados en los métodos de la presente invención opcionalmente se pueden conjugar a un agente inhibidor de crecimiento o a un agente citotóxico como una toxina, incluyendo, por ejemplo, un maitansinoide o calicheamicina, un antibiótico, un isótopo radioactivo, una enzima nucleolítica o similares. Los anticuerpos y los oligopéptidos empleados en los métodos de la presente invención opcionalmente se pueden producir en células CHO o células bacterianas.

C. Otras Incorporaciones Adicionales

En aún otras incorporaciones, la invención se relaciona con el siguiente conjunto de reivindicaciones potenciales para esta solicitud:

1. Ácido nucleico aislado que tiene una secuencia nucleótida que tiene una identidad de secuencia de ácido nucleico por lo menos de alrededor de 80%, con:

(a) una molécula de ADN codificando una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) una molécula de ADN codificando una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) una molécula de ADN codificando un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos79-154), con su péptido de señal asociado;

(d) una molécula de ADN codificando un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78);

(f) la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

491

(g) el complemento de (a), (b), (c), (d), (e) o (f).

2. Ácido nucleico aislado teniendo:

(a) una secuencia nucleótida que codifica una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) una secuencia nucleótida que codifica una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) una secuencia nucleótida que codifica un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

(d) una secuencia nucleótida que codifica un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78);

(f) la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(g) el complemento de (a), (b), (c), (d), (e) o (f).

3. Ácido nucleico aislado que se hibrida con:

(a) un ácido nucleico que codifica una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) un ácido nucleico que codifica una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) un ácido nucleico que codifica un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

(d) un ácido nucleico que codifica un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78);

(f) la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(g) el complemento de (a), (b), (c), (d), (e) o (f).

4. El ácido nucleico de la Reivindicación 3, donde la hibridación ocurre bajo condiciones severas.

5. El ácido nucleico de la Reivindicación 3 que tiene por lo menos alrededor de 5 nucleótidos de largo.

6. Un vector de expresión que comprende el ácido nucleico de la Reivindicación 1, 2 o 3.

7. El vector de expresión de la Reivindicación 6, donde el citado ácido nucleico está enlazado operacionalmente a las secuencias de control reconocidas por una célula hospedera transformada con el vector.

8. Una célula hospedera que comprende el vector de expresión de la Reivindicación 7.

9. La célula hospedera de la Reivindicación 8 que es una célula CHO, una célula *E. coli* o una célula de levadura.

10. Un proceso para producir un polipéptido que comprende cultivar la célula hospedera de la Reivindicación 8 bajo condiciones apropiadas para la expresión del citado polipéptido y recuperar el citado polipéptido del cultivo celular.

11. Un polipéptido aislado teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con:

(a) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

(d) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) un polipéptido codificado por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

12. Un polipéptido aislado teniendo:

(a) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(c) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su de péptido de señal asociado;

(d) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(e) una secuencia de aminoácidos codificada por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos:1-78); o

(f) una secuencia de aminoácidos codificada por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

13. Un polipéptido quimérico que comprende el polipéptido de la Reivindicación 11 o 12 fusionado a un polipéptido heterólogo.

14. El polipéptido quimérico de la Reivindicación 13, donde el citado polipéptido heterólogo es una secuencia marcadora de epítopo o una región Fc de una inmunoglobulina.

15. Un anticuerpo aislado que se fija a un polipéptido teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con:

(a) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

d) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) un polipéptido codificado por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

16. Un anticuerpo aislado que se fija a un polipéptido teniendo:

(a) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(c) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su secuencia de péptido de señal asociado;

(d) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(e) una secuencia de aminoácidos codificada por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(f) una secuencia de aminoácidos codificada por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

17. El anticuerpo de la Reivindicación 15 o 16 el cual es un anticuerpo monoclonal.

18. El anticuerpo de la Reivindicación 15 o 16 el cual es un fragmento de anticuerpo.

19. El anticuerpo de la Reivindicación 15 o 16 el cual es un anticuerpo quimérico o humanizado.

20. El anticuerpo de la Reivindicación 15 o 16 el cual se conjuga con un agente inhibidor de crecimiento.

21. El anticuerpo de la Reivindicación 15 o 16 el cual se conjuga con un agente citotóxico.

22. El anticuerpo de la Reivindicación 21, donde el agente citotóxico se selecciona entre el grupo conformado por toxinas, antibióticos, isótopos radioactivos y enzimas nucleolíticas.

23. El anticuerpo de la Reivindicación 21, donde el agente citotóxico es una toxina.

24. El anticuerpo de la Reivindicación 23, donde la toxina se selecciona entre el grupo conformado por maitansinoide y calicheamicina.

25. El anticuerpo de la Reivindicación 23, donde la toxina es un maitansinoide.

26. El anticuerpo de la Reivindicación 15 o 16 el cual se produce en bacterias.

27. El anticuerpo de la Reivindicación 15 o 16 el cual se produce en células CHO.

28. El anticuerpo de la Reivindicación 15 o 16 que induce la muerte de una célula a la cual se fija.

29. El anticuerpo de la Reivindicación 15 o 16 el cual se marca en forma detectable.

30. Un ácido nucleico aislado que tiene una secuencia nucleótida que codifica el anticuerpo de la Reivindicación 15 o 16.

31. Un vector de expresión que comprende el ácido nucleico de la Reivindicación 30 operacionalmente enlazado a las secuencias de control reconocidas por una célula hospedera transformada con el vector.

32. Una célula hospedera que comprende el vector de expresión de la Reivindicación 31.

33. La célula hospedera de la Reivindicación 32 la cual es una célula CHO, una célula *E. coli* o una célula de levadura.

34. Un proceso para producir un anticuerpo que comprende cultivar la célula hospedera de la Reivindicación 32 bajo condiciones apropiadas para la expresión del citado anticuerpo y recuperar el citado anticuerpo del cultivo celular.

35. Un oligopéptido aislado que se fija a un polipéptido teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con:

(a) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

(d) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) un polipéptido codificado por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

36. Un oligopéptido aislado que se fija a un polipéptido teniendo:

(a) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(c) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su secuencia de péptido de señal asociado;

(d) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(e) una secuencia de aminoácidos codificada por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos:1-78); o

(f) una secuencia de aminoácidos codificada por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

37. El oligopéptido de la Reivindicación 35 o 36 el cual se conjuga con un agente inhibidor de crecimiento.

38. El oligopéptido de la Reivindicación 35 o 36 el cual se conjuga con un agente citotóxico.

39. El oligopéptido de la Reivindicación 38, donde el agente citotóxico se selecciona entre el grupo conformado por toxinas, antibióticos, isótopos radioactivos y enzimas nucleolíticas.

40. El oligopéptido de la Reivindicación 38, donde el agente citotóxico es una toxina.

41. El oligopéptido de la Reivindicación 40, donde la toxina se selecciona entre el grupo conformado por maitansinoide y calicheamicina.

42. El oligopéptido de la Reivindicación 40, donde la toxina es un maitansinoide.

43. El oligopéptido de la Reivindicación 35 o 36 que induce la muerte de una célula a la cual se fija.

44. El oligopéptido de la Reivindicación 35 o 36 el cual se marca en forma detectable.

45. Una molécula orgánica TAT de unión que se fija a un polipéptido teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con:

(a) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

(d) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) un polipéptido codificado por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

46. La molécula orgánica de la Reivindicación 45 que se fija a un polipéptido teniendo:

(a) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(c) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su secuencia de péptido de señal asociado;

(d) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(e) una secuencia de aminoácidos codificada por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos:1-78); o

(f) una secuencia de aminoácidos codificada por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

47. La molécula orgánica de la Reivindicación 45 o 46 la cual se conjuga con un agente inhibidor de crecimiento.

48. La molécula orgánica de la Reivindicación 45 o 46 la cual se conjuga con un agente citotóxico.

49. La molécula orgánica de la Reivindicación 48, donde el agente citotóxico se selecciona entre el grupo conformado por toxinas, antibióticos, isótopos radioactivos y enzimas nucleolíticas.

50. La molécula orgánica de la Reivindicación 48, donde el agente citotóxico es una toxina.

51. La molécula orgánica de la Reivindicación 50, donde la toxina se selecciona entre el grupo conformado por maitansinoide y calicheamicina.

52. La molécula orgánica de la Reivindicación 50, donde la toxina es un maitansinoide.

53. La molécula orgánica de la Reivindicación 45 o 46 que induce la muerte de una célula a la cual se fija.

54. La molécula orgánica de la Reivindicación 45 o 46 la cual se marca en forma detectable.

55. Una composición de sustancias que comprende:

(a) el polipéptido de la Reivindicación 11;

(b) el polipéptido de la Reivindicación 12;

(c) el polipéptido quimérico de la Reivindicación 13;

(d) el anticuerpo de la Reivindicación 15;

(e) el anticuerpo de la Reivindicación 16;

(f) el oligopéptido de la Reivindicación 35;

(g) el oligopéptido de la Reivindicación 36;

(h) la molécula orgánica TAT de unión de la Reivindicación 45; o

(i) la molécula orgánica TAT de unión de la Reivindicación 46; en combinación con un excipiente.

56. La composición de las sustancias de la Reivindicación 55, donde el citado excipiente es un excipiente farmacéuticamente aceptable.

57. Un artículo de manufactura que comprende:

(a) un envase; y

(b) la composición de las sustancias de la Reivindicación 55 contenida en el citado envase.

58. El artículo de manufactura de la Reivindicación 57 que además comprende una etiqueta adherida al citado envase, o un inserto en el paquete incluido con el citado envase, haciendo referencia al uso de la citada composición de sustancias para el tratamiento terapéutico o la detección diagnóstica de un cáncer.

59. Un método para inhibir el crecimiento de una célula que expresa una proteína teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con:

(a) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

(d) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) un polipéptido codificado por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos:1-78); o

(f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78), el citado método que

comprende poner en contacto la citada célula con un anticuerpo, un oligopéptido o una molécula orgánica que se fija a la citada proteína, la fijación del citado anticuerpo, oligopéptido o molécula orgánica a la citada proteína causando en esta forma una inhibición del crecimiento de la citada célula.

60. El método de la Reivindicación 59, donde el citado anticuerpo es un anticuerpo monoclonal.

61. El método de la Reivindicación 59, donde el citado anticuerpo es un fragmento de anticuerpo.

62. El método de la Reivindicación 59, donde el citado anticuerpo es un anticuerpo quimérico o humanizado.

63. El método de la Reivindicación 59, donde el citado anticuerpo, el citado oligopéptido o la citada molécula orgánica se conjuga a un agente inhibidor de crecimiento.

64. El método de la Reivindicación 59, donde el citado anticuerpo, el citado oligopéptido o la citada molécula orgánica se conjuga a un agente citotóxico.

65. El método de la Reivindicación 64, donde el citado agente citotóxico se selecciona entre el grupo conformado por toxinas, antibióticos, isótopos radioactivos y enzimas nucleolíticas.

66. El método de la Reivindicación 64, donde el agente citotóxico es una toxina.

67. El método de la Reivindicación 66, donde la toxina se selecciona entre el grupo conformado por maitansinoide y calicheamicina.

68. El método de la Reivindicación 66, donde la toxina es un maitansinoide.

69. El método de la Reivindicación 59, donde el citado anticuerpo se produce en bacterias.

70. El método de la Reivindicación 59, donde el citado anticuerpo se produce en células CHO.

71. El método de la Reivindicación 59, donde la citada célula es una célula cancerosa.

502

72. El método de la Reivindicación 71, donde la citada célula cancerosa se expone además a un tratamiento de radiación o un agente quimioterapéutico.

73. El método de la Reivindicación 71, donde la citada célula cancerosa se selecciona entre el grupo conformado por una célula de cáncer de mama, una célula de cáncer colorrectal, una célula de cáncer de pulmón, una célula de cáncer ovárico, una célula de cáncer del sistema nervioso central, una célula de cáncer de hígado, una célula de cáncer de vejiga, una célula de cáncer pancreático, una célula de cáncer cervical, una célula de melanoma y una célula de leucemia.

74. El método de la Reivindicación 71, donde la citada proteína se expresa en forma más abundante por la citada célula cancerosa cuando se compara con una célula normal del mismo tejido de origen.

75. El método de la Reivindicación 59 que causa la muerte de la citada célula.

76. El método de la Reivindicación 59, donde la citada proteína tiene:

(a) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(c) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su secuencia de péptido de señal asociado;

(d) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(e) una secuencia de aminoácidos codificada por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

503

(f) una secuencia de aminoácidos codificada por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

77. Un método para tratar terapéuticamente a un mamífero teniendo un tumor canceroso que comprende células que expresan una proteína teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con:

(a) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

(d) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) un polipéptido codificado por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos:1-78); o

(f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos:1-78), el citado método que comprende administrar al citado mamífero una cantidad terapéuticamente efectiva de un anticuerpo, un oligopéptido o una molécula orgánica que se fija a la citada proteína, tratando efectivamente en esta forma al citado mamífero.

78. El método de la Reivindicación 77, donde el citado anticuerpo es un anticuerpo monoclonal.

79. El método de la Reivindicación 77, donde el citado anticuerpo es un fragmento de anticuerpo.

80. El método de la Reivindicación 77, donde el citado anticuerpo es un anticuerpo quimérico o humanizado.

81. El método de la Reivindicación 77, donde el citado anticuerpo, oligopéptido o molécula orgánica se conjuga a un agente inhibidor de crecimiento.

82. El método de la Reivindicación 77, donde el citado anticuerpo, oligopéptido o molécula orgánica se conjuga a un agente citotóxico.

83. El método de la Reivindicación 82, donde el citado agente citotóxico se selecciona entre el grupo conformado por toxinas, antibióticos, isótopos radioactivos y enzimas nucleolíticas.

84. El método de la Reivindicación 82, donde el agente citotóxico es una toxina.

85. El método de la Reivindicación 84, donde la toxina se selecciona entre el grupo conformado por maitansinoide y calicheamicina.

86. El método de la Reivindicación 84, donde la toxina es un maitansinoide.

87. El método de la Reivindicación 77, donde el citado anticuerpo se produce en bacterias.

88. El método de la Reivindicación 77, donde el citado anticuerpo se produce en células CHO.

89. El método de la Reivindicación 77, donde el citado tumor se expone además a un tratamiento de radiación o un agente quimioterapéutico.

90. El método de la Reivindicación 77, donde el citado tumor es un tumor de mama, un tumor colorrectal, un tumor de pulmón, un tumor ovárico, un tumor del sistema nervioso central, un tumor de hígado, un tumor de vejiga, un tumor pancreático o un tumor cervical.

91. El método de la Reivindicación 77, donde la citada proteína se expresa en forma más abundante por las células cancerosas del citado tumor cuando se compara con una célula normal del mismo tejido de origen.

92. El método de la Reivindicación 77, donde la citada proteína tiene:

(a) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(c) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su secuencia de péptido de señal asociado;

(d) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(e) una secuencia de aminoácidos codificada por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(f) una secuencia de aminoácidos codificada por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

93. Un método para determinar la presencia de una proteína en una muestra que se sospecha contiene la citada proteína, donde la citada proteína tiene una identidad de secuencia de aminoácidos por lo menos de 80% con:

(a) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

(d) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) un polipéptido codificado por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78), el citado método que comprende exponer la citada muestra a un anticuerpo, a un oligopéptido o a una molécula orgánica que se fija a la citada proteína y determinar la fijación del citado anticuerpo, oligopéptido o molécula orgánica a la citada proteína en la citada muestra, donde la fijación del anticuerpo, del oligopéptido o de la molécula orgánica a la citada proteína indica la presencia de la citada proteína en la citada muestra.

94. El método de la Reivindicación 93, donde la citada muestra comprende una célula que se sospecha expresa la citada proteína.

95. El método de la Reivindicación 94, donde la citada célula es una célula cancerosa.

96. El método de la Reivindicación 93, donde el citado anticuerpo, oligopéptido o molécula orgánica se marca en forma detectable.

97. El método de la Reivindicación 93, donde la citada proteína tiene:

(a) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(c) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su secuencia de péptido de señal asociado;

(d) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(e) una secuencia de aminoácidos codificada por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(f) una secuencia de aminoácidos codificada por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

98. Un método para diagnosticar la presencia de un tumor en un mamífero, el citado método que comprende determinar el nivel de expresión de un gen codificando una proteína teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con:

(a) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

(d) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) un polipéptido codificado por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78), en una muestra de prueba de células de tejido obtenida del citado mamífero y en una muestra de control de células normales del mismo tejido de origen que se sabe son normales, donde un mayor nivel de expresión de la citada proteína en la muestra de prueba, cuando se compara con la muestra de control, indica la presencia de un tumor en el mamífero del cual se obtuvo de la muestra de prueba.

99. El método de la Reivindicación 98, donde el paso de determinar el nivel de expresión de un gen codificando la citada proteína comprende emplear un oligonucleótido en una hibridación *in situ* o un análisis RT-PCR.

100. El método de la Reivindicación 98, donde el paso de determinar el nivel de expresión de un gen codificando la citada proteína comprende emplear un anticuerpo en un análisis de inmunohistoquímica o de inmunostransferencia.

101. El método de la Reivindicación 98, donde la citada proteína tiene:

(a) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(c) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su secuencia de péptido de señal asociado;

(d) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(e) una secuencia de aminoácidos codificada por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos.-1-78); o

(f) una secuencia de aminoácidos codificada por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos:1-78).

102. Un método para diagnosticar la presencia de un tumor en un mamífero, el citado método que comprende poner en contacto una muestra de prueba de células de tejido obtenida del citado mamífero con un anticuerpo, un oligopéptido o una molécula orgánica que se fija a una proteína teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con:

(a) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

(d) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) un polipéptido codificado por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos:1-78); o

(f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78), y detectar la formación de un complejo entre el citado anticuerpo, oligopéptido o molécula orgánica y la citada proteína en la muestra de prueba, donde la formación de un complejo indica la presencia de un tumor en el citado mamífero.

103. El método de la Reivindicación 102, donde el citado anticuerpo, oligopéptido o molécula orgánica se marca en forma detectable.

104. El método de la Reivindicación 102, donde la citada muestra de prueba de células de tejido se obtiene de un individuo que se sospecha tiene un tumor canceroso.

105. El método de la Reivindicación 102, donde la citada proteína tiene:

(a) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(c) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su secuencia de péptido de señal asociado;

(d) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(e) una secuencia de aminoácidos codificada por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos:1-78); o

(f) una secuencia de aminoácidos codificada por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

106. Un método para tratar o prevenir un desorden de proliferación celular asociado con la expresión o actividad aumentada de una proteína teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con:

(a) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

(d) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) un polipéptido codificado por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos:1-78), el citado método que comprende administrar a un sujeto que necesite el citado tratamiento una cantidad efectiva de un antagonista de la citada proteína, y en esta forma tratar efectivamente o prevenir el citado desorden de proliferación celular.

107. El método de la Reivindicación 106, donde el citado desorden de proliferación celular es cáncer.

108. El método de la Reivindicación 106, donde el citado antagonista es un anticuerpo polipéptido anti-TAT, un oligopéptido TAT de unión, una molécula orgánica TAT de unión o un oligonucleótido antisentido.

109. Un método para fijar un anticuerpo, un oligopéptido o una molécula orgánica a una célula que expresa una proteína teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con:

(a) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

(d) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) un polipéptido codificado por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78), el citado método que comprende poner en contacto la citada célula con un anticuerpo, un oligopéptido o una molécula orgánica que se fija a la citada proteína y permitir que suceda la unión del anticuerpo, del oligopéptido o de la molécula orgánica a la citada proteína, en esa forma fijando el citado anticuerpo, oligopéptido o molécula orgánica a la citada célula.

110. El método de la Reivindicación 109, donde el citado anticuerpo es un anticuerpo monoclonal.

111. El método de la Reivindicación 109, donde el citado anticuerpo es un fragmento de anticuerpo.

112. El método de la Reivindicación 109, donde el citado anticuerpo es un anticuerpo quimérico o humanizado.

113. El método de la Reivindicación 109, donde el citado anticuerpo, oligopéptido o molécula orgánica se conjuga a un agente inhibidor de crecimiento.

114. El método de la Reivindicación 109, donde el citado anticuerpo, oligopéptido o molécula orgánica se conjuga a un agente citotóxico.

115. El método de la Reivindicación 114, donde el citado agente citotóxico se selecciona entre el grupo conformado por toxinas, antibióticos, isótopos radioactivos y enzimas nucleolíticas.

116. El método de la Reivindicación 114, donde el agente citotóxico es una toxina.

117. El método de la Reivindicación 116, donde la toxina se selecciona entre el grupo conformado por maitansinoide y calicheamicina.

118. El método de la Reivindicación 116, donde la toxina es un maitansinoide.

119. El método de la Reivindicación 109, donde el citado anticuerpo se produce en bacterias.

120. El método de la Reivindicación 109, donde el citado anticuerpo se produce en células CHO.

121. El método de la Reivindicación 109, donde la citada célula es una célula cancerosa.

122. El método de la Reivindicación 121, donde la citada célula cancerosa se expone además a un tratamiento de radiación o a un agente quimioterapéutico.

123. El método de la Reivindicación 121, donde la citada célula cancerosa se selecciona entre el grupo conformado por una célula de cáncer de mama, una célula de cáncer colorrectal, una célula de cáncer de pulmón, una célula de cáncer ovárico, una célula de cáncer del sistema nervioso

central, una célula de cáncer de hígado, una célula de cáncer de vejiga, una célula de cáncer pancreático, una célula de cáncer cervical, una célula de melanoma y una célula de leucemia.

124. El método de la Reivindicación 123, donde la citada proteína se expresa en forma más abundante por la citada célula cancerosa cuando se compara con una célula normal del mismo tejido de origen.

125. El método de la Reivindicación 109 que causa la muerte de la citada célula.

126. El uso de un ácido nucleico como se reivindica en cualquiera de las Reivindicaciones 1 a 5 o 30 en la preparación de un medicamento para el tratamiento terapéutico o la detección diagnóstica de un cáncer.

127. El uso de un ácido nucleico como se reivindica en cualquiera de las Reivindicaciones 1 a 5 o 30 en la preparación de un medicamento para tratar un tumor.

128. El uso de un ácido nucleico como se reivindica en cualquiera de las Reivindicaciones 1 a 5 o 30 en la preparación de un medicamento para el tratamiento o la prevención de un desorden de proliferación celular.

129. El uso de un vector de expresión como se reivindica en cualquiera de las Reivindicaciones 6, 7 o 31 en la preparación de un medicamento para el tratamiento terapéutico o la detección diagnóstica de un cáncer.

130. El uso de un vector de expresión como se reivindica en cualquiera de las Reivindicaciones 6, 7 o 31 en la preparación de un medicamento para tratar un tumor.

131. El uso de un vector de expresión como se reivindica en cualquiera de las Reivindicaciones 6, 7 o 31 en la preparación de un medicamento para el tratamiento o la prevención de un desorden de proliferación celular.

132. El uso de una célula hospedera como se reivindica en cualquiera de las Reivindicaciones 8, 9, 32 o 33 en la preparación de un medicamento para el tratamiento terapéutico o la detección diagnóstica de un cáncer.

133. El uso de una célula hospedera como se reivindica en cualquiera de las Reivindicaciones 8, 9, 32 o 33 en la preparación de un medicamento para tratar un tumor.

134. El uso de una célula hospedera como se reivindica en cualquiera de las Reivindicaciones 8, 9, 32 o 33 en la preparación de un medicamento para el tratamiento o la prevención de un desorden de proliferación celular.

135. El uso de un polipéptido como se reivindica en cualquiera de las Reivindicaciones 11 a 14 en la preparación de un medicamento para el tratamiento terapéutico o la detección diagnóstica de un cáncer.

136. El uso de un polipéptido como se reivindica en cualquiera de las Reivindicaciones 11 a 14 en la preparación de un medicamento para tratar un tumor.

137. El uso de un polipéptido como se reivindica en cualquiera de las Reivindicaciones 11 a 14 en la preparación de un medicamento para el tratamiento o la prevención de un desorden de proliferación celular.

138. El uso de un anticuerpo como se reivindica en cualquiera de las Reivindicaciones 15 a 29 en la preparación de un medicamento para el tratamiento terapéutico o la detección diagnóstica de un cáncer.

139. El uso de un anticuerpo como se reivindica en cualquiera de las Reivindicaciones 15 a 29 en la preparación de un medicamento para tratar un tumor.

140. El uso de un anticuerpo como se reivindica en cualquiera de las Reivindicaciones 15 a 29 en la preparación de un medicamento para el tratamiento o la prevención de un desorden de proliferación celular.

141. El uso de un oligopéptido como se reivindica en cualquiera de las Reivindicaciones 35 a 44 en la preparación de un medicamento para el tratamiento terapéutico o la detección diagnóstica de un cáncer.

142. El uso de un oligopéptido como se reivindica en cualquiera de las Reivindicaciones 35 a 44 en la preparación de un medicamento para tratar un tumor.

143. El uso de un oligopéptido como se reivindica en cualquiera de las Reivindicaciones 35 a 44 en la preparación de un medicamento para el tratamiento o la prevención de un desorden de proliferación celular.

144. El uso de una molécula orgánica TAT de unión como se reivindica en cualquiera de las Reivindicaciones 45 a 54 en la preparación de un medicamento para el tratamiento terapéutico o la detección diagnóstica de un cáncer.

145. El uso de una molécula orgánica TAT de unión como se reivindica en cualquiera de las Reivindicaciones 45 a 54 en la preparación de un medicamento para tratar un tumor.

146. El uso de una molécula orgánica TAT de unión como se reivindica en cualquiera de las Reivindicaciones 45 a 54 en la preparación de un medicamento para el tratamiento o la prevención de un desorden de proliferación celular.

147. El uso de una composición de sustancias como se reivindica en cualquiera de las Reivindicaciones 55 o 56 en la preparación de un medicamento para el tratamiento terapéutico o la detección diagnóstica de un cáncer.

148. El uso de una composición de sustancias como se reivindica en cualquiera de las Reivindicaciones 55 o 56 en la preparación de un medicamento para tratar un tumor.

149. El uso de una composición de sustancias como se reivindica en cualquiera de las Reivindicaciones 55 o 56 en la preparación de un medicamento para el tratamiento o la prevención de un desorden de proliferación celular.

150. El uso de un artículo de manufactura como se reivindica en cualquiera de las Reivindicaciones 57 o 58 en la preparación de un medicamento para el tratamiento terapéutico o la detección diagnóstica de un cáncer.

151. El uso de un artículo de manufactura como se reivindica en cualquiera de las Reivindicaciones 57 o 58 en la preparación de un medicamento para tratar un tumor.

152. El uso de un artículo de manufactura como se reivindica en cualquiera de las Reivindicaciones 57 o 58 en la preparación de un medicamento para el tratamiento o la prevención de un desorden de proliferación celular.

153. Un método para inhibir el crecimiento de una célula, donde el crecimiento de la citada célula depende por lo menos en parte de un efecto potenciador del crecimiento de una proteína teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con:

(a) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

(d) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) un polipéptido codificado por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78), el citado método que comprende poner en contacto la citada proteína con un anticuerpo, oligopéptido o molécula orgánica que se fija a la citada proteína, en esta forma inhibiendo el crecimiento de la citada célula.

154. El método de la Reivindicación 153, donde la citada célula es una célula cancerosa.

155. El método de la Reivindicación 153, donde la citada proteína se expresa por la citada célula.

156. El método de la Reivindicación 153, donde la fijación del citado anticuerpo, oligopéptido o molécula orgánica a la citada proteína

antagoniza una actividad potenciadora del crecimiento celular de la citada proteína.

157. El método de la Reivindicación 153, donde la fijación del citado anticuerpo, oligopéptido o molécula orgánica a la citada proteína induce la muerte de la citada célula.

158. El método de la Reivindicación 153, donde el citado anticuerpo es un anticuerpo monoclonal.

159. El método de la Reivindicación 153, donde el citado anticuerpo es un fragmento de anticuerpo.

160. El método de la Reivindicación 153, donde el citado anticuerpo es un anticuerpo quimérico o humanizado.

161. El método de la Reivindicación 153, donde el citado anticuerpo, oligopéptido o molécula orgánica se conjuga a un agente inhibidor de crecimiento.

162. El método de la Reivindicación 153, donde el citado anticuerpo, oligopéptido o molécula orgánica se conjuga a un agente citotóxico.

163. El método de la Reivindicación 162, donde el citado agente citotóxico se selecciona entre el grupo conformado por toxinas, antibióticos, isótopos radioactivos y enzimas nucleolíticas.

164. El método de la Reivindicación 162, donde el agente citotóxico es una toxina.

165. El método de la Reivindicación 164, donde la toxina se selecciona entre el grupo conformado por maitansinoide y calicheamicina.

166. El método de la Reivindicación 164, donde la toxina es un maitansinoide.

167. El método de la Reivindicación 153, donde el citado anticuerpo se produce en bacterias.

168. El método de la Reivindicación 153, donde el citado anticuerpo se produce en células CHO.

169. El método de la Reivindicación 153, donde la citada proteína tiene:

(a) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(c) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su secuencia de péptido de señal asociado;

(d) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(e) una secuencia de aminoácidos codificada por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(f) una secuencia de aminoácidos codificada por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

170. Un método para tratar terapéuticamente un tumor en un mamífero, donde el crecimiento del citado tumor depende por lo menos en parte de un efecto potenciador del crecimiento de una proteína teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con:

(a) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

(d) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) un polipéptido codificado por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos:1-78), el citado método que comprende poner en contacto la citada proteína con un anticuerpo, oligopéptido o molécula orgánica que se fija a la citada proteína, en esta forma tratando efectivamente el citado tumor.

171. El método de la Reivindicación 170, donde la citada proteína se expresa por células del citado tumor.

172. El método de la Reivindicación 170, donde la fijación del citado anticuerpo, oligopéptido o molécula orgánica a la citada proteína antagoniza una actividad potenciadora del crecimiento celular de la citada proteína.

173. El método de la Reivindicación 170, donde el citado anticuerpo es un anticuerpo monoclonal.

174. El método de la Reivindicación 170, donde el citado anticuerpo es un fragmento de anticuerpo.

175. El método de la Reivindicación 170, donde el citado anticuerpo es un anticuerpo quimérico o humanizado.

176. El método de la Reivindicación 170, donde el citado anticuerpo, oligopéptido o molécula orgánica se conjuga a un agente inhibidor de crecimiento.

177. El método de la Reivindicación 170, donde el citado anticuerpo, oligopéptido o molécula orgánica se conjuga a un agente citotóxico.

178. El método de la Reivindicación 177, donde el citado agente citotóxico se selecciona entre el grupo conformado por toxinas, antibióticos, isótopos radioactivos y enzimas nucleolíticas.

179. El método de la Reivindicación 177, donde el agente citotóxico es una toxina.

180. El método de la Reivindicación 179, donde la toxina se selecciona entre el grupo conformado por maitansinoide y calicheamicina.

181. El método de la Reivindicación 179, donde la toxina es un maitansinoide.

182. El método de la Reivindicación 170, donde el citado anticuerpo se produce en bacterias.

183. El método de la Reivindicación 170, donde el citado anticuerpo se produce en células CHO.

184. El método de la Reivindicación 170, donde la citada proteína tiene:

(a) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(c) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su secuencia de péptido de señal asociado;

(d) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(e) una secuencia de aminoácidos codificada por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

(f) una secuencia de aminoácidos codificada por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

185. Un anticuerpo aislado que se fija a un polipéptido teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con:

(a) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(c) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

(d) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

(e) un polipéptido codificado por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos:1-78); o

(f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

186. Un anticuerpo aislado que se fija a un polipéptido teniendo:

(a) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

(b) una secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(c) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su secuencia de péptido de señal asociado;

(d) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

(e) una secuencia de aminoácidos codificada por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos:1-78); o

(f) una secuencia de aminoácidos codificada por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos:1-78).

522

187. El anticuerpo de la Reivindicación 185 el cual es un anticuerpo monoclonal.

188. El anticuerpo de la Reivindicación 185 el cual es un fragmento de anticuerpo.

189. El anticuerpo de la Reivindicación 185 el cual es un anticuerpo quimérico o humanizado.

190. El anticuerpo de la Reivindicación 185 el cual se conjuga con un agente inhibidor de crecimiento.

191. El anticuerpo de la Reivindicación 185 el cual se conjuga con un agente citotóxico.

192. El anticuerpo de la Reivindicación 191, donde el agente citotóxico se selecciona entre el grupo conformado por toxinas, antibióticos, isótopos radioactivos y enzimas nucleolíticas.

193. El anticuerpo de la Reivindicación 191, donde el agente citotóxico es una toxina.

194. El anticuerpo de la Reivindicación 193, donde la toxina se selecciona entre el grupo conformado por auristatina, maitansinoide y calicheamicina.

195. El anticuerpo de la Reivindicación 193, donde la toxina es un maitansinoide.

196. El anticuerpo de la Reivindicación 185 el cual se produce en bacterias.

197. El anticuerpo de la Reivindicación 185 el cual se produce en células CHO.

198. El anticuerpo de la Reivindicación 185 que induce la muerte de una célula a la cual se fija.

199. El anticuerpo de la Reivindicación 185 el cual se marca en forma detectable.

200. El anticuerpo de la Reivindicación 193, donde la toxina se selecciona entre el grupo conformado por MMAE (mono-metil auristatina E), MMAF, (auristatina E valeril bencihidrazona) y AFP (Auristatina F fenileno diamina).

201. El anticuerpo de la Reivindicación 193, donde la toxina se adhiere en forma covalente al anticuerpo mediante un enlazante.

202. El anticuerpo de la Reivindicación 201, donde el enlazante se selecciona entre el grupo conformado por maleimidocaproilo (MC), valina-citrulina (val-cit, vc), citrulina (2-amino-5-ureido ácido pentanoico), PAB (p-aminobencilcarbamoilo), Me (N-metil-valina citrulina), MC(PEG)6-OH (maleimidocaproilo-polietilenglicol), SPP (N-Succinimidil 4-(2-piridiltio)pentanoato), SMCC (N-Succinimidil 4-(N-maleimidometil) ciclohexano-1 carboxilato) y MC-vc-PAB.

203. El anticuerpo de la Reivindicación 200, donde la toxina es MMAE.

204. El anticuerpo de la Reivindicación 200, donde la toxina es MMAF.

205. El anticuerpo de la Reivindicación 202, donde el enlazante es MC-vc-PAB y la toxina es MMAE o MMAF.

206. El anticuerpo de una cualquiera de las Reivindicaciones 185-205, donde el anticuerpo fija el polipéptido TAT188.

207. El anticuerpo de la Reivindicación 206, donde el anticuerpo inhibe la proliferación o promueve la muerte celular de una célula que expresa el TAT188.

208. El anticuerpo de la Reivindicación 207, donde la célula es una célula cancerosa.

209. El anticuerpo de la Reivindicación 208, donde la célula cancerosa se selecciona entre el grupo de mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado.

210. Un anticuerpo aislado que compite con la fijación de los epítopos del polipéptido TAT188 unido por un anticuerpo producido por el hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No. PTA-6193), 12B9.1 (ATCC Acceso No. PTA-6194) y 12G12.1 (ATCC Acceso No. PTA-6195).

211. Un anticuerpo aislado teniendo la actividad biológica de un anticuerpo producido por el hibridoma seleccionado entre el grupo

conformado por 3B5.1 (ATCC Acceso No. PTA-6193), 12B9.1 (ATCC Acceso No. PTA-6194) y 12G12.1 (ATCC Acceso No. PTA- 6195), donde la actividad biológica es la inhibición de la proliferación celular o la promoción de la muerte celular en una célula que expresa el TAT188.

212. Un anticuerpo aislado que comprende en una región de determinante complementario (CDR) correspondiente, una secuencia de aminoácidos teniendo por lo menos 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de una secuencia de aminoácidos de por lo menos 1, 2, 3, 4, 5 o 6 de los CDR(s) del anticuerpo producido por un hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No. PTA-6193), 12B9.1 (ATCC Acceso No. PTA-6194) y 12G12.1 (ATCC Acceso No. PTA-6195).

213. El anticuerpo de una cualquiera de las Reivindicaciones 185-205, donde el anticuerpo comprende en una región de determinante complementario (CDR) correspondiente, una secuencia de aminoácidos teniendo por lo menos 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de una secuencia de aminoácidos de por lo menos 1, 2, 3, 4, 5 o 6 de los CDR(s) del anticuerpo producido por un hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No. PTA- 6193), 12B9.1 (ATCC Acceso No. PTA-6194) y 12G12.1 (ATCC Acceso No. PTA-6195).

214. El anticuerpo de una cualquiera de las Reivindicaciones 185-205, donde el anticuerpo exhibe la actividad biológica de un anticuerpo producido por el hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No. PTA-6193), 12B9.1 (ATCC Acceso No. PTA-6194) y 12G12.1 (ATCC Acceso No. PTA-6195), donde la actividad biológica es la inhibición de la proliferación celular o la promoción de la muerte celular en una célula que expresa el TAT188.

215. El anticuerpo de la Reivindicación 214, donde la célula es una célula cancerosa.

216. El anticuerpo de la Reivindicación 215, donde la célula cancerosa se selecciona entre el grupo conformado por mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado.

217. Un método para inhibir el crecimiento de una célula que expresa el TAT188, el método que comprende poner una célula en contacto con un anticuerpo de una cualquiera de las Reivindicaciones 185-205.

218. El método de la Reivindicación 217, donde la célula es una célula cancerosa.

219. El método de la Reivindicación 218, donde la célula cancerosa se selecciona entre el grupo conformado por mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado.

220. El método de la Reivindicación 219, donde la célula cancerosa es una célula de mamífero.

221. El método de la Reivindicación 220, donde la célula de mamífero es una célula humana.

222. Un método para inhibir el crecimiento de una célula que expresa el TAT188, el método que comprende poner una célula en contacto con un anticuerpo de una cualquiera de las Reivindicaciones 185-205, donde el anticuerpo comprende en una región de determinante complementario (CDR) correspondiente, una secuencia de aminoácidos teniendo por lo menos 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de una secuencia de aminoácidos de por lo menos 1, 2, 3, 4, 5 o 6 de las CDR(s) del anticuerpo producido por un hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No. PTA-6193), 12B9.1 (ATCC Acceso No. PTA-6194) y 12G12.1 (ATCC Acceso No. PTA-6195).

223. Un método para detectar el nivel del polipéptido TAT188 expresado en una célula de prueba en relación con una célula de control, el método que comprende:

(a) poner en contacto la célula de prueba y la célula de control con un anticuerpo anti-TAT188 aislado de la Reivindicación 210;

(b) detectar la fijación del anticuerpo; y

526

(c) determinar la fijación relativa del anticuerpo a la célula de prueba y de control.

224. El método de la Reivindicación 223, donde la célula de prueba y la célula de control se someten a lisis.

225. El método de la Reivindicación 223, donde la célula de prueba está en un tejido.

226. El método de la Reivindicación 225, donde el tejido es un tejido de tumor.

227. El método de la Reivindicación 226, donde el tejido de tumor se selecciona entre el grupo conformado por mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado.

228. Un método para detectar el nivel del polipéptido TAT188 o un polipéptido teniendo una identidad de la secuencia por lo menos de 80% con una secuencia de aminoácidos que se muestra en la Figura 115 (SEQ ID No: 115) en una célula de prueba en relación con una célula de control, el método que comprende:

(a) poner en contacto la célula de prueba y la célula de control con un anticuerpo aislado de una cualquiera de las Reivindicaciones 185-189, 196-199 y 210-212;

(b) detectar la fijación del anticuerpo; y

(c) determinar la fijación relativa del anticuerpo a la célula de prueba y de control.

229. El método de la Reivindicación 228, donde el nivel del polipéptido TAT188 en la célula de prueba es mayor que el de la célula de control.

230. El método de la Reivindicación 229, donde el método diagnostica el cáncer en un tejido que contiene o que tiene contenida la célula de prueba.

231. Un ARN de interferencia de unión TAT (ARNsi) que se fija a un ácido nucleico teniendo una identidad de la secuencia por lo menos de 80% con:

(a) una secuencia nucleótida que se muestra en la Figura 37 (SEQ ID No:37; y

(b) el complemento de (a), donde el ARNsi reduce la expresión del TAT188.

232. Un vector de expresión que comprende el ARNsi de la Reivindicación 231.

233. El vector de expresión de la Reivindicación 232, donde el citado ARNsi está enlazado operacionalmente a las secuencias de control reconocidas por una célula hospedera transinfectada con el vector.

234. Una célula hospedera que comprende el vector de expresión de la Reivindicación 233.

235. Una composición de sustancias que comprende:

(a) el anticuerpo de la Reivindicación 185, o

(b) el ARNsi de la Reivindicación 231, en combinación con un excipiente.

236. La composición de las sustancias de la Reivindicación 235, donde el citado excipiente es un excipiente farmacéuticamente aceptable.

237. Un artículo de manufactura:

(a) un envase; y

(b) la composición de las sustancias de la Reivindicación 235 contenida en el citado envase.

238. El artículo de manufactura de la Reivindicación 237, que además comprende una etiqueta adherida al citado envase, o un inserto en el paquete incluido con el citado envase, haciendo referencia al uso de la citada composición de sustancias para el tratamiento terapéutico o la detección diagnóstica de un cáncer.

239. Un método para inhibir el crecimiento de una célula cancerosa que expresa un polipéptido teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con una secuencia de aminoácidos que se muestra en la Figura 115 (SEQ ID No: 115), el citado método que comprende poner en contacto la citada célula cancerosa con un ARNsi que se fija a un ácido nucleico codificando el aminoácido en la citada célula

528

cancerosa, en esta forma inhibiendo el crecimiento de la citada célula cancerosa.

240. El método de la Reivindicación 239, donde el ácido nucleico tiene la secuencia que se muestra en la Figura 37 (SEQ ID No:37).

241. El método de la Reivindicación 228, donde la detección del nivel de expresión del polipéptido comprende emplear un anticuerpo en un análisis de inmunohistoquímica.

242. Un método para tratar o prevenir un desorden de proliferación celular asociado con una expresión o actividad aumentada de un polipéptido teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con una secuencia de aminoácidos que se muestra en la Figura 115 (SEQ ID No: 115), el citado método que comprende administrar a un sujeto que necesite el citado tratamiento una cantidad efectiva de un antagonista de un polipéptido TAT188.

243. El método de la Reivindicación 242, donde el citado antagonista es un anticuerpo aislado de un polipéptido anti-TAT188 de una cualquiera de las Reivindicaciones 1-21 y 23-28.

244. El método de la Reivindicación 243, donde el desorden de proliferación celular es cáncer.

245. El método de la Reivindicación 244, donde el cáncer se selecciona entre el grupo conformado por mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado.

Aún otras incorporaciones de la presente invención serán evidentes para un artesano con destreza cuando lea la presente especificación.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

La Figura 1 muestra una secuencia nucleótida (SEQ ID No:1) de un TAT161 ADNc, donde la SEQ ID No:1 es un clon denominado en la presente "ADN77507".

Las Figuras 2A-B muestran una secuencia nucleótida (SEQ ID No:2) de un TAT101 ADNc, donde la SEQ ID No:2 es un clon denominado en la presente "ADN80894".

529

La Figura 3 muestra una secuencia nucleótida (SEQ ID No:3) de un TAT157 ADNc, donde la SEQ ID No:3 es un clon denominado en la presente "ADN82343".

La Figura 4 muestra una secuencia nucleótida (SEQ ID No:4) de un TAT160 ADNc, donde la SEQ ID No:4 es un clon denominado en la presente "ADN87994".

La Figura 5 muestra una secuencia nucleótida (SEQ ID No:5) de un TAT158 ADNc, donde la SEQ ID No:5 es un clon denominado en la presente "ADN88131".

La Figura 6 muestra una secuencia nucleótida (SEQ ID No:6) de un TAT110 ADNc, donde la SEQ ID No:6 es un clon denominado en la presente "ADN95930".

La Figura 7 muestra una secuencia nucleótida (SEQ ID No:7) de un TAT210 ADNc, donde la SEQ ID No:7 es un clon denominado en la presente "ADN95930-1".

La Figura 8 muestra una secuencia nucleótida (SEQ ID No:8) de un TAT159 ADNc, donde la SEQ ID No:8 es un clon denominado en la presente "ADN96917".

La Figura 9 muestra una secuencia nucleótida (SEQ ID No:9) de un TAT112 ADNc, donde la SEQ ID No:9 es un clon denominado en la presente "ADN96930".

La Figura 10 muestra una secuencia nucleótida (SEQ ID No: 10) de un TAT147 ADNc, donde la SEQ ID No: 10 es un clon denominado en la presente "ADN96936".

La Figura 11 muestra una secuencia nucleótida (SEQ ID No: 11) de un TAT145 ADNc, donde la SEQ ID No:11 es un clon denominado en la presente "ADN98565".

La Figura 12 muestra una secuencia nucleótida (SEQ ID No:12) de un TAT152 ADNc, donde la SEQ ID No:12 es un clon denominado en la presente "ADN246435".

La Figura 13 muestra una secuencia nucleótida (SEQ ID No:13) de un TAT162 ADNc, donde la SEQ ID No:13 es un clon denominado en la presente "ADN98591".

La Figura 14 muestra una secuencia nucleótida (SEQ ID No: 14) de un TAT114 ADNc, donde la SEQ ID No:14 es un clon denominado en la presente "ADN108809".

La Figura 15 muestra una secuencia nucleótida (SEQ ID No:15) de un TAT119 ADNc, donde la SEQ ID No: 15 es un clon denominado en la presente "ADN119488".

La Figura 16 muestra una secuencia nucleótida (SEQ ID No: 16) de un TAT103 ADNc, donde la SEQ ID No:16 es un clon denominado en la presente "ADN143493".

Las Figuras 17A-B muestran una secuencia nucleótida (SEQ ID No: 17) de un TAT130 ADNc, donde la SEQ ID No: 17 es un clon denominado en la presente "ADN167234".

La Figura 18 muestra una secuencia nucleótida (SEQ ID No: 18) de un TAT 166 ADNc, donde la SEQ ID No:18 es un clon denominado en la presente "ADN235621".

La Figura 19 muestra una secuencia nucleótida (SEQ ID No: 19) de un TAT132 ADNc, donde la SEQ ID No:19 es un clon denominado en la presente "ADN176766".

La Figura 20 muestra una secuencia nucleótida (SEQ ID No:20) de un TAT150 ADNc, donde la SEQ ID No:20 es un clon denominado en la presente "ADN236463".

La Figura 21 muestra una secuencia nucleótida (SEQ ID No:21) de un TAT129 ADNc, donde la SEQ ID No:21 es un clon denominado en la presente "ADN181162".

La Figura 22 muestra una secuencia nucleótida (SEQ ID No:22) de un TAT111 ADNc, donde la SEQ ID No:22 es un clon denominado en la presente "ADN188221".

La Figura 23 muestra una secuencia nucleótida (SEQ ID No:23) de un TAT146 ADNc, donde la SEQ ID No:23 es un clon denominado en la presente "ADN233876".

La Figura 24 muestra una secuencia nucleótida (SEQ ID No:24) de un TAT148 ADNc, donde la SEQ ID No:24 es un clon denominado en la presente "ADN193891".

La Figura 25 muestra una secuencia nucleótida (SEQ ID No:25) de un TAT187 ADNc, donde la SEQ ID No:25 es un clon denominado en la presente "ADN248170".

La Figura 26 muestra una secuencia nucleótida (SEQ ID No:26) de un TAT118 ADNc, donde la SEQ ID No:26 es un clon denominado en la presente "ADN194628".

La Figura 27 muestra una secuencia nucleótida (SEQ ID No:27) de un TAT167 ADNc, donde la SEQ ID No:27 es un clon denominado en la presente "ADN246415".

La Figura 28 muestra una secuencia nucleótida (SEQ ID No:28) de un TAT123 ADNc, donde la SEQ ID No:28 es un clon denominado en la presente "ADN210499".

La Figura 29 muestra una secuencia nucleótida (SEQ ID No:29) de un TAT211 ADNc, donde la SEQ ID No:29 es un clon denominado en la presente "ADN219894".

La Figura 30 muestra una secuencia nucleótida (SEQ ID No:30) de un TAT113 ADNc, donde la SEQ ID No:30 es un clon denominado en la presente "ADN215609".

La Figura 31 muestra una secuencia nucleótida (SEQ ID No:31) de un TAT128 ADNc, donde la SEQ ID No:31 es un clon denominado en la presente "ADN220432".

Las Figuras 32A-B muestran una secuencia nucleótida (SEQ ID No:32) de un TAT164 ADNc, donde la SEQ ID No:32 es un clon denominado en la presente "ADN226094".

La Figura 33 muestra una secuencia nucleótida (SEQ ID No:33) de un TAT122 ADNc, donde la SEQ ID No:33 es un clon denominado en la presente "ADN226165".

La Figura 34 muestra una secuencia nucleótida (SEQ ID No:34) de un TAT117 ADNc, donde la SEQ ID No:34 es un clon denominado en la presente "ADN226237".

La Figura 35 muestra una secuencia nucleótida (SEQ ID No:35) de un TAT168 ADNc, donde la SEQ ID No:35 es un clon denominado en la presente "ADN246450".

La Figura 36 muestra una secuencia nucleótida (SEQ ID No:36) de un TAT144 ADNc, donde la SEQ ID No:36 es un clon denominado en la presente "ADN226456".

La Figura 37 muestra una secuencia nucleótida (SEQ ID No:37) de un TAT188 ADNc, donde la SEQ ID No:37 es un clon denominado en la presente "ADN237637".

La Figura 38 muestra una secuencia nucleótida (SEQ ID No:38) de un TAT126 ADNc, donde la SEQ ID No:38 es un clon denominado en la presente "ADN226539".

La Figura 39 muestra una secuencia nucleótida (SEQ ID No:39) de un TAT151 ADNc, donde la SEQ ID No:39 es un clon denominado en la presente "ADN236511".

La Figura 40 muestra una secuencia nucleótida (SEQ ID No:40) de un TAT115 ADNc, donde la SEQ ID No:40 es un clon denominado en la presente "ADN226771".

La Figura 41 muestra una secuencia nucleótida (SEQ ID No:41) de un TAT163 ADNc, donde la SEQ ID No:41 es un clon denominado en la presente "ADN227087".

La Figura 42 muestra una secuencia nucleótida (SEQ ID No:42) de un TAT227 ADNc, donde la SEQ ID No:42 es un clon denominado en la presente "ADN266307".

La Figura 43 muestra una secuencia nucleótida (SEQ ID No:43) de un TAT228 ADNc, donde la SEQ ID No:43 es un clon denominado en la presente "ADN266311".

La Figura 44 muestra una secuencia nucleótida (SEQ ID No:44) de un TAT229 ADNc, donde la SEQ ID No:44 es un clon denominado en la presente "ADN266312".

La Figura 45 muestra una secuencia nucleótida (SEQ ID No:45) de un TAT230 ADNc, donde la SEQ ID No:45 es un clon denominado en la presente "ADN266313".

La Figura 46 muestra una secuencia nucleótida (SEQ ID No:46) de un TAT121 ADNc, donde la SEQ ID No:46 es un clon denominado en la presente "ADN227224".

La Figura 47 muestra una secuencia nucleótida (SEQ ID No:47) de un TAT183 ADNc, donde la SEQ ID No:47 es un clon denominado en la presente "ADN247486".

La Figura 48 muestra una secuencia nucleótida (SEQ ID No:48) de un TAT165 ADNc, donde la SEQ ID No:48 es un clon denominado en la presente "ADN227578".

La Figura 49 muestra una secuencia nucleótida (SEQ ID No:49) de un TAT131 ADNc, donde la SEQ ID No:49 es un clon denominado en la presente "ADN227800".

La Figura 50 muestra una secuencia nucleótida (SEQ ID No:50) de un TAT140 ADNc, donde la SEQ ID No:50 es un clon denominado en la presente "ADN227904".

La Figura 51 muestra una secuencia nucleótida (SEQ ID No:51) de un TAT127 ADNc, donde la SEQ ID No:51 es un clon denominado en la presente "ADN228199".

La Figura 52 muestra una secuencia nucleótida (SEQ ID No:52) de un TAT116 ADNc, donde la SEQ ID No:52 es un clon denominado en la presente "ADN228201".

La Figura 53 muestra una secuencia nucleótida (SEQ ID No:53) de un TAT189 ADNc, donde la SEQ ID No:53 es un clon denominado en la presente "ADN247488".

La Figura 54 muestra una secuencia nucleótida (SEQ ID No:54) de un TAT190 ADNc, donde la SEQ ID No:54 es un clon denominado en la presente "ADN236538".

La Figura 55 muestra una secuencia nucleótida (SEQ ID No:55) de un TAT191 ADNc, donde la SEQ ID No:55 es un clon denominado en la presente "ADN247489".

La Figura 56 muestra una secuencia nucleótida (SEQ ID No:56) de un TAT133 ADNc, donde la SEQ ID No:56 es un clon denominado en la presente "ADN228211".

La Figura 57 muestra una secuencia nucleótida (SEQ ID No:57) de un TAT186 ADNc, donde la SEQ ID No:57 es un clon denominado en la presente "ADN233937".

La Figura 58 muestra una secuencia nucleótida (SEQ ID No:58) de un TAT120 ADNc, donde la SEQ ID No:58 es un clon denominado en la presente "ADN228993".

La Figura 59 muestra una secuencia nucleótida (SEQ ID No:59) de un TAT124 ADNc, donde la SEQ ID No:59 es un clon denominado en la presente "ADN228994".

La Figura 60 muestra una secuencia nucleótida (SEQ ID No:60) de un TAT105 ADNc, donde la SEQ ID No:60 es un clon denominado en la presente "ADN229410".

Las Figuras 61A-B muestran una secuencia nucleótida (SEQ ID No:61) de un TAT107 ADNc, donde la SEQ ID No:61 es un clon denominado en la presente "ADN229411".

Las Figuras 62A-B muestran una secuencia nucleótida (SEQ ID No:62) de un TAT108 ADNc, donde la SEQ ID No:62 es un clon denominado en la presente "ADN229413".

Las Figuras 63A-B muestran una secuencia nucleótida (SEQ ID No:63) de un TAT139 ADNc, donde la SEQ ID No:63 es un clon denominado en la presente "ADN229700".

La Figura 64 muestra una secuencia nucleótida (SEQ ID No:64) de un TAT143 ADNc, donde la SEQ ID No:64 es un clon denominado en la presente "ADN231312".

La Figura 65 muestra una secuencia nucleótida (SEQ ID No:65) de un TAT100 ADNc, donde la SEQ ID No:65 es un clon denominado en la presente "ADN231542".

La Figura 66 muestra una secuencia nucleótida (SEQ ID No:66) de un TAT284 ADNc, donde la SEQ ID No:66 es un clon denominado en la presente "ADN231542-1".

La Figura 67 muestra una secuencia nucleótida (SEQ ID No:67) de un TAT285 ADNc, donde la SEQ ID No:67 es un clon denominado en la presente "ADN231542-2".

La Figura 68 muestra una secuencia nucleótida (SEQ ID No:68) de un TAT285-1 ADNc, donde la SEQ ID No:68 es un clon denominado en la presente "ADN297393".

La Figura 69 muestra una secuencia nucleótida (SEQ ID No:69) de un TAT125 ADNc, donde la SEQ ID No:69 es un clon denominado en la presente "ADN232754".

La Figura 70 muestra una secuencia nucleótida (SEQ ID No:70) de un TAT149 ADNc, donde la SEQ ID No:70 es un clon denominado en la presente "ADN234833".

La Figura 71 muestra una secuencia nucleótida (SEQ ID No:71) de un TAT231 ADNc, donde la SEQ ID No:71 es un clon denominado en la presente "ADN268022".

La Figura 72 muestra una secuencia nucleótida (SEQ ID No:72) de un TAT153 ADNc, donde la SEQ ID No:72 es un clon denominado en la presente "ADN236246".

La Figura 73 muestra una secuencia nucleótida (SEQ ID No:73) de un TAT104 ADNc, donde la SEQ ID No:73 es un clon denominado en la presente "ADN236343".

La Figura 74 muestra una secuencia nucleótida (SEQ ID No:74) de un TAT141 ADNc, donde la SEQ ID No:74 es un clon denominado en la presente "ADN236493".

La Figura 75 muestra una secuencia nucleótida (SEQ ID No:75) de un TAT102 ADNc, donde la SEQ ID No:75 es un clon denominado en la presente "ADN236534".

La Figura 76 muestra una secuencia nucleótida (SEQ ID No:76) de un TAT109 ADNc, donde la SEQ ID No:76 es un clon denominado en la presente "ADN246430".

La Figura 77 muestra una secuencia nucleótida (SEQ ID No:77) de un TAT142 ADNc, donde la SEQ ID No:77 es un clon denominado en la presente "ADN247480".

Las Figuras 78A-B muestran una secuencia nucleótida (SEQ ID No:78) de un TAT106 ADNc, donde la SEQ ID No:78 es un clon denominado en la presente "ADN264454".

La Figura 79 muestra una secuencia de aminoácidos (SEQ ID No:79) derivada de la secuencia de codificación de SEQ ID No:1 que se muestra en la Figura 1.

La Figura 80 muestra una secuencia de aminoácidos (SEQ ID No:80) derivada de la secuencia de codificación de SEQ ID No:2 que se muestra en la Figura 2.

La Figura 81 muestra una secuencia de aminoácidos (SEQ ID No:81) derivada de la secuencia de codificación de SEQ ID No:3 que se muestra en la Figura 3.

La Figura 82 muestra una secuencia de aminoácidos (SEQ ID No:82) derivada de la secuencia de codificación de SEQ ID No:4 que se muestra en la Figura 4.

La Figura 83 muestra una secuencia de aminoácidos (SEQ ID No:83) derivada de la secuencia de codificación de SEQ ID No:5 que se muestra en la Figura 5.

La Figura 84 muestra una secuencia de aminoácidos (SEQ ID No:84) derivada de la secuencia de codificación de SEQ ID No:6 que se muestra en la Figura 6.

La Figura 85 muestra una secuencia de aminoácidos (SEQ ID No:85) derivada de la secuencia de codificación de SEQ ID No:7 que se muestra en la Figura 7.

La Figura 86 muestra una secuencia de aminoácidos (SEQ ID No: 86) derivada de la secuencia de codificación de SEQ ID No:8 que se muestra en la Figura 8.

La Figura 87 muestra una secuencia de aminoácidos (SEQ ID No:87) derivada de la secuencia de codificación de SEQ ID No:9 que se muestra en la Figura 9.

La Figura 88 muestra una secuencia de aminoácidos (SEQ ID No:88) derivada de la secuencia de codificación de SEQ ID No: 10 que se muestra en la Figura 10.

La Figura 89 muestra una secuencia de aminoácidos (SEQ ID No:89) derivada de la secuencia de codificación de SEQ ID No: 11 que se muestra en la Figura 11.

La Figura 90 muestra una secuencia de aminoácidos (SEQ ID No:90) derivada de la secuencia de codificación de SEQ ID No: 12 que se muestra en la Figura 12.

La Figura 91 muestra una secuencia de aminoácidos (SEQ ID No:91) derivada de la secuencia de codificación de SEQ ID No:13 que se muestra en la Figura 13.

La Figura 92 muestra una secuencia de aminoácidos (SEQ ID No:92) derivada de la secuencia de codificación de SEQ ID No: 14 que se muestra en la Figura 14.

La Figura 93 muestra una secuencia de aminoácidos (SEQ ID N :93) derivada de la secuencia de codificación de SEQ ID No: 15 que se muestra en la Figura 15.

La Figura 94 muestra una secuencia de aminoácidos (SEQ ID No:94) derivada de la secuencia de codificación de SEQ ID No: 16 que se muestra en la Figura 16.

La Figura 95 muestra una secuencia de aminoácidos (SEQ ID No:95) derivada de la secuencia de codificación de SEQ ID No: 17que se muestra en las Figuras 17 A-B.

La Figura 96 muestra una secuencia de aminoácidos (SEQ ID No:96) derivada de la secuencia de codificación de SEQ ID No:18 que se muestra en la Figura 18.

La Figura 97 muestra una secuencia de aminoácidos (SEQ ID No: 97) derivada de la secuencia de codificación de SEQ ID No:19 que se muestra en la Figura 19.

La Figura 98 muestra una secuencia de aminoácidos (SEQ ID No:98) derivada de la secuencia de codificación de SEQ ID No:20 que se muestra en la Figura 20.

La Figura 99 muestra una secuencia de aminoácidos (SEQ ID No:99) derivada de la secuencia de codificación de SEQ ID No:21 que se muestra en la Figura 21.

La Figura 100 muestra una secuencia de aminoácidos (SEQ ID No: 100) derivada de la secuencia de codificación de SEQ ID No:22 que se muestra en la Figura 22.

La Figura 101 muestra una secuencia de aminoácidos (SEQ ID No: 101) derivada de la secuencia de codificación de SEQ ID No:23 que se muestra en la Figura 23.

La Figura 102 muestra una secuencia de aminoácidos (SEQ ID No: 102) derivada de la secuencia de codificación de SEQ ID No:24 que se muestra en la Figura 24.

La Figura 103 muestra una secuencia de aminoácidos (SEQ ID No: 103) derivada de la secuencia de codificación de SEQ ID No:25 que se muestra en la Figura 25.

La Figura 104 muestra una secuencia de aminoácidos (SEQ ID No: 104) derivada de la secuencia de codificación de SEQ ID No:26 que se muestra en la Figura 26.

La Figura 105 muestra una secuencia de aminoácidos (SEQ ID No: 105) derivada de la secuencia de codificación de SEQ ID No:27 que se muestra en la Figura 27.

La Figura 106 muestra una secuencia de aminoácidos (SEQ ID No: 106) derivada de la secuencia de codificación de SEQ ID No:28 que se muestra en la Figura 28.

La Figura 107 muestra una secuencia de aminoácidos (SEQ ID No: 107) derivada de la secuencia de codificación de SEQ ID No:29 que se muestra en la Figura 29.

La Figura 108 muestra una secuencia de aminoácidos (SEQ ID No: 108) derivada de la secuencia de codificación de SEQ ID No:30 que se muestra en la Figura 30.

La Figura 109 muestra una secuencia de aminoácidos (SEQ ID No: 109) derivada de la secuencia de codificación de SEQ ID No:31 que se muestra en la Figura 31.

Las Figuras 110A-B muestran una secuencia de aminoácidos (SEQ ID No: 110) derivada de la secuencia de codificación de SEQ ID No:32 que se muestra en las Figuras 32A-B.

La Figura 111 muestra una secuencia de aminoácidos (SEQ ID No:111) derivada de la secuencia de codificación de SEQ ID No:33 que se muestra en la Figura 33.

La Figura 112 muestra una secuencia de aminoácidos (SEQ ID No: 112) derivada de la secuencia de codificación de SEQ ID No:34 que se muestra en la Figura 34.

La Figura 113 muestra una secuencia de aminoácidos (SEQ ID No: 113) derivada de la secuencia de codificación de SEQ ID No:35 que se muestra en la Figura 35.

La Figura 114 muestra una secuencia de aminoácidos (SEQ ID No: 114) derivada de la secuencia de codificación de SEQ ID No:36 que se muestra en la Figura 36.

La Figura 115 muestra una secuencia de aminoácidos (SEQ ID No: 115) derivada de la secuencia de codificación de SEQ ID No: 37 que se muestra en la Figura 37.

La Figura 116 muestra una secuencia de aminoácidos (SEQ ID No: 116) derivada de la secuencia de codificación de SEQ ID No:38 que se muestra en la Figura 38.

La Figura 117 muestra una secuencia de aminoácidos (SEQ ID No: 117) derivada de la secuencia de codificación de SEQ ID No:39 que se muestra en la Figura 39.

La Figura 118 muestra una secuencia de aminoácidos (SEQ ID No: 118) derivada de la secuencia de codificación de SEQ ID No:40 que se muestra en la Figura 40.

La Figura 119 muestra una secuencia de aminoácidos (SEQ ID No: 119) derivada de la secuencia de codificación de SEQ ID No:41 que se muestra en la Figura 41.

La Figura 120 muestra una secuencia de aminoácidos (SEQ ID No: 120) derivada de la secuencia de codificación de SEQ ID No:42 que se muestra en la Figura 42.

La Figura 121 muestra una secuencia de aminoácidos (SEQ ID No: 121) derivada de la secuencia de codificación de SEQ ID No:43 que se muestra en la Figura 43.

La Figura 122 muestra una secuencia de aminoácidos (SEQ ID No: 122) derivada de la secuencia de codificación de SEQ ID No:44 que se muestra en la Figura 44.

La Figura 123 muestra una secuencia de aminoácidos (SEQ ID No: 123) derivada de la secuencia de codificación de SEQ ID No:45 que se muestra en la Figura 45.

La Figura 124 muestra una secuencia de aminoácidos (SEQ ID No: 124) derivada de la secuencia de codificación de SEQ ID No:46 que se muestra en la Figura 46.

La Figura 125 muestra una secuencia de aminoácidos (SEQ ID No: 125) derivada de la secuencia de codificación de SEQ ID No:47 que se muestra en la Figura 47.

La Figura 126 muestra una secuencia de aminoácidos (SEQ ID No: 126) derivada de la secuencia de codificación de SEQ ID No:48 que se muestra en la Figura 48.

La Figura 127 muestra una secuencia de aminoácidos (SEQ ID No:127) derivada de la secuencia de codificación de SEQ ID No:49 que se muestra en la Figura 49.

La Figura 128 muestra una secuencia de aminoácidos (SEQ ID No:128) derivada de la secuencia de codificación de SEQ ID No:50 que se muestra en la Figura 50.

La Figura 129 muestra una secuencia de aminoácidos (SEQ ID No: 129) derivada de la secuencia de codificación de SEQ ID No:51 que se muestra en la Figura 51.

La Figura 130 muestra una secuencia de aminoácidos (SEQ ID No: 130) derivada de la secuencia de codificación de SEQ ID No:52 que se muestra en la Figura 52.

La Figura 131 muestra una secuencia de aminoácidos (SEQ ID No:131) derivada de la secuencia de codificación de SEQ ID No:53 que se muestra en la Figura 53.

La Figura 132 muestra una secuencia de aminoácidos (SEQ ID No: 132) derivada de la secuencia de codificación de SEQ ID No:54 que se muestra en la Figura 54.

La Figura 133 muestra una secuencia de aminoácidos (SEQ ID No: 133) derivada de la secuencia de codificación de SEQ ID No:55 que se muestra en la Figura 55.

La Figura 134 muestra una secuencia de aminoácidos (SEQ ID No: 134) derivada de la secuencia de codificación de SEQ ID No:56 que se muestra en la Figura 56.

La Figura 135 muestra una secuencia de aminoácidos (SEQ ID No: 135) derivada de la secuencia de codificación de SEQ ID No:57 que se muestra en la Figura 57.

La Figura 136 muestra una secuencia de aminoácidos (SEQ ID No:136) derivada de la secuencia de codificación de SEQ ID No:58 que se muestra en la Figura 58.

La Figura 137 muestra una secuencia de aminoácidos (SEQ ID No: 137) derivada de la secuencia de codificación de SEQ ID No:59 que se muestra en la Figura 59.

La Figura 138 muestra una secuencia de aminoácidos (SEQ ID No:138) derivada de la secuencia de codificación de SEQ ID No:60 que se muestra en la Figura 60.

La Figura 139 muestra una secuencia de aminoácidos (SEQ ID No: 139) derivada de la secuencia de codificación de SEQ ID No:61 que se muestra en las Figuras 61A-B.

La Figura 140 muestra una secuencia de aminoácidos (SEQ ID No: 140) derivada de la secuencia de codificación de SEQ ID No:62 que se muestra en las Figuras 62A-B.

La Figura 141 muestra una secuencia de aminoácidos (SEQ ID No:141) derivada de la secuencia de codificación de SEQ ID No:63 que se muestra en las Figuras 63 A-B.

La Figura 142 muestra una secuencia de aminoácidos (SEQ ID No: 142) derivada de la secuencia de codificación de SEQ ID No:64 que se muestra en la Figura 64.

La Figura 143 muestra una secuencia de aminoácidos (SEQ ID No: 143) derivada de la secuencia de codificación de SEQ ID No:66 que se muestra en la Figura 66.

La Figura 144 muestra una secuencia de aminoácidos (SEQ ID No: 144) derivada de la secuencia de codificación de SEQ ID No:67 que se muestra en la Figura 67.

La Figura 145 muestra una secuencia de aminoácidos (SEQ ID No: 145) derivada de la secuencia de codificación de SEQ ID No:68 que se muestra en la Figura 68.

La Figura 146 muestra una secuencia de aminoácidos (SEQ ID No: 146) derivada de la secuencia de codificación de SEQ ID No:69 que se muestra en la Figura 69.

La Figura 147 muestra una secuencia de aminoácidos (SEQ ID No: 147) derivada de la secuencia de codificación de SEQ ID No:70 que se muestra en la Figura 70.

La Figura 148 muestra una secuencia de aminoácidos (SEQ ID No: 148) derivada de la secuencia de codificación de SEQ ID No:71 que se muestra en la Figura 71.

La Figura 149 muestra una secuencia de aminoácidos (SEQ ID No: 149) derivada de la secuencia de codificación de SEQ ID No:73 que se muestra en la Figura 73.

La Figura 150 muestra una secuencia de aminoácidos (SEQ ID No: 150) derivada de la secuencia de codificación de SEQ ID No:74 que se muestra en la Figura 74.

La Figura 151 muestra una secuencia de aminoácidos (SEQ ID No: 151) derivada de la secuencia de codificación de SEQ ID No:75 que se muestra en la Figura 75.

La Figura 152 muestra una secuencia de aminoácidos (SEQ ID No: 152) derivada de la secuencia de codificación de SEQ ID No:76 que se muestra en la Figura 76.

La Figura 153 muestra una secuencia de aminoácidos (SEQ ID No: 153) derivada de la secuencia de codificación de SEQ ID No:77 que se muestra en la Figura 77.

La Figura 154 muestra una secuencia de aminoácidos (SEQ ID No: 154) derivada de la secuencia de codificación de SEQ ID No:78 que se muestra en las Figuras 78A-B.

La Figura 155 es un diagrama que representa la estructura tridimensional del polipéptido E16 (TAT188) como una subunidad del sistema transportador de aminoácidos neutros grandes, independiente de iones de sodio.

La Figura 156 muestra los resultados gráficos de los diagramas FACS que muestran la fijación de anticuerpos anti-TAT188 (anti-E16) a la superficie de la célula PC3 (diagrama en verde) y la reducción de la fijación con la reducción de la expresión del TAT188 (diagrama rojo).

La Figura 157 muestra la internalización del anticuerpo anti-TAT188 después de fijarse a la superficie del TAT188- que expresa las células PC3.

La Figura 158 es una gráfica de barras que muestra los cambios en la actividad de transporte de aminoácidos en el TAT188 en presencia del anticuerpo anti-TAT188 con el tiempo.

La Figura 159 dispone diagramas de la viabilidad de células en presencia de un incremento de las cantidades de MC-vc-PAB- MMAE anticuerpo anti-TAT188 conjugado a toxina en células PC3 y Colo205.

La Figura 160 dispone diagramas de la viabilidad de células en presencia de un incremento de las cantidades de MC-vc-PAB-MMAE o MC-vc-PAB-MMAF anticuerpo anti-TAT188 conjugado a toxina en células Colo205.

La Figura 161 muestra los resultados de las células de fijación endógena de anticuerpo E16 anti-humano que expresan el polipéptido E16, donde las células son células de mono COS7, células humanas de cáncer de mama MCF10A y células de ratón NIH-3T3. La Figura 161A muestra los resultados FACS de la fijación de anticuerpos 3B5anti-E16, 12B9 y 12G12 a monos COS. La Figura 161B muestra la fijación de anti-E16 3B5 a la

estirpe celular humana de cáncer de mama MCF10A y no hay fijación a las células NIH-3T3 de ratón.

La Figura 162 es un diagrama de los cambios en el volumen medio del tumor durante un tiempo en ratones xenotransplantados a los cuales se administró anticuerpos anti-E16 desnudos (sin conjugar con toxina).

La Figura 163 es una gráfica de los cambios en el volumen medio de un tumor durante un tiempo en ratones xenotransplantados a los cuales se administraron anticuerpos anti-E16 conjugados con toxina, donde la toxina era MC-vc-PAB-MMAE o MC-vc-PAB-MMAF.

La Figura 164 muestra que la expresión del polipéptido E16-GFP de fusión se reduce en las células transinfectadas con ARNsi que tienen por objetivo el gen E16. El panel superior contiene unas fotografías de los resultados de análisis de bioluminiscencia para notar la presencia de GFP. El panel inferior contiene fotografías de la visualización del contraste de fases que muestran que la reducción de bioluminiscencia GFP no fue causada por la reducción del número de células.

La Figura 165 es un análisis de inmunotransferencia que muestra que la expresión del polipéptido E16-GFP de fusión se inhibe en las células PC3 transinfectadas con E16 ARNsi. β-tubulina es un gel de control de carga.

La Figura 166 muestra que la transinfección transiente de las células PC3 con E16 ARNsi está asociada a la reducción de la función de transporte de aminoácidos, consistente con una expresión reducida de E16.

La Figura 167 es un diagrama de proliferación celular durante un tiempo con y sin una transinfección transiente con E16 ARNsi. La proliferación celular se reduce en las células PC3 transinfectadas transientemente con E16.

DESCRIPCIÓN DETALLADA DE LAS INCORPORACIONES PREFERIDAS

I. Definiciones

Los términos "polipéptido TAT" y "TAT" como se emplean en la presente, y cuando están seguídos inmediatamente de una denominación

numérica, hacen referencia a varios polipéptidos, donde la denominación completa (es decir TAT/número) hace referencia a unas secuencias específicas de polipéptidos como se describe en la presente. Los términos "polipéptido TAT/número" y "TAT/número" donde el término "número" se dispone como una denominación numérica real como se emplea en la presente, abarcan las secuencias naturales de polipéptidos, variantes de polipéptidos y fragmentos de secuencias naturales de polipéptidos y variantes de polipéptidos (que se definen en más detalle en la presente). Los polipéptidos TAT descritos en la presente se pueden aislar de una variedad de fuentes, como de tipos de tejidos humanos o de otra fuente, o se pueden preparar mediante métodos recombinantes o sintéticos. El término "polipéptido TAT" hace referencia a cada polipéptido individual TAT/número que se divulga en la presente. Todas las divulgaciones en esta especificación que hacen referencia al "polipéptido TAT" se refieren a cada uno de los polipéptidos individualmente, al igual que conjuntamente. Por ejemplo, descripciones de la preparación de, de la purificación de, de la derivación de, de la formación de anticuerpos para o contra la formación de oligopéptidos TAT de unión para o contra la formación de moléculas orgánicas TAT de unión para o contra, la administración de, las composiciones conteniendo, el tratamiento de una enfermedad con, etc. corresponden a cada polipéptido de la invención individualmente. El término "polipéptido TAT" también incluye variantes de los polipéptidos TAT/número como se divulga en la presente.

"La secuencia natural del polipéptido TAT" comprende un polipéptido teniendo la misma secuencia de aminoácidos que el polipéptido TAT correspondiente derivado de la naturaleza. Dichas secuencias naturales de polipéptidos TAT se pueden aislar de la naturaleza o se pueden producir mediante técnicas recombinantes o sintéticas. El término "la secuencia natural del polipéptido TAT" específicamente comprende formas truncadas naturalmente o secretadas del polipéptido TAT específico (por ejemplo, una secuencia de dominio extracelular), formas variantes en la naturaleza (por ejemplo, alternativamente formas empalmadas) y variantes

546
547

alélicas naturales del polipéptido. En ciertas incorporaciones de la invención, las secuencias naturales de los polipéptidos TAT como se divulga en la presente son polipéptidos maduros o son secuencias naturales de polipéptidos de longitud total que comprenden las secuencias de aminoácidos de longitud total que se muestran en las Figuras adjuntas. Los codones de inicio y de parada (si se indican) se muestran en negrilla y subrayado en las Figuras. Los residuos de ácido nucleico indicados como "N" en las Figuras adjuntas son cualquier residuo de ácido nucleico. Sin embargo aunque los polipéptidos TAT divulgados en las Figuras adjuntas se muestran con un comienzo de residuos de metionina denominados en la presente un aminoácido en posición 1 en las Figuras, es concebible y posible emplear otros residuos de metionina ubicados en sentido ascendente o descendente del aminoácido en posición 1 en las Figuras como el residuo de inicio para los polipéptidos TAT.

El "dominio extracelular" del polipéptido TAT o "ECD" hace referencia a una forma del polipéptido TAT el cual está esencialmente libre de los dominios transmembrana y citoplásmicos. Normalmente, un polipéptido TAT ECD tendrá menos de 1% de los citados dominios de transmembrana y/o citoplásmicos y preferiblemente, tendrá menos de 0.5% de los citados dominios. Se entenderá que cualquier dominio transmembrana identificado para los polipéptidos TAT de la presente invención se identifica de conformidad con unos criterios empleados de manera rutinaria en el arte para identificar ese tipo de dominio hidrófobo. Los límites exactos de un dominio transmembrana pueden variar, pero muy probablemente no más de alrededor de 5 aminoácidos en cualquiera de los extremos del dominio como se identificó inicialmente en la presente. Opcionalmente, por lo tanto, un dominio extracelular de un polipéptido TAT puede contener desde alrededor de 5 o menos aminoácidos en cualquier lado del límite del dominio transmembrana/dominio extracelular como se identifica en los Ejemplos o en la especificación, y los citados polipéptidos, con o sin el péptido de señal asociada y el ácido nucleico codificándolos, se contemplan en la presente invención.

La ubicación aproximada de los "péptidos de señal" de los diferentes polipéptidos TAT como se divulga en la presente se puede mostrar en la presente especificación y/o en las Figuras adjuntas. Sin embargo se nota que el límite del terminal C de un péptido de señal puede variar, pero seguramente en no más de alrededor de 5 aminoácidos en cualquiera de los lados del límite del terminal C del péptido de señal como se identifica inicialmente en la presente, donde el límite del terminal C del péptido de señal se puede identificar de conformidad con unos criterios empleados en forma rutinaria en el arte para identificar ese tipo de elementos de la secuencia de aminoácidos (por ejemplo, Nielsen y colaboradores, Prot. Eng. 10:1-6 (1997) y von Heinje y colaboradores, Nucl. Acids. Res. 14:4683-4690 (1986)). Por otra parte, igualmente se reconoce que en algunos casos, la segmentación de una secuencia de señal de un polipéptido secretado no es completamente uniforme, resultando en más de una especie secretada. Estos polipéptidos maduros, donde el péptido de señal se segmenta con no más de alrededor de 5 aminoácidos en cada lado del límite del terminal C del péptido de señal como se identifica en la presente y los polinucleótidos codificándolos, se contemplan en la presente invención.

"Variante de polipéptido TAT" significa un polipéptido TAT, preferiblemente un polipéptido TAT activo, como se define en la presente, que tiene una identidad de secuencia de aminoácidos por lo menos alrededor del 80% con una secuencia natural de longitud total de secuencia de polipéptido TAT como se divulga en la presente, una secuencia de polipéptido TAT que carece del péptido de señal como se divulga en la presente, un dominio extracelular de un polipéptido TAT, con o sin el péptido de señal, como se divulga en la presente o cualquier otro fragmento de secuencia de polipéptido TAT de longitud total como se divulga en la presente (como aquellas codificadas por un ácido nucleico que representa únicamente una porción de la secuencia completa de codificación de un polipéptido TAT de longitud total). Las citadas variantes del polipéptido TAT incluyen por ejemplo, polipéptidos TAT donde uno o varios residuos de aminoácidos se agregan o se eliminan en el terminal N o C de la

secuencia natural de aminoácidos de longitud total. Normalmente, una variante de polipéptido TAT tendrá por lo menos una identidad de la secuencia de aminoácidos alrededor de 80%, alternativamente una identidad de la secuencia de aminoácidos por lo menos alrededor de 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99%, con una secuencia natural del polipéptido TAT de longitud total como se divulga en la presente, una secuencia de polipéptido TAT que carece del péptido de señal como se divulga en la presente, un dominio extracelular de un polipéptido TAT, con o sin el péptido de señal, como se divulga en la presente o cualquier otro fragmento específicamente definido de una secuencia de polipéptido TAT de longitud total como se divulga en la presente. Normalmente, las variantes de polipéptidos TAT tienen por lo menos alrededor de 10 aminoácidos de largo, alternativamente por lo menos alrededor de 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600 aminoácidos de largo, o más. Opcionalmente, las variantes de polipéptidos TAT tendrán no más de una sustitución conservadora de aminoácido cuando se comparan con la secuencia natural del polipéptido TAT, alternativamente no más de 2, 3, 4, 5, 6, 7, 8, 9, o 10 sustituciones conservadoras de aminoácidos cuando se comparan con la secuencia natural del polipéptido TAT.

"Porcentaje (%) de identidad de la secuencia de aminoácidos" con respecto a las secuencias de polipéptidos TAT identificadas en la presente, se define como el porcentaje de residuos de aminoácidos de una secuencia candidata que son todos idénticos a los residuos de aminoácido en la secuencia específica del polipéptido TAT, después de alinear las secuencias e introducir espacios, de ser necesario, para alcanzar el porcentaje máximo de identidad de la secuencia y sin considerar ninguna sustitución conservadora como parte de la identidad. La alineación, para efectos de determinar el porcentaje de identidad de la secuencia de aminoácidos, se

puede lograr de varias formas dentro de las destrezas del arte, por ejemplo, utilizando un software de computador disponible para el público como BLAST, BLAST-2, ALIGN o Megalign (DNASTAR). Aquellos con destreza en el arte pueden determinar los parámetros apropiados para medir la alineación, incluyendo cualquier algoritmo que se requiera para alcanzar la alineación máxima en toda la longitud de la secuencia que se está comparando. Sin embargo, para los fines de la presente, los valores del % de identidad de la secuencia de aminoácidos se generan empleando el programa de computador para la comparación de secuencias ALIGN-2, donde se dispone el código completo de fuente para el programa ALIGN-2 en la Tabla 1 abajo. Genentech, Inc. es el autor del programa de computador para la comparación de secuencias ALIGN-2 y el código de fuente que aparece en la Tabla 1 fue presentado junto con la documentación de usuario a la Oficina de Derechos de Autor de Washington D.C., 20559, Estados Unidos donde se encuentra registrado bajo el Registro de Derechos de Autor de Estados Unidos No. TXU510087. El programa ALIGN-2 está disponible para el público a través de Genentech, Inc., South San Francisco, California o se puede compilar del código de fuente dispuesto en la Tabla 1 abajo. El programa ALIGN-2 se debe compilar para ser utilizado en un sistema operativo UNIX, preferiblemente digital UNIX V4.0D. El programa ALIGN-2 fija todos los parámetros de comparación de secuencias y estos no varían.

En situaciones donde se utiliza el ALIGN-2 para las comparaciones de secuencias de aminoácidos, el % de identidad de la secuencia de aminoácidos para determinada secuencia de aminoácidos A, con o frente a determinada secuencia de aminoácidos B (que alternativamente se puede expresar como determinada secuencia de aminoácidos A que tiene o comprende cierto % de identidad de la secuencia de aminoácidos con frente a determinada secuencia de aminoácidos B) se calcula de la siguiente forma:

100 veces la fracción X/Y

donde X es el número de residuos de aminoácidos totalizados como compatibilidades idénticas por el programa ALIGN-2 de alineación de secuencia de A y B, y donde Y es el número total de residuos de aminoácidos en B. Se apreciará que cuando la longitud de la secuencia de aminoácidos A no es igual a la longitud de la secuencia de aminoácidos B, el % de identidad de la secuencia de aminoácidos de A a B no será igual al % de identidad de la secuencia de aminoácidos de B a A. Como Ejemplos de cálculos de un % de identidad de la secuencia de aminoácidos empleando este método, las Tablas 2 y 3 muestran como calcular el % de identidad de una secuencia de aminoácidos denominada "Proteína de Comparación" a una secuencia de aminoácidos denominada "TAT", donde "TAT" representa una secuencia de aminoácidos de un polipéptido TAT determinado de interés, la "Proteína de Comparación" representa una secuencia de aminoácidos de un polipéptido contra el cual se está comparando el polipéptido "TAT" determinado de interés y cada uno de "X", "Y" y "Z" representa unos residuos hipotéticos diferentes de aminoácidos. Salvo si se indica específicamente en contrario, todos los valores de % de identidad de la secuencia de aminoácidos empleados en la presente se obtienen como se describe en el párrafo inmediatamente anterior, empleando el programa de computador ALIGN-2.

"Variante de polinucleótido TAT", o "variante de secuencia de ácido nucleico TAT" significa una molécula de ácido nucleico que codifica un polipéptido TAT, preferiblemente un polipéptido TAT activo, como se define en la presente, y que tiene una identidad de secuencia de ácido nucleico por lo menos de alrededor de 80%, con una secuencia de ácido nucleótido codificando una secuencia natural de longitud total de secuencia de polipéptido TAT como se divulga en la presente, una secuencia natural de longitud total de secuencia de polipéptido TAT que carece del péptido de señal como se divulga en la presente, un dominio extracelular de un polipéptido TAT, con o sin el péptido de señal, como se divulga en la presente o cualquier otro fragmento de una secuencia de polipéptido TAT

de longitud total como se divulga en la presente (como aquellos codificados por un ácido nucleico que únicamente representa una porción de la secuencia de codificación completa de un polipéptido TAT de longitud total). Normalmente, una variante de polinucleótido TAT tendrá una identidad de secuencia de ácido nucleico por lo menos de alrededor de 80%, alternativamente una identidad de secuencia de ácido nucleico por lo menos alrededor de 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% con una secuencia de ácido nucleico codificando una secuencia natural de longitud total de secuencia de polipéptido TAT como se divulga en la presente, una secuencia natural de longitud total de secuencia de polipéptido TAT que carece del péptido de señal como se divulga en la presente, un dominio extracelular de un polipéptido TAT, con o sin la secuencia de señal, como se divulga en la presente o o cualquier otro fragmento de una secuencia de polipéptido TAT de longitud total como se divulga en la presente. Las variantes no abarcan la secuencia nucleótida natural.

Normalmente, las variantes de polinucleótidos TAT tienen por lo menos alrededor de 5 nucleótidos de largo, alternativamente por lo menos alrededor de 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, o 1000 nucleótidos de largo, donde en este contexto, el término "alrededor de" significa que la citada longitud de secuencia nucleótida es más o menos el 10% de la citada longitud.

"Porcentaje (%) de identidad de secuencia de ácido nucleico" con respecto a las secuencias de ácido nucleico codificando TAT identificadas

554
553

en la presente, se define como el porcentaje de nucleótidos en una secuencia candidata que son idénticos a los nucleótidos en la secuencia del ácido nucleico TAT de interés, después de alinear las secuencias e introducir espacios, de ser necesario, para alcanzar el porcentaje máximo de identidad de la secuencia. La alineación, para efectos de determinar el porcentaje de identidad de secuencia de ácido nucleico, se puede lograr de varias formas dentro de las destrezas del arte, por ejemplo, utilizando un software de computador disponible para el público como BLAST, BLAST-2, ALIGN o Megalign (DNASTAR). Sin embargo, para los fines de la presente, los valores del % de identidad de secuencia de ácido nucleico se generan empleando el programa de computador para la comparación de secuencias ALIGN-2, donde se dispone el código completo de fuente para el programa ALIGN-2 en la Tabla 1 abajo. Genentech, Inc. es el autor del programa de computador para la comparación de secuencias ALIGN-2 y el código de fuente que aparece en la Tabla 1 fue presentado junto con la documentación de usuario a la Oficina de Derechos de Autor de Washington D.C., 20559, Estados Unidos donde se encuentra registrado bajo el Registro de Derechos de Autor de Estados Unidos No. TXU510087. El programa ALIGN-2 está disponible para el público a través de Genentech, Inc., South San Francisco, California o se puede compilar del código de fuente dispuesto en la Tabla 1 abajo. El programa ALIGN-2 se debe compilar para ser utilizado en un sistema operativo UNIX, preferiblemente digital UNIX V4.0D. El programa ALIGN-2 fija todos los parámetros de comparación de secuencias y estos no varían.

En situaciones donde se utiliza ALIGN-2 para las comparaciones de secuencias de ácido nucleico, el % de identidad de secuencia de ácido nucleico para determinada secuencia de ácido nucleico C, con o frente a determinada secuencia de ácidos nucleico D (que alternativamente se puede expresar como determinada secuencia de ácido nucleico C que tiene o comprende cierto % de identidad de secuencia de ácido nucleico con o frente a determinada secuencia de ácido nucleico D) se calcula de la siguiente forma:

100 veces la fracción W/Z

donde W es el número de nucleótidos totalizados como compatibilidades idénticas por el programa ALIGN-2 de alineación de secuencia de C y D y donde Z es el número total de nucleótidos en D. Se apreciará que cuando la longitud de la secuencia de ácido nucleico C no es igual a la longitud de la secuencia de ácido nucleico D, el % de identidad de secuencia de ácido nucleico C a D no será igual al % de identidad de la secuencia de ácido nucleico D a C. Como Ejemplos de cálculos de un % de identidad de secuencia de ácido nucleico las Tablas 4 y 5 muestran como calcular el % de identidad de una secuencia de ácido nucleico denominada "ADN de Comparación" a una secuencia de ácido nucleico denominada "TAT-ADN", donde "TAT-ADN" representa una secuencia hipotética de ácido nucleico codificando el TAT de interés, el "ADN de Comparación" representa la secuencia de nucleótidos de una molécula de ácido nucleico contra la cual se está comparando la molécula de ácido nucleido "TAT-ADN" de interés y cada uno de "N", "L" y "V" representa diferentes nucleótidos hipotéticos. Salvo si se indica específicamente en contrario, todos los valores de % de identidad de secuencia de ácido nucleico empleados en la presente se obtienen como se describe en el párrafo inmediatamente anterior, empleando el programa de computador ALIGN-2.

En otras incorporaciones, los variantes de polinucleótidos TAT son las moléculas de ácido nucleico que codifican un polipéptido TAT y que son capaces de hibridación, preferiblemente bajo condiciones severas de hibridación y lavado con las secuencias nucleótidas que codifican un polipéptido TAT de longitud total como se divulga en la presente. Variantes de polipéptidos TAT pueden ser aquellos codificados por una variante de polinucleótido TAT.

El término "región de codificación de longitud total" cuando se emplea con respecto a un ácido nucleico codificando un polipéptido TAT hace referencia a la secuencia de nucleótidos que codifican el polipéptido TAT de longitud total de la invención (el cual con frecuencia se muestra

entre los codones de inicio y de parada, inclusivos de estos, en las Figuras adjuntas). El término "región de codificación de longitud total" cuando se emplea con respecto a un ácido nucleico depositado en la ATCC, hace referencia a la porción codificando el polipéptido codificando el TAT del ADNc que se inserta en el vector depositado en la ATCC (el cual con frecuencia se muestra entre los codones de inicio y de parada, inclusivos de estos, en las Figuras adjuntas).

"Aislado", cuando se utiliza para describir los diferentes polipéptidos TAT como se divulga en la presente, significa el polipéptido que ha sido identificado y separado y/o recuperado de un componente de su entorno natural. Los componentes contaminantes de su entorno natural son materiales que típicamente interferirían con los usos de diagnóstico o terapéutico del polipéptido y pueden incluir enzimas, hormonas y otros solutos proteináceos o no-proteináceos. En unas incorporaciones preferidas, el polipéptido será purificado (hasta un grado suficiente para obtener por lo menos 15 residuos de secuencia de aminoácidos de terminal N o internos mediante el uso de un secuenciador de copa giratoria o (2) hasta su homogeneidad por SDS-PAGE bajo condiciones de reducción o sin reducción empleando azul Coomassie o preferiblemente mancha plateada. Los polipéptidos aislados incluyen el polipéptido *in situ* en células recombinantes puesto que por lo menos un componente del medio natural del polipéptido TAT no estará presente. Sin embargo, por lo general el polipéptido aislado se preparará mediante por lo menos un paso de purificación.

Un polipéptido TAT codificando ácido nucleico "aislado" u otro polipéptido codificando ácido nucleico es una molécula de ácido nucleico que se identifica y se separa de por lo menos una molécula de ácido nucleico contaminante que normalmente se asocia a la fuente natural del polipéptido codificando ácido nucleico. Un polipéptido aislado codificando la molécula de ácido nucleico tiene una forma y en entorno diferente al que tiene en la naturaleza. Por lo tanto, los polipéptidos TAT codificando las moléculas de ácido nucleico se distinguen del polipéptido específico

codificando la molécula de ácido nucleico puesto que existe en las células naturales. Sin embargo, un polipéptido aislado codificando la molécula de ácido nucleico incluye un polipéptido codificando las moléculas de ácido nucleicos contenido en las células que normalmente expresan el polipéptido donde, por ejemplo, la molécula de ácido nucleico se encuentra en una ubicación cromosomal diferente a la de las células naturales.

El término "secuencias de control" hace referencia a secuencias de ADN necesarias para la expresión de una secuencia de codificación operacionalmente enlazada en un organismo hospedero en particular. Las secuencias de control apropiadas para las procariotas incluyen por ejemplo un promotor, opcionalmente una secuencia operadora y un punto de unión al ribosoma. Se sabe que las células eucariotas utilizan promotores, señales de poliadenilación e intensificadores.

El ácido nucleico está "operacionalmente enlazado" cuando se sitúa en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, el ADN para una pre-secuencia o un líder secretor está enlazado operacionalmente al ADN para un polipéptido si está expresado como una pre-proteína que participa en la secreción del polipéptido; un promotor o un intensificador está enlazado operacionalmente a una secuencia de codificación si afecta la transcripción de la secuencia; o un punto de unión al ribosoma está enlazado operacionalmente a una secuencia de codificación si está posicionado de forma que facilita la traslación. Generalmente, "operacionalmente enlazado" significa que las secuencias de ADN que se están enlazando están contiguas, y en el caso de un líder secretor, contiguas y en fase de lectura. Sin embargo, los intensificadores no tienen que ser contiguos. El enlace se logra mediante un ligado en unos sitios de restricción apropiados. Si dichos sitios no existen, se emplean los adaptadores o enlazantes oligonucleótidos sintéticos de conformidad con la práctica convencional.

Una persona con destreza en el arte fácilmente puede determinar las reacciones de hibridación "severas" y generalmente es un cálculo empírico que depende de la longitud de sonda, de la temperatura de lavado y la

concentración de sal. Por lo general, unas sondas más largas requieren unas temperaturas más altas para su hibridación correcta, mientras que unas sondas más cortas necesitan unas temperaturas más bajas. Generalmente la hibridación depende de la habilidad del ADN desnaturalizado para hibridarse cuando se encuentran presentes unas hebras complementarias en un entorno por debajo de su temperatura de fusión. Cuanto mayor es el grado de homología deseado entre la sonda y la secuencia de hibridación, más alta será la temperatura relativa que puede emplear. Como resultado se sabe que unas temperaturas relativas más altas van a tender a hacer las condiciones de reacción más severas, mientras que las temperaturas más bajas, lo harán menos. Para detalles y explicaciones adicionales sobre la severidad de las reacciones de hibridación, consultar Ausubel y colaboradores, Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

"Condiciones severas" o "condiciones altamente severas", como se definen en la presente, se pueden identificar por aquellos que: (1) emplean una baja fortaleza iónica y una temperatura alta para lavado, por ejemplo 0.015 M cloruro de sodio/0.0015 M citrato de sodio/dodecil sulfato sódico al 0.1% a 50°C; (2) utilizan un agente desnaturalizante durante la hibridación como formamida, por ejemplo, formamida al 50% (v/v) con albúmina de suero bovino al 0.1%/Ficoll al 0.1%/polivinilpirrolidona al 0.1% /50mM solución amortiguadora de fosfato de sodio con un pH 6.5 con 750 mM cloruro de sodio, 75 mM citrato de sodio a 42°C; o (3) hibridación durante la noche en una solución que emplea formamida al 50%, 5 x SSC (0.75 M NaCl, 0.075 M citrato de sodio), 50 mM fosfato de sodio (pH 6.8), pirofosfato de sodio al 0.1%, 5 x solución de Denhardt, ADN de esperma de salmón sonicado (50 µg/ml), SDS al 0.1% y sulfato de dextrano al 10% a 42°C, con un lavado de 10 minutos a 42°C en 0.2 x SSC (cloruro de sodio/citrato de sodio) seguido por un lavado altamente riguroso de 10 minutos que consiste en 0.1 x SSC conteniendo EDTA a 55°C.

Las "condiciones moderadamente severas" se pueden identificar como se describe por Sambrook y colaboradores, Molecular Cloning: A

Laboratory Manual, New York: Cold Spring Harbor Press, 1989 e incluyen el uso de una solución de lavado y unas condiciones de hibridación (por ejemplo, temperatura, fortaleza iónica y %SDS) menos severas que las descritas arriba. Un ejemplo de unas condiciones moderadamente severas es la incubación durante la noche a 37°C en una solución que comprende: formamida al 20%, 5 x SSC (150 mM NaCl, 15 mM tricitrato de sodio), 50 mM fosfato de sodio (pH 7.6), 5 x solución de Denhardt, sulfato de dextrano al 10% y 20 mg/ml de ADN de esperma de salmón desnaturalizado cortado seguido por el lavado de los filtros en 1 x SSC a alrededor de 37-50°C. El artesano con destrezas va a reconocer cómo ajustar la temperatura, la fortaleza iónica, etc. según sea necesario para ajustar factores como la longitud de sonda y similares.

El término "marcador de epítopo" cuando se emplea en la presente, hace referencia a un polipéptido quimérico que comprende un polipéptido TAT o un anticuerpo anti-TAT fusionado a un "polipéptido marcador". El polipéptido marcador tiene suficientes residuos para proveer un epítopo contra el cual se puede hacer un anticuerpo, pero es lo suficientemente corto para no interferir con la actividad del polipéptido al cual está fusionado. Preferiblemente el polipéptido marcador también es bastante único de manera que el anticuerpo no realiza una reacción cruzada con otros epítopos. Generalmente los polipéptidos marcadores apropiados tienen por lo menos seis residuos de aminoácidos, y usualmente entre alrededor de 8 y 50 residuos de aminoácidos (preferiblemente, entre alrededor de 10 y 20 residuos de aminoácidos).

"Activo" o "actividad" para los efectos de la presente hace referencia a forma(s) de un polipéptido TAT que conserva(n) una actividad biológica y/o inmunológica de los TAT naturales o que se encuentran en forma natural, donde actividad "biológica" hace referencia a una función biológica (ya sea inhibidora o estimulante) causada por el TAT natural o que se encuentra en forma natural, diferente a la habilidad de inducir la producción de un anticuerpo contra un epítopo antigénico que posee un TAT natural o que se encuentra en forma natural y una actividad "inmunológica" hace

referencia a la habilidad de inducir la producción de un anticuerpo contra un epítopo antigénico que posee un TAT natural o que se encuentra en forma natural.

El término "antagonista" se emplea en el sentido más amplio e incluye cualquier molécula que parcial o totalmente bloquea, inhibe o neutraliza una actividad biológica de un polipéptido TAT natural que se divulga en la presente. En forma similar, el término "agonista" se utiliza en el sentido más amplio e incluye cualquier molécula que imita una actividad biológica de un polipéptido TAT natural que se divulga en la presente. Unas moléculas agonistas o antagonistas apropiadas específicamente incluyen anticuerpos agonistas o antagonistas o un fragmento de anticuerpos, fragmentos o variantes de secuencias de aminoácidos de polipéptidos TAT naturales, péptidos, oligonucleótidos antisentido, moléculas orgánicas pequeñas, etc. Los métodos para identificar los agonistas o antagonistas de un polipéptido TAT pueden comprender poner en contacto un polipéptido TAT con una molécula agonista o antagonista candidata y medir un cambio detectable en una o en varias actividades biológicas normalmente asociadas con el polipéptido TAT.

"Tratar", "tratamiento" o "aliviar" hacen referencia tanto a tratamientos terapéuticos como a medidas profilácticas o preventivas, donde el objetivo es prevenir o retardar (reducir) la condición o el desorden patológico objetivo. Aquellos que necesitan el tratamiento incluyen los que ya tienen el desorden al igual que los que son propensos a tener el desorden o aquellos en los cuales se debe prevenir. Un sujeto o mamífero se "trata" con éxito por un polipéptido TAT-que expresa cáncer si después de recibir una cantidad terapéutica de un anticuerpo anti-TAT, el oligopéptido TAT de unión o la molécula orgánica TAT de unión de acuerdo con los métodos de la presente invención, el paciente muestra una reducción observable y/o medible en o la ausencia de uno o varios de los siguientes: la reducción del número de células cancerosas o una ausencia de células cancerosas; la reducción del tamaño del tumor; la inhibición (es decir el retardo y preferiblemente la detención) de la infiltración de células cancerosas en

órganos periféricos incluyendo la propagación del cáncer en tejidos blandos y huesos; la inhibición (es decir el retardo y preferiblemente la detención) de la metástasis del tumor; la inhibición hasta cierto punto, del crecimiento del tumor y/o el alivio hasta cierto punto de uno o varios de los síntomas asociados con el cáncer específico; una reducción de la morbimortalidad y una mejoría en temas relacionados con la calidad de vida. Hasta el punto que el anticuerpo anti-TAT o el oligopéptido TAT de unión puede prevenir el crecimiento y/o destruir las células cancerosas existentes, puede ser citostático y/o citotóxico. El paciente también puede sentir la reducción de estas señales o síntomas.

Los anteriores parámetros para evaluar un tratamiento exitoso y la mejoría de la enfermedad son fácilmente medibles mediante unos procedimientos rutinarios conocidos por un galeno. Para la terapia contra el cáncer, se puede medir la eficacia por ejemplo evaluando el tiempo de desarrollo de la enfermedad (TTP) y/o determinar la tasa de respuesta (RR). Se puede determinar la metástasis mediante unos análisis de estadificación y mediante un escaneo de huesos y unos exámenes para conocer el nivel de calcio y otras enzimas para determinar la propagación al hueso. Igualmente se pueden efectuar escaneos CT para mirar si se propagó a la pelvis y a los nódulos linfáticos en el área. Se pueden realizar radiografías de tórax y se pueden medir las enzimas del hígado mediante unos métodos conocidos para buscar una metástasis en pulmón y en hígado, respectivamente. Otros métodos rutinarios para monitorear la enfermedad incluyen ultrasonografías transrectales (TRUS) y biopsias transrectales con aguja (TRNB).

Para el cáncer de vejiga, que es un cáncer más localizado, los métodos para determinar el progreso de la enfermedad incluyen una citología urinaria por cistoscopia, monitorear la presencia de sangre en la orina, visualizar el tracto urotelial mediante tomografía o por medio de un pielograma intravenoso, una tomografía computarizada (CT) y una imagen de resonancia magnética (MRI). La presencia de metástasis distantes se puede evaluar mediante una CT de abdomen, una radiografía de tórax o una gammagrafía radionúclida del esqueleto.

SBZ 561

Administración "crónica" hace referencia a la administración del (los) agente(s) en un modo continuo, en contraposición a un modo agudo, con el fin de mantener el efecto (la actividad) terapéutico(a) durante un largo periodo de tiempo. Administración "intermitente" es un tratamiento que no se hace en forma consecutiva sin interrupción, sino más bien es de naturaleza cíclica.

"Mamífero" para los efectos del tratamiento, de aliviar los síntomas o de diagnosticar un cáncer hace referencia a cualquier animal clasificado como un mamífero, incluyendo humanos, animales domésticos y de granja y zoológicos, de deportes, mascotas como perros, gatos, ganado, caballos, ovejas, cerdos, cabras, conejos, etc. Preferiblemente, el mamífero es humano.

Administración "en combinación con" uno o varios agentes terapéuticos incluye la administración simultánea (concurrente) y consecutiva en cualquier orden.

"Excipientes" como se emplea en la presente incluyen excipientes o estabilizantes farmacéuticamente aceptables que no son tóxicos para la célula o el mamífero que se ve expuesto a estos y a las dosis y las concentraciones empleadas. Con frecuencia un excipiente fisiológicamente aceptable es una solución amortiguadora de pH acuoso. Ejemplos de excipientes fisiológicamente aceptables incluyen soluciones amortiguadoras como fosfato, citrato y otros ácidos orgánicos; antioxidantes; ácido ascórbico; polipéptidos de bajo peso molecular (menos de alrededor de 10 residuos); proteínas, como suero de albúmina, gelatina o inmunoglobulinas; polímeros hidrofílicos como polivinilpirrolidona; aminoácidos como glicina, glutamina, asparagina, arginina o lisina; monosacáridos, disacáridos y otros carbohidratos incluyendo glucosa, manosa o dextrinas, agentes quelantes como EDTA; alcoholes de azúcar como manitol o sorbitol; contraiones que forman sales como sodio y/o surfactantes no iónicos como TWEEN®, polietilenglicol (PEG) y PLURONICS®.

Con "fase sólida" o "soporte sólido" se indica una matriz no acuosa a la cual se puede adherir o unir un anticuerpo, un oligopéptido TAT de unión

563
562

o una molécula orgánica TAT de unión de la presente invención. Ejemplos de fases sólidas comprendidas en la presente incluyen aquellas formadas parcial o totalmente por vidrio (por ejemplo, vidrio de poro controlado), polisacáridos (por ejemplo, agarosa), poliacrilamidas, poliestireno, polivinil alcohol y siliconas. En ciertas incorporaciones, dependiendo del contexto, la fase sólida puede comprender el pozo de una placa de análisis; en otros una columna de purificación (por ejemplo, una cromatografía en columna por afinidad). Este término también incluye una fase sólida discontinua de partículas discretas, como aquellas descritas en la Patente de Estados Unidos No. 4,275,149.

Un "liposoma" es una pequeña vesícula compuesta por varios tipos de lípidos, fosfolípidos y/o surfactantes que es útil para administrar un fármaco (como un polipéptido TAT, un anticuerpo del mismo o un oligopéptido TAT de unión) a un mamífero. Los componentes del liposoma comúnmente están ordenados en una formación de dos capas, similar a la distribución de las membranas biológicas.

Una molécula "pequeña" o una molécula orgánica "pequeña" se define en la presente como teniendo un peso molecular por debajo de alrededor de 500 Daltons.

Una "cantidad efectiva" de un polipéptido, de un anticuerpo, del oligopéptido TAT de unión, de la molécula orgánica TAT de unión o de un agonista o antagonista de los mismos, como se divulga en la presente, es una cantidad suficiente para cumplir específicamente un propósito determinado. Una "cantidad efectiva" se puede determinar en forma empírica y rutinaria, en relación con el propósito determinado.

El término "cantidad terapéuticamente efectiva" hace referencia a una cantidad de un anticuerpo, de un polipéptido, del oligopéptido TAT de unión, de la molécula orgánica TAT de unión u otro fármaco efectivo para "tratar" una enfermedad o un desorden en un sujeto o mamífero. En el caso del cáncer, la cantidad terapéuticamente efectiva del fármaco puede reducir el número de células cancerosas; reducir el tamaño del tumor; inhibir (es decir, hasta cierto punto retardar y preferiblemente detener) la infiltración

de células cancerosas en órganos periféricos; inhibir (es decir retardar hasta cierto punto y preferiblemente detener) la metástasis del tumor; inhibir hasta cierto punto, el crecimiento del tumor y/o aliviar hasta cierto punto uno o varios de los síntomas asociados con el cáncer. Consultar la definición de "tratar" en la presente. Hasta el punto que el anticuerpo anti-TAT o el oligopéptido TAT de unión puede prevenir el crecimiento y/o destruir las células cancerosas existentes, puede ser citostático y/o citotóxico.

Una "cantidad inhibidora de crecimiento" de un anticuerpo anti-TAT, de un polipéptido TAT, del oligopéptido TAT de unión, de un TAT ARNsi (como un TAT188 ARNsi) o de la molécula orgánica TAT de unión es una cantidad capaz de inhibir el crecimiento de una célula, especialmente de un tumor, por ejemplo, una célula cancerosa, ya sea *in vitro* o *in vivo*. Una "cantidad inhibidora de crecimiento" de un anticuerpo anti-TAT, de un polipéptido TAT, del oligopéptido TAT de unión, de un TAT ARNsi (como un TAT188 ARNsi) o de la molécula orgánica TAT de unión con el fin de inhibir el crecimiento celular neoplásico se puede determinar de forma empírica y rutinaria.

Una "cantidad citotóxica" de un anticuerpo anti-TAT, de un polipéptido TAT, de un oligopéptido TAT de unión, de un TAT ARNsi (como un TAT188 ARNsi) o de la molécula orgánica TAT de unión es una cantidad capaz de causar la destrucción de una célula, especialmente de un tumor, por ejemplo, una célula cancerosa, bien sea *in vitro* o *in vivo*. Una "cantidad citotóxica" de un anticuerpo anti-TAT, de un polipéptido TAT, de un oligopéptido TAT de unión, de un TAT ARNsi (como un TAT188 ARNsi) o de la molécula orgánica TAT de unión con el fin de inhibir el crecimiento celular neoplásico se puede determinar de forma empírica y rutinaria.

El término "anticuerpo" se utiliza en el sentido más amplio y abarca específicamente, por ejemplo anticuerpos monoclonales simples anti-TAT (incluyendo agonista, antagonista y anticuerpos neutralizantes), composiciones de anticuerpos anti-TAT con especificidad poliepitópica,

anticuerpos policlonales, anticuerpos anti-TAT de una sola cadena y fragmentos de anticuerpos anti-TAT (ver abajo) siempre que exhiban la actividad biológica o inmunológica deseada. El término "inmunoglobulina" (Ig) se emplea en forma intercambiable en la presente con anticuerpo.

Un "anticuerpo aislado" es uno que fue identificado y separado y/o recuperado de un componente de su entorno natural. Los componentes contaminantes de su entorno natural son materiales que interferirían con los usos diagnósticos y terapéuticos del anticuerpo y pueden incluir enzimas, hormonas y otros solutos proteináceos o nonproteináceo. En unas incorporaciones preferidas, el anticuerpo será purificado (1) hasta que se determine por lo menos el 95% por peso del anticuerpo como se determina por el método de Lowry, y más preferiblemente por lo menos hasta el 99% por peso; (2) hasta un grado suficiente para obtener por lo menos 15 residuos de secuencia de aminoácidos de terminal N o internos mediante el uso de un secuenciador de copa giratoria o (3) hasta su homogeneidad por SDS-PAGE bajo condiciones de reducción o sin reducción empleando azul Coomassie o preferiblemente mancha plateada. Los anticuerpos aislados incluyen el anticuerpo *in situ* en células recombinantes puesto que por lo menos un componente del medio natural del anticuerpo no estará presente. Sin embargo, por lo general el anticuerpo aislado se preparará mediante por lo menos un paso de purificación.

La unidad del anticuerpo básico de 4 cadenas es una glicoproteína heterotetramérica compuesta por dos cadenas livianas idénticas (L) y dos cadenas pesadas idénticas (H) (un anticuerpo IgM consiste en 5 de las unidades básicas heterotetraméricas junto con un polipéptido adicional, denominado cadena J, y por lo tanto contiene 10 puntos de unión de antígenos, mientras que los anticuerpos IgA secretados se pueden polimerizar para formar unos conjuntos polivalentes que comprenden 2-5 de las unidades básicas de 4 cadenas junto con la cadena J). En el caso de los IgGs, generalmente la unidad de 4 cadenas es de alrededor de 150,000 daltons. Cada cadena L está enlazada a una cadena G por medio de un enlace covalente disúlfido, mientras que las dos cadenas H se enlazan entre

sí por medio de uno o dos enlaces disúlfidos dependiendo del isótopo de la cadena H. Tanto la cadena H como la L también tienen unos puentes disúlfidos entre cadenas espaciados en forma regular. Cada cadena H tiene un dominio variable (VD según la sigla en inglés) en el terminal N seguido por tres dominios constantes (C_H) para cada una de las cadenas α y γ y cuatro dominios C_H para μ y los isotipos ε. Cada cadena L tiene un dominio variable (V_L) en el terminal N seguido por un dominio constante (C_L) en su otro extremo. El V_L está alineado con el V_H y el C_L está alineado con el primer dominio constante de la cadena pesada (C_H1). Se cree que unos residuos particulares de aminoácidos forman una interfaz entre los dominios variables de la cadena liviana y de la cadena pesada. Las parejas formadas por V_H y V_L conforman un solo punto de unión del antígeno. Para conocer la estructura y las propiedades de los diferentes tipos de anticuerpos, consultar por ejemplo, Basic and Clinical Immunology, Edición 8, Daniel P. Stites, Abba I. Terr and Tristram G. Parslow (eds.), Appleton & Lange, Norwalk, CT, 1994, página 71 y Capítulo 6.

La cadena L de cualquier especie vertebrada se puede asignar a uno de dos tipos claramente distinguidos, denominados kappa y lambda, con base en una secuencia de aminoácidos de sus dominios constantes. Dependiendo de una secuencia de aminoácidos del dominio constante de sus cadenas pesadas (CH), las inmunoglobulinas se pueden asignar a diferentes tipos o isotipos. Existen cinco clases de inmunoglobulinas: IgA, IgD, IgE, IgG e IgM, teniendo cadenas pesadas denominadas α, δ, ε, γ y μ, respectivamente. Las clases γ y α se dividen además en subclases sobre la base de diferencias menores en la secuencia C_H y la función, por ejemplo, los humanos expresan las siguientes subclases: IgG1, IgG2, IgG3, IgG4, IgA1 e IgA2.

El término "variable" hace referencia al hecho de que ciertos segmentos de los dominios variables difieren enormemente en la secuencia entre anticuerpos. El dominio V media la fijación del antígeno y define la especificidad de un anticuerpo en particular para su antígeno particular. Sin embargo, la variabilidad no se distribuye de forma pareja en el tramo de

110-aminoácidos de los dominios variables. En cambio las regiones V constan de unas porciones relativamente invariadas denominadas regiones de soporte (FRs) de 15-30 aminoácidos separados por unas regiones más cortas extremadamente variables denominadas "regiones hipervariables" que tienen cada una 9-12 aminoácidos de largo. Los dominios variables de las cadenas naturales pesadas y livianas comprenden cada uno cuatro FRs, básicamente adoptando una configuración tipo hoja β conectados por tres regiones hipervariables que forman bucles conectando y en algunos casos formando parte de la estructura β en forma de hoja. Las regiones hipervariables de cada cadena se mantienen juntas en fuerte proximidad por las FRs y con las regiones hipervariables de la otra cadena, contribuyen a la formación del punto de unión de antígeno de anticuerpos (consultar Kabat y colaboradores, Sequences of Proteins of Immunological Interest, Edición 5, Public Health Service, National Institutes of Health, Bethesda, MD. (1991)). Los dominios constantes no están comprometidos directamente en la fijación de un anticuerpo a un antígeno, pero exhiben varias funciones efectoras como la participación del anticuerpo en la citotoxicidad celular dependiente de anticuerpo (ADCC).

El término "región hipervariable" cuando se emplea en la presente hace referencia a residuos del aminoácido de un anticuerpo que son responsables de la fijación del antígeno. Generalmente la región hipervariable (HVR) comprende residuos de aminoácidos de una "región determinante de complementariedad" o "CDR" () (por ejemplo alrededor de los residuos 24-34 (L1), 50-56 (L2) y 89-97 (L3) en el V_L y alrededor de 1-35 (H1), 50-65 (H2) y 95-102 (H3) en el V_H; Kabat y colaboradores, Sequences of Proteins of Immunological Interest, Edición 5, Public Health Service, National Institutes of Health, Bethesda, MD. (1991)) y/o los residuos de un "bucle hipervariable" (por ejemplo residuos 26-32 (L1), 50-52 (L2) y 91-96 (L3) en el V_L y 26-32 (H1), 53-55 (H2) y 96-101 (H3) en el V_H; Chothia and Lesk J. Mol. Biol. 196:901-917 (1987)). Los HVRs también se definen como que abarcan las secuencias como se divulga en la Solicitud de Estados Unidos Serial No., presentada el 19 de febrero de

2004, incorporada en su totalidad a la presente por referencia, donde los HVRs incluyen algunas posiciones de una secuencia de aminoácidos en las regiones variables que se definen como que están en contacto con un ligando de fijación del anticuerpo con base en el análisis de unas estructuras complejas de cristal. Como se emplean en la presente, los términos "HVR" y "CDR" se utilizan en forma intercambiable para abarcar la totalidad de los HVRs definidos como sigue.

El término "región hipervariable" cuando se emplea en la presente hace referencia a las regiones de un anticuerpo de dominio variable cuya secuencia es hipervariable y/o forman unos bucles estructuralmente definidos. Generalmente, los anticuerpos comprenden seis regiones hipervariables; tres en el VH (H1, H2, H3) y tres en el VL (L1, L2, L3). Se emplea un número de delineaciones de región hipervariable y se comprenden en la presente. Las Regiones Determinantes de Complementariedades de Kabat (CDRs) se basan en la variabilidad de la secuencia y son los que se utilizan con mayor frecuencia (Kabat *y colaboradores, Sequences of Proteins of Immunological Interest*, Edición 5, Public Health Service, National Institutes of Health, Bethesda, MD. (1991)). En cambio, Chothia hace referencia a la ubicación de los bucles estructurales (Chothia and Lesk *J. Mol. Biol.* 196:901-917 (1987)). Las regiones hipervariables AbM representan un compromiso entre los CDRs de Kabat y los bucles estructurales de Chothia y se utilizan en el software de modelos de anticuerpo AbM Molecular de Oxford. Las regiones hipervariables de "contacto" se basan en el análisis de las estructuras complejas de cristal disponibles. Los residuos de cada una de estas regiones hipervariables se enumeran a continuación.

Tabla HVR

Bucle	Kabat	AbM	Chothia	Contacto
L1	L24-L34	L24-L34	L26-L32	L30-L36
L2	L50-L56	L50-L56	L50-L52	L46-L55
L2	L89-L97	L89-L97	L91-L96	L89-L96

H1 (Numeración Kabat)	H31-H35B	H26-H35B	H26-H32..34	H30-H35
H1 (Numeración Chothia)	H31-H35	H26-H35	H26-H32	H30-H35
H2	H50-H65	H50-H58	H53-H55	H47-H58
H3	H95-H102	H95-H102	H96-H101	H93-H101

Las regiones hipervariables pueden comprender "regiones hipervariables extendidas" como sigue: 24-34 (L1), 50-56 (L2) y 89-97 (L3) en el VL y 26-35 (H1), 50-65 o 49-65 (H2) y 93-102, 94-102 o 95-102 (H3) en el VH. Los residuos del dominio variable se enumeran de acuerdo con Kabat *y colaboradores, supra* para cada una de estas definiciones.

Los residuos de "soporte" o "FR" son esos residuos de dominio variable diferentes de los residuos de región hipervariable como se definen en la presente.

El término "anticuerpo monoclonal" como se utiliza en la presente hace referencia a un anticuerpo obtenido de una población de anticuerpos sustancialmente homogéneos, es decir, los anticuerpos individuales que comprenden la población son idénticos excepto por unas mutaciones que pueden ocurrir en la naturaleza y que pueden estar presentes en cantidades menores. Los anticuerpos monoclonales son altamente específicos, y se dirigen hacia un solo sitio de antígeno. Adicionalmente, en contraste con preparaciones de anticuerpos policlonales que incluyen diferentes anticuerpos dirigidos a diferentes determinantes (epítopos), cada anticuerpo monoclonal se dirige hacia un solo determinante en el antígeno. Además de su especificidad, los anticuerpos monoclonales tienen la ventaja de que se pueden sintetizar sin verse contaminados por otros anticuerpos. El modificador "monoclonal" no se deberá interpretar como que requiere producción del anticuerpo por un método en particular. Por ejemplo, los anticuerpos monoclonales útiles en la presente invención se pueden preparar mediante la metodología de hibridoma descrita en primer lugar por Kohler y colaboradores, Nature, 256:495 (1975), o se puede manufacturar

empleando métodos de ADN recombinante en células de bacterias, eucariotas de animal o de planta (consultar por ejemplo, Patente de Estados Unidos No. 4,816,567). Los "anticuerpos monoclonales" también se pueden aislar de bibliotecas de anticuerpo bacteriófago empleando las técnicas descritas en Clackson y colaboradores, Nature, 352:624-628 (1991) y Marks y colaboradores, J. Mol. Biol., 222:581-597 (1991), por ejemplo.

Los anticuerpos monoclonales de la presente incluyen anticuerpos "quiméricos" en los cuales una porción de la cadena pesada y/o liviana es idéntica u homóloga a las secuencias correspondientes en los anticuerpos derivados de una especie en particular, o perteneciendo a una clase o sub-clase de anticuerpo en particular, mientras que el resto de la(s) cadena(s) es (son) idéntica(s) u homóloga(s) a las secuencias correspondientes en los anticuerpos derivados de otras especies o perteneciendo a otra clase o sub-clase de anticuerpo, al igual que fragmentos de los citados anticuerpos, siempre que muestren la actividad biológica deseada (Patente de Estados Unidos No. 4,816,567; y Morrison y colaboradores, Proc. Natl. Acad. Sci USA 81:6851-6855 (1984)). Los anticuerpos quiméricos de interés de la presente comprenden anticuerpos "primatizados" que comprenden secuencias de unión de antígeno de dominio variable derivadas de un primate no humano (*por ejemplo* Mono del Viejo Mundo, Simio etc.) y secuencias humanas de región constante.

Un anticuerpo "intacto" es uno que comprende un punto de unión de antígeno al igual que un C_L y por lo menos dominios constantes de cadena pesada C_H1, C_H2 y C_H3. Los dominios constantes pueden ser dominios constantes de secuencia natural (por ejemplo dominios constantes de secuencia natural humanos) o variantes de secuencias de aminoácidos de los mismos. Preferiblemente, el anticuerpo intacto tiene una o varias funciones efectoras.

"Fragmentos de anticuerpos" comprende una porción de un anticuerpo intacto, preferiblemente la región de unión de antígeno o la región variable del anticuerpo intacto. Ejemplos de un fragmento de anticuerpo incluyen Fab, Fab', $F(ab')_2$ y fragmentos Fv; diacuerpos;

anticuerpos lineales (consultar Patente de Estados Unidos No. 5,641,870, Ejemplo 2; Zapata y colaboradores, Protein Eng. 8(10): 1057-1062 [1995]); moléculas de anticuerpos de una sola cadena; y anticuerpos multiespecíficos formados a partir de un fragmento de anticuerpos.

La digestión de papaína de los anticuerpos produce dos fragmentos idénticos de unión de antígeno denominados fragmentos "Fab" y un fragmento residual "Fc", una denominación que refleja la capacidad de fácil cristalización. El fragmento Fab consiste en una cadena L completa junto con la región de dominio variable de la cadena H (V_H) y el primer dominio constante de una cadena pesada (C_H1). Cada fragmento Fab es monovalente con respecto a la unión de antígeno, es decir que tiene un solo punto de unión de antígeno. El tratamiento de un anticuerpo con pepsina produce un solo fragmento grande $F(ab')_2$ que aproximadamente corresponde a dos fragmentos Fab enlazados por disúlfido teniendo una actividad divalente de unión de antígeno y aún es capaz de enlazar el antígeno en cruz. Los fragmentos Fab' difieren de los fragmentos Fab en que tienen unos pocos residuos adicionales en el terminal carboxi del dominio C_H1 incluyendo una o varias cisteínas de la región bisagra del anticuerpo. Fab'-SH es la denominación en la presente de Fab' donde el (los) residuo(s) de cisteína de los dominios constantes tienen un grupo tiol libre. Los fragmentos del anticuerpo $F(ab')_2$ originalmente se produjeron como pares de fragmentos Fab' que tienen unas cisteínas bisagra entre ellas. También se conocen otros acoplamientos químicos de fragmentos de anticuerpos.

El fragmento Fc comprende las porciones del terminal caboxi de las dos cadenas H unidas por disúlfidos. Las funciones efectoras de los anticuerpos se determinan por secuencias en la región Fc, cuya región es también parte de los receptores Fc reconocidos (FcR) que se han encontrado en ciertos tipos de células.

"Fv" es el fragmento mínimo de un anticuerpo que contiene un reconocimiento total de antígeno y un punto de unión. Este fragmento consiste en un dímero de una región de dominio variable con una cadena

pesada y una cadena liviana, una asociación no covalente. Del plegamiento de estos dos dominios emergen seis bucles hipervariables (3 bucles de la cadena H y 3 de la cadena L) que aportan los residuos de aminoácido para la unión del antígeno y confieren la especificidad de unión de antígeno al anticuerpo. Sin embargo, cada uno de los dominios variables (o la mitad de un Fv que comprende únicamente tres CDRs específicos para un antígeno) tiene la capacidad de reconocer y fijar el antígeno, aunque con una afinidad menor en la totalidad del punto de unión.

"Fv de una sola cadena" también abreviado como "sFv" o "scFv" son unos fragmentos de anticuerpo que comprenden los dominios del anticuerpo V_H y V_L conectados en una sola cadena de polipéptido. Preferiblemente, el sFv polipéptido comprende además un enlazante polipéptido entre los dominios V_H y V_L que permite que el sFv forme la estructura deseada para la unión del antígeno. Para revisar los sFv, consultar Pluckthun en The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenburg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994); Borrebaeck 1995, infra.

El término "diacuerpos" hace referencia a pequeños fragmentos de anticuerpos preparados al contribuir fragmentos sFv (ver el párrafo anterior) con enlazantes cortos (alrededor de 5-10 residuos) entre los dominios V_H y V_L de manera que se formen las parejas entre cadenas y no entre la misma cadena de los dominios V, resultando en un fragmento bivalente, es decir un fragmento teniendo dos puntos de unión de antígeno. Los diacuerpos biespecíficos con heterodímeros de dos fragmentos sFv "que se cruzan" donde los dominios V_H y V_L de los dos anticuerpos se encuentran presentes en dos cadenas de polipéptidos. Los diacuerpos se describen con mayor detalle por ejemplo en EP 404,097; WO 93/11161; y Hollinger y colaboradores, Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993).

Formas "humanizadas" de anticuerpos no humanos (*por ejemplo*, roedores) son anticuerpos quiméricos que contienen una secuencia mínima derivada del anticuerpo no humano. La mayoría de las veces, los anticuerpos humanizados son inmunoglobulinas humanas (anticuerpo

receptor) donde los residuos de una región hipervariable del receptor se reemplazan con residuos de una región hipervariable de una especie no humana (anticuerpo donante) como ratón, rata, conejo o primate no humano que tiene la especificidad, la afinidad y la capacidad deseadas. En algunos casos, los residuos de la región de soporte (FR) de la inmunoglobulina humana se reemplazan con los residuos no-humanos correspondientes. Adicionalmente, los anticuerpos humanizados pueden comprender residuos que no se encuentran en el anticuerpo receptor ni en el anticuerpo donante. Estas modificaciones se efectúan para refinar aún más el desempeño del anticuerpo. En general, el anticuerpo humanizado sustancialmente va a comprender todos de por lo menos uno y típicamente dos dominios variables, donde todos o sustancialmente todos los bucles hipervariables corresponden a aquellos de la inmunoglobulina no-humana y donde todos o sustancialmente todos los FRs son aquellos de una secuencia de inmunoglobulina humana. El anticuerpo humanizado opcionalmente también va a comprender por lo menos una porción de una inmunoglobulina de región constante (Fc), típicamente aquella de una inmunoglobulina humana. Para mayores detalles, consultar Jones y colaboradores, Nature 321:522-525 (1986); Riechmann y colaboradores, Nature 332:323-329 (1988); y Presta, Curr. Op. Struct. Biol. 2:593-596 (1992).

Un "anticuerpo dependiente de la especie" por ejemplo, un anticuerpo mamífero anti-humano IgE, es un anticuerpo que tiene una afinidad de fijación más fuerte para un antígeno de una primera especie de mamífero que la que tiene para un homólogo de ese antígeno de una segunda especie de mamífero. Normalmente el anticuerpo dependiente de la especie se "fija específicamente" a un antígeno humano (es decir que tiene un valor de afinidad de fijación (Kd) de no más de alrededor de 1×10^{-7} M, preferiblemente no más de alrededor de 1×10^{-8} y más preferiblemente no más de alrededor de 1×10^{-9} M) pero tiene una afinidad de fijación para un homólogo del antígeno de una segunda especie de mamífero no humano, por lo menos alrededor de 50 veces, o por lo menos alrededor de 500 veces,

o por lo menos alrededor de 1000 veces más débil que su afinidad de fijación para el antígeno humano. El anticuerpo dependiente de la especie puede ser cualquiera de los diferentes tipos de anticuerpos como se definen arriba, pero preferiblemente es un anticuerpo humanizado o humano.

Un "oligopéptido TAT de unión" es un oligopéptido que se fija, preferiblemente específicamente, con un polipéptido TAT como se describe en la presente. Los oligopéptidos TAT de unión se pueden sintetizar químicamente empleando una metodología conocida para la síntesis de oligopéptidos o se pueden preparar y purificar utilizando tecnología recombinante. Normalmente, los oligopéptidos TAT de unión tienen por lo menos alrededor de 5 aminoácidos de largo, alternativamente por lo menos alrededor de 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 o 100 aminoácidos de largo o más, donde esos oligopéptidos tienen la capacidad de fijarse, preferiblemente específicamente, con un polipéptido TAT como se describe en la presente. Los oligopéptidos TAT de unión se pueden identificar sin una experimentación indebida utilizando técnicas bien conocidas. Con respecto a esto, igualmente se nota que las técnicas para el cribado de las bibliotecas de oligopéptidos en busca de oligopéptidos que tienen la capacidad de fijarse específicamente a un polipéptido objetivo son bien conocidas en el arte (consultar por ejemplo, Patente de Estados Unidos Nos. 5,556,762, 5,750,373, 4,708,871, 4,833,092, 5,223,409, 5,403,484, 5,571,689, 5,663,143; PCT Publicación Nos. WO 84/03506 y WO84/03564; Geysen y colaboradores, Proc. Natl. Acad. Sci. U.S.A., 81:3998-4002 (1984); Geysen y colaboradores, Proc. Natl. Acad. Sci. U.S.A., 82:178-182 (1985); Geysen y colaboradores, en Synthetic Peptides as Antigens, 130-149 (1986); Geysen y colaboradores, J. Immunol. Meth., 102:259-274 (1987); Schoofs y colaboradores, J. Immunol., 140:611-616 (1988), CwMa, S. E. y colaboradores (1990) Proc. Natl. Acad. Sci. USA,

87:6378; Lowman, H.B. y colaboradores (1991) Biochemistry, 30:10832; Clackson, T. y colaboradores (1991) Nature, 352: 624; Marks, J. D. y colaboradores (1991), J. Mol. Biol., 222:581; Kang, A.S. y colaboradores (1991) Proc. Natl. Acad. Sci. USA, 88:8363 y Smith, G. P. (1991) Current Opin. Biotechnol., 2:668).

Una "molécula orgánica TAT de unión" es una molécula orgánica diferente a un oligopéptido o a un anticuerpo como se define en la presente, que se fija, preferiblemente específicamente, a un polipéptido TAT como se describe en la presente. Las moléculas orgánicas TAT de unión se pueden identificar y químicamente se pueden sintetizar empleando una metodología conocida (consultar por ejemplo, PCT Publicación Nos. WO00/00823 y WO00/39585). Normalmente, las moléculas orgánicas TAT de unión tienen un tamaño inferior a alrededor de 2000 daltons, alternativamente menos de alrededor de 1500, 750, 500, 250 o 200 daltons de tamaño, donde las citadas moléculas orgánicas que son capaces de fijarse, preferiblemente específicamente, con un polipéptido TAT como se describe en la presente, se pueden identificar sin una experimentación indebida utilizando técnicas muy conocidas. Con respecto a esto, igualmente se nota que las técnicas para el cribado de las bibliotecas de moléculas orgánicas en busca de moléculas que tienen la capacidad de fijarse a un polipéptido objetivo son bien conocidas en el arte (consultar por ejemplo, PCT Publicación Nos. WO00/00823 and WO00/39585).

Un "ARN de interferencia" o un "ARN pequeño de interferencia (ARNsi)" es una molécula bicatenaria de ARN con menos de 30 nucleótidos de largo que reduce la expresión de un gen objetivo. Un "ARN TAT de interferencia" o un "TAT ARNsi" se fija, preferiblemente específicamente, a un ácido nucleico TAT y reduce su expresión. Esto significa que la expresión de la molécula TAT es menor con el ARN de interferencia presente cuando se compara con la expresión de la molécula TAT en el control donde el ARN de interferencia no se encuentra presente. Los TAT ARN de interferencia se pueden identificar y sintetizar empleando

métodos conocidos (Shi Y. *Trends in Genetics* 19(1):9-12 (2003), WO/2003056012 y WO2003064621).

Un anticuerpo, un oligopéptido, ARNsi u otra molécula orgánica "que se unen" a un antígeno de interés, por ejemplo un antígeno polipéptido objetivo asociado a un tumor es uno que fija el antígeno con suficiente afinidad de manera que el anticuerpo, el oligopéptido u otra molécula orgánica es útil como agente diagnóstico y/o terapéutico para apuntar a una célula o a un tejido que expresa el antígeno y no reacciona en forma cruzada de manera significativa con otras proteínas. En estas incorporaciones, el grado de unión del anticuerpo, del oligopéptido, del ARNsi o de otra molécula orgánica a una proteína "no objetivo" será menor de alrededor de 10% de la unión del anticuerpo, oligopéptido, ARNsi u otra molécula orgánica a su proteína objetivo particular, como se determina por un separador de células afectadas por fluorescencia (FACS) o una radioinmunoprecipitación (RIA). Con respecto a la unión de un anticuerpo, de un oligopéptido u otra molécula orgánica a una molécula objetivo, el término "unión específica" o "específicamente se fija a" o es "especifico para" un polipéptido particular o un epítopo en un polipéptido particular objetivo, significa una unión mediblemente diferente de una interacción no específica. La unión específica se puede medir, por ejemplo, al determinar la fijación de una molécula comparada con la fijación de una molécula de control, que generalmente es una molécula con una estructura similar que no tiene una actividad de fijación. Por ejemplo, la unión específica se puede determinar mediante la competencia con una molécula de control que es similar al objetivo, por ejemplo, un exceso de objetivo sin marcación. En este caso, la unión específica se indica si la fijación del objetivo marcado a una sonda está competitivamente inhibida por un exceso de objetivo sin marcar. El término "unión específica" o "específicamente se fija a" o es "especifico para" un polipéptido particular o un epítopo en un objetivo polipéptido particular como se emplea en la presente se puede ver, por ejemplo, por una molécula teniendo un Kd para el objetivo de por lo menos alrededor de 10^{-4} M, alternativamente por lo menos alrededor de 10^{-5}

M, alternativamente por lo menos alrededor de 10^{-6} M, alternativamente por lo menos alrededor de 10^{-7} M, alternativamente por lo menos alrededor de 10^{-8} M, alternativamente por lo menos alrededor de 10^{-9} M, alternativamente por lo menos alrededor de 10^{-10} M, alternativamente por lo menos alrededor de 10^{-11} M, alternativamente por lo menos alrededor de 10^{-12} M o mayor. En una incorporación, el término "unión específica" hace referencia a la fijación donde una molécula se une a un polipéptido particular o a un epítopo en un polipéptido particular sin unirse sustancialmente a cualquier otro polipéptido o polipéptido epítopo.

Un anticuerpo, oligopéptido, ARNsi u otra molécula orgánica que "inhibe el crecimiento de células tumorales que expresan un polipéptido TAT" o un anticuerpo, un oligopéptido u otra molécula orgánica "que inhibe el crecimiento" es uno que resulta en una inhibición del crecimiento medible de las células cancerosas que expresa o sobreexpresa el polipéptido TAT apropiado. El polipéptido TAT puede ser un polipéptido transmembrana expresado en la superficie de una célula cancerosa o puede ser un polipéptido producido y secretado por una célula cancerosa. Los anticuerpos anti-TAT que inhiben el crecimiento, los oligopéptidos o las moléculas orgánicas preferidos inhiben el crecimiento de células tumorales que expresan el TAT en más de 20%, preferiblemente desde alrededor de 20% hasta alrededor de 50% y aún más preferiblemente en más de 50% (por ejemplo, desde alrededor de 50% hasta alrededor de 100%) cuando se compara con el control respectivo, típicamente el control siendo células tumorales no tratadas con el anticuerpo, oligopéptido u otra molécula orgánica bajo prueba. En una incorporación, la inhibición del crecimiento se puede medir a una concentración de anticuerpo de alrededor de 0.1 a 30 μg/ml o alrededor de 0.5 nM a 200 nM en un cultivo de células, donde la inhibición del crecimiento se determina 1-10 días después de exponer las células tumorales al anticuerpo. La inhibición del crecimiento de las células tumorales *in vivo* se puede determinar de varias maneras como las descritas en los Ejemplos Experimentales en la sección más abajo. El anticuerpo es un inhibidor de crecimiento *in vivo* si la administración del anticuerpo anti-

TAT a alrededor de 1 μg/kg hasta alrededor de 100 mg/kg por peso corporal resulta en la reducción del tamaño del tumor o de la proliferación de las células del tumor en alrededor de 5 días a 3 meses a partir de la primera administración del anticuerpo, preferiblemente en alrededor de 5 a 30 días.

Un anticuerpo, oligopéptido, ARNsi u otra molécula orgánica que "induce la apoptosis" induce la muerte celular programada como se determina por la fijación de anexina V, la fragmentación de ADN, el encogimiento de las células, la dilatación del retículo endoplásmico, la fragmentación celular y/o la formación de vesículas en la membrana (denominados cuerpos apoptóticos). La célula normalmente sobreexpresa un polipéptido TAT. Preferiblemente la célula es una célula de tumor, por ejemplo, una célula de próstata, mama, ovario, estómago, endometrio, pulmón, riñón, colon, colorrectal, vejiga. Hay varios métodos disponibles para evaluar los eventos celulares asociados con la apoptosis. Por ejemplo, la translocación de fosfatidil serina (PS) se puede medir mediante la fijación a anexina; la fragmentación de ADN se puede evaluar por medio de una escalera de ADN; y la condensación nuclear/cromatina junto con la fragmentación de ADN se pueden evaluar mediante cualquier incremento en las células hipodiploides. Preferiblemente, el anticuerpo, oligopéptido u otra molécula orgánica que induce la apoptosis es uno que resulta en alrededor de 2 a 50 veces, preferiblemente alrededor de 5 a 50 veces y más preferiblemente alrededor de 10 a 50 veces, la inducción de fijación de anexina con relación a una célula sin tratar en un análisis de fijación de anexina.

Las "funciones efectoras" del anticuerpo hacen referencia a aquellas funciones biológicas que se pueden atribuir a la región Fc (una región Fc de secuencia natural o una región Fc de variante de secuencia de aminoácidos) de un anticuerpo y varían con el isotipo del anticuerpo. Ejemplos de las funciones efectoras de los anticuerpos incluyen: fijación C1q y complementan la citotoxicidad dependiente; la fijación del receptor Fc; la citotoxicidad mediada por la célula dependiente de anticuerpo (ADCC); la

fagocitosis; la regulación por disminución de receptores en la superficie de la célula (por ejemplo, receptor célula B); y la activación de la célula B.

"Citotoxicidad celular dependiente de anticuerpo" o "ADCC" hace referencia a una forma de citotoxicidad donde el Ig secretado fijado en los receptores Fc (FcRs) presente en ciertas células citotóxicas (por ejemplo, los linfocitos citolíticos (NK) neutrófilos y macrófagos) permite que estas células citotóxicas efectoras se fijen específicamente a una célula objetivo que tiene el antígeno y posteriormente destruyan la célula objetivo con citotoxinas. Los anticuerpos "arman" las células citotóxicas y son absolutamente indispensables para esta destrucción. Las células primarias para mediar la ADCC, las células NK, únicamente expresan FcγRIII, mientras que los monolitos expresan FcγRI, FcγRII y FcγRIII. La expresión de FcR en células hematopoyéticas se resume en la Tabla 3 en la página 464 de Ravetch and Kinet, Annu. Rev. Immunol. 9:457-92 (1991). Para evaluar la actividad ADCC de una molécula de interés, se puede realizar un análisis *in vitro* de ADCC como el que se describe en la Patente de Estados Unidos No. 5,500,362 o 5,821,337. Unas células efectoras útiles para dichos análisis incluyen las células sanguíneas periféricas mononucleares (PBMC) y los linfocitos citolíticos (NK). Alternativamente, o adicionalmente, la actividad ADCC de la molécula de interés se puede evaluar *in vivo*, por ejemplo, en un modelo animal como el que se divulga en Clynes y colaboradores (USA) 95:652-656 (1998).

"Fc receptor" o "FcR" describe un receptor que se fija a la región Fc de un anticuerpo. El FcR preferido es una secuencia humana natural FcR. Por otra parte, un FcR preferido es uno que se fija a un anticuerpo IgG (un receptor gamma) e incluye receptores de las subclases FcγRI, FcγRII y FcγRIII, incluyendo variantes alélicas y alternativamente formas empalmadas de esos receptores. Los receptores FcγRII incluyen FcγRIIA (un "receptor activador") y FcγRIIB (un "receptor inhibidor"), que tienen secuencias similares de aminoácidos que difieren principalmente del dominio citoplásmico de los mismos. El receptor activador FcγRIIA contiene un motivo de activación inmunoreceptor a base de tirosina (ITAM)

en su dominio citoplásmico. El receptor inhibidor FcγRIIB contiene un motivo de activación inmunoreceptor a base de tirosina (ITIM) en su dominio citoplásmico, (consultar revisión M. in Daëron, Annu. Rev. Immunol. 15:203-234 (1997)). Los FcRs se revisan en Ravetch and Kinet, Annu. Rev. Immunol. 9:457-492 (1991); Capel y colaboradores, Immunomethods 4:25-34 (1994); y de Haas y colaboradores, J. Lab. Clin. Med. 126:330-41 (1995). Otros FcRs, incluyendo aquellos que se identificarán en el futuro, se comprenden en el término "FcR" de la presente. El término igualmente incluye el receptor neonatal FcRn, que es el responsable de la transferencia de los IgGs maternos al feto (Guyer y colaboradores, J. Immunol. 117:587 (1976) y Kim y colaboradores, J. Immunol. 24:249 (1994)).

"Células efectoras humanas" son leucocitos que expresan uno o varios FcRs y cumplen funciones efectoras. Preferiblemente, las células expresan por lo menos FcγRIII y cumplen la función efectora ADCC. Ejemplos de leucocitos humanos que median la ADCC incluyen células sanguíneas periféricas mononucleares (PBMC), linfocitos citolíticos (NK), monocitos, células T citotóxicas y neutrófilos; y se prefieren las PBMCs y las NK. Las células efectoras se pueden aislar de una fuente natural, por ejemplo, de la sangre.

"Citotoxicidad dependiente complementaria" o "CDC" hace referencia a la lisis de una célula objetivo en presencia de un complemento. La activación de la ruta complementaria clásica se inicia por la fijación del primer componente del sistema complementario (C1q) a los anticuerpos (de la subclase apropiada) que se unen a su antígeno afín. Para evaluar la activación complementaria, se puede realizar un análisis CDC, por ejemplo, como se describe en Gazzano-Santoro y colaboradores, J. Immunol. Methods 202:163 (1996).

Los términos "cáncer" y "canceroso" hacen referencia a o describen la condición fisiológica de los mamíferos que se caracteriza típicamente por un crecimiento celular desregulado. Ejemplos de cáncer incluyen, pero no se limitan a, carcinoma, linfoma, blastoma, sarcoma y leucemia o cáncer

linfático. Ejemplos más particulares de dichas neoplasias incluyen carcinoma celular escamoso (por ejemplo, carcinoma celular epitelial escamoso), cáncer de pulmón incluyendo carcinoma microcítico de pulmón, carcinoma no microcítico de pulmón, adenocarcinoma de pulmón y carcinoma escamoso de pulmón, cáncer de peritoneo, cáncer hepatocelular, cáncer gástrico o de estómago incluyendo cáncer gastrointestinal, cáncer pancreático, glioblastoma, cáncer cervical, cáncer de ovario, cáncer de hígado, cáncer de vejiga, cáncer del tracto urinario, hepatoma, cáncer de mama, cáncer de colon, cáncer rectal, cáncer colorrectal, carcinoma de endometrio o uterino, carcinoma de la glándula salival, cáncer de riñón o renal, cáncer de próstata, cáncer de vulva, cáncer de tiroides, carcinoma hepático, carcinoma anal, carcinoma peniano, melanoma, mieloma múltiple y linfoma de células B, cáncer de cerebro, cabeza y cuello y las metástasis asociadas.

Los términos "desorden celular proliferativo" y "desorden proliferativo" hacen referencia a los desórdenes asociados a cierto grado de proliferación celular anormal. En una incorporación, el desorden de proliferación celular es cáncer.

"Tumor", como se utiliza en la presente, hace referencia a todos los crecimientos celulares neoplásicos y a la proliferación, bien sea maligna o benigna, y a todas las células y los tejidos precancerosos y cancerosos.

Un anticuerpo, oligopéptido u otra molécula orgánica que "induce la muerte celular" es uno(a) que causa que una célula viable se vuelva inviable. La célula es una que expresa un polipéptido TAT, preferiblemente una célula que sobreexpresa un polipéptido TAT cuando se compara con una célula normal del mismo tipo de tejido. El polipéptido TAT puede ser un polipéptido transmembrana expresado en la superficie de una célula cancerosa o puede ser un polipéptido producido y secretado por una célula cancerosa. Preferiblemente, la célula es una célula cancerosa, por ejemplo, una célula de mama, ovario, estómago, endometrio, glándula salival, pulmón, riñón, colon, colorrectal, tiroides, pancreática o de vejiga. La muerte celular *in vitro* se puede determinar en la ausencia de células

efectoras de complemento e inmunes para distinguir la muerte celular inducida por la citotoxicidad celular dependiente de anticuerpo (ADCC) o la citoxicidad complemento dependiente (CDC). Por lo tanto, el análisis de la muerte celular se puede efectuar empleando un suero inactivado por calor (es decir en ausencia de complemento) y en ausencia de células efectoras inmunes. Para determinar si el anticuerpo, oligopéptido u otra molécula orgánica es capaz de inducir la muerte celular, se puede evaluar la pérdida de la integridad de la membrana mediante la absorción de yoduro de propidio (PI), triptano azul (consultar Moore y colaboradores Cytotechnology 17:1-11 (1995)) o 7AAD con respecto a células sin tratar. Los anticuerpos, oligopéptidos u otras moléculas orgánicas preferidos que inducen la muerte celular son los que inducen la absorción PI en el análisis de absorción PI en las células BT474.

Una "célula que expresa TAT" es una célula que expresa un polipéptido TAT endógeno o transinfectado bien sea en la superficie de la célula o en forma secretada. Un "TAT que expresa cáncer" es un cáncer que comprende células que tienen un polipéptido TAT presente en la superficie de la célula o que produce y secreta un polipéptido TAT. Un "TAT que expresa cáncer" opcionalmente produce suficientes niveles de ese polipéptido TAT en la superficie de las células del mismo, de manera que un anticuerpo anti-TAT, un oligopéptido u otra molécula orgánica se puede fijar al mismo y tener un efecto terapéutico con respecto al cáncer. En otra incorporación, un "TAT-que expresa cáncer" opcionalmente produce y secreta suficientes niveles de polipéptido TAT de manera que el anticuerpo anti-TAT, el oligopéptido u otra molécula orgánica antagonista se puede fijar al mismo y tener un efecto terapéutico con respecto al cáncer. Con respecto a esto último, el antagonista puede ser un oligonucleótido antisentido que reduce, inhibe o previene la producción y secreción del polipéptido TAT secretado por las células tumorales. Un cáncer que "sobreexpresa" un polipéptido TAT es uno que tiene unos niveles significativamente mayores de polipéptido TAT en la superficie de la célula del mismo, o produce y secreta, comparado con una célula no cancerosa del

mismo tejido. Esta sobreexpresión puede ser causada por la amplificación del gen o por una transcripción o traslación aumentada. La sobreexpresión de polipéptido TAT se puede determinar en un análisis diagnóstico o de pronóstico evaluando los niveles aumentados de la proteína TAT presente en la superficie de una célula, o secretada por la célula (por ejemplo, mediante un análisis de inmunohistoquímica empleando anticuerpos anti-TAT preparados contra un polipéptido TAT aislado que se pueden preparar empleando tecnología recombinante de ADN de un ácido nucleico aislado codificando el polipéptido TAT; análisis FACS etc.). Alternativamente, o adicionalmente, se pueden medir los niveles del polipéptido TAT codificando ácido nucleico o ARNm en la célula, por ejemplo, por medio de una hibridación fluorescente *in situ* empleando una sonda a base de ácido nucleico correspondiente a un ácido nucleico codificando TAT o el complemento del mismo (FISH; consultar WO98/45479 publicada en octubre, 1998), método Southern, método Northern o técnicas de reacción en cadena de polimerasa (PCR), como PCR cuantitativo en tiempo real (RT-PCR). Igualmente se puede estudiar esa sobreexpresión del polipéptido TAT midiendo el antígeno circulante en un fluido biológico como suero por ejemplo, utilizando un análisis a base de anticuerpo (consultar también, por ejemplo, Patente de Estados Unidos No. 4,933,294 otorgada el 12 de junio, 1990; WO91/05264 publicada el 18 de abril, 1991; Patente de Estados Unidos 5,401,otorgada el 18 de marzo de 1995; y Sias y colaboradores, J. Immunol. Methods 132:73-80 (1990)). Aparte de los análisis descritos arriba, existen varios análisis *in vivo* disponibles para el médico tratante con destrezas. Por ejemplo, se pueden exponer células en el cuerpo del paciente a un anticuerpo el cual está opcionalmente marcado con una marcación detectable, por ejemplo, un isótopo radioactivo y la fijación del anticuerpo a las células en el paciente se puede evaluar, por ejemplo, por medio de un escaneo externo en busca de radioactividad o analizando una biopsia tomada de un paciente previamente expuesto al anticuerpo.

Como se utiliza en la presente, el término "inmunoadhesina" denomina moléculas similares a anticuerpos que combinan la especificidad

de unión de una proteína heteróloga (una "adhesina") con las funciones efectoras de los dominios constantes de la inmunoglobulina. Estructuralmente, las inmunoadhesinas comprenden una fusión de una secuencia de aminoácidos con la especificidad de unión deseada, que es diferente al reconocimiento del antígeno y del punto de unión de un anticuerpo (es decir, es "heterólogo") y una secuencia de dominio constante de inmunoglobulina. La parte de adhesina de una molécula de inmunoadhesina típicamente es una secuencia contigua de aminoácidos que comprende por lo menos el punto de unión de un receptor o de un ligando. La secuencia de dominio constante de la inmunoglobulina en la inmunoadhesina se puede obtener a partir de cualquier inmunoglobulina, como de los subtipos IgG-1I, IgG-2, IgG-3, o IgG-4 e IgA (incluyendo IgA-1 eIgA-2), IgE, IgD o IgM.

La palabra "marca" cuando se emplea en la presente hace referencia a un compuesto o composición detectable que se conjuga directamente o indirectamente con el anticuerpo, oligopéptido u otra molécula orgánica para generar un anticuerpo, un oligopéptido u otra molécula orgánica "marcado/a". La marca puede ser detectable por sí misma (por ejemplo, las marcas radioisótopas o las marcas fluorescentes) o en el caso de una marca enzimática, puede catalizar la alteración química de un sustrato de compuesto o una composición que es detectable.

El término "agente citotóxico" como se emplea en la presente, hace referencia a una sustancia que inhibe o previene la función de las células y/o causa la destrucción de las células. El término incluye isótopos radioactivos (por ejemplo, At^{211}, I^{131}, I^{125}, Y^{90}, Re^{186}, Re^{188}, Sm^{153}, Bi^{212}, P^{32} e isótopos radioactivos de Lu), agentes quimioterapéuticos por ejemplo metotrexato, adriamicina, alcaloides vinca (vincristina, vinblastina, etoposida), doxorubicina, melfalano, mitomicina C, clorambucil, daunorubicina u otros agentes intercalantes, enzimas y fragmentos de los mismos como enzimas nucleótidas, antibióticos y toxinas como toxinas de molécula pequeña o toxinas enzimáticamente activas de origen bacteriano, fungal, de plantas o animales incluyendo fragmentos y/o variantes de los

mismos y los diferentes agentes antitumor o anti cancerígenos divulgados abajo. Otros agentes citotóxicos se describen abajo. Un agente tumoricida causa la destrucción de las células tumorales. Las citotoxinas se pueden adherir en forma covalente a un anticuerpo para apuntar la toxina hacia una célula en particular de interés que expresa el antígeno. Las citotoxinas útiles y su enlazante incluyen, sin limitación, los siguientes:

ENLAZANTES:

MC = maleimidocaproilo

Val Cit = valina-citrulina, sitio dipéptido en un enlazante segmentable de proteasa.

Citrulina = 2-amino-5-ureido ácido pentanoico

PAB = p-aminobencilcarbamoilo (porción de enlazante "autodestructivo")

Me = N-metil-valina citrulina donde el enlace peptídico enlazante se modificó para prevenir su segmentación por catepsina B

MC(PEG)6-OH = maleimidocaproil-polietilenglicol, adherido a las cisteínas del anticuerpo.

SPP = N-Succinimidil 4-(2-piridiltio) pentanoato

SMCC = N-Succinimidil 4-(N-maleimidometil) ciclohexano-1 carboxilato

FÁRMACOS CITOTÓXICOS

MMAE = mono-metil auristatina E (MW 718)

MMAF = variante de auristatina E (MMAE) con una fenilalanina en el terminal C del fármaco (MW 731.5)

AEVB = auristatina E valeril bencilhidrazona, enlazante ácido lábil a través del terminal C de AE (MW 732)

AFP =Auristatina F fenileno diamina; (la variante fenilalanina enlazada al anticuerpo a través del terminal C- mediante un espaciador fenileno diamina) (MW 732).

Un "agente inhibidor de crecimiento" cuando se emplea en la presente, hace referencia a un compuesto o composición que inhibe el crecimiento de una célula, especialmente un TAT que expresa una célula de cáncer, ya sea *in vitro* o *in vivo*. Por lo tanto, el agente inhibidor de crecimiento puede ser uno que reduce significativamente el porcentaje del

TAT que expresan las células en la fase S. Ejemplos de agentes inhibidores de crecimiento incluyen agentes que bloquean la progresión del ciclo celular (en un lugar diferente al de la fase S), como agentes que inducen la detención de la fase G1 y la detención de la fase M. Los bloqueadores clásicos de la fase M incluyen las vincas (vincristina y vinblastina), taxanos e inhibidores de topoisomerasa II como doxorubicina, epirubicina, daunorubicina, etoposida y bleomicina. Esos agentes que detienen el G1 también se trasladan a la detención de la fase S, por ejemplo, los agentes alquilantes de ADN como tamoxifén, prednisona, dacarbazina, mecloretamina, cisplatina, metotrexato, 5-fluorouracilo y ara-C. Se puede encontrar mayor información en The Molecular Basis of Cancer, Mendelsohn and Israel, eds., Capítulo 1, titulado "Cell cycle regulation, oncogenes, and antineoplastic drugs" por Murakami y colaboradores (WB Saunders: Philadelphia, 1995), especialmente p. 13. Los taxanos (paclitaxel y docetaxel) son fármacos anticáncer derivados ambos del árbol de tejo. Docetaxel (TAXOTERE®, Rhone-Poulenc Rorer), derivado del tejo europeo, es un análogo semisintético de paclitaxel (TAXOL®, Bristol-Myers Squibb). Paclitaxel y docetaxel promueven la formación de un conjunto de microtúbulos de dímeros de tubulina y estabilizan los microtúbulos evitando la despolimerización, que resulta en la inhibición de la mitosis en las células.

"Doxorubicina" es un antibiótico antraciclina. El nombre químico completo de la doxorubicina es (8S-cis)-10-[(3-amino-2,3,6-trideoxi-α-L-lixo-hexapiranosil)oxi]-7,8,9,10-tetrahidro-6,8,11-trihidroxi-8-(hidroxiacetil)-1-metoxi-5,12-naftacenediona.

El término "citocina" es un término genérico para las proteínas liberadas por una población de células que actúan sobre otra célula como mediadores intercelulares. Ejemplos de dichas citocinas son linfoquinas, monoquinas y las hormonas tradicionales de polipéptido. Incluidas en las citocinas se encuentran las hormonas de crecimiento como la hormona de crecimiento humana, la hormona de crecimiento humana N-metionilo y la hormona de crecimiento bovino; la hormona paratiroide; la tiroxina; la

insulina; la proinsulina; la relaxina; la prorelaxina; las hormonas glicoproteínas como la hormona folículo estimulante (FSH), la hormona tiroide estimulante (TSH) y la hormona leutinizante (LH); el factor de crecimiento hepático; el factor de crecimiento de fibroblasto; la prolactina; el lactógeno placental; el factor-α y -β de necrosis de tumor; sustancia inhibidora de Müller; péptido de ratón asociado a gonadotropina; inhibina; activina; factor de crecimiento endotelial vascular; integrina; trombopoyetina (TPO); factores de crecimiento de nervios como NGF-β; factor de crecimiento de trombocitos; factores de crecimiento transformantes (TGFs) como TGF-α y TGF-β; factor-I y -II de crecimiento similar a insulina; eritropoyetina (EPO); factores osteoinductivos; interferonas como interferon -α, -β y -γ; factores de estimulación de colonias (CSFs) como CSF-macrófagos (M- CSF); granulocito-macrófago-CSF (GM-CSF); y granulocito-CSF (G-CSF); interleucinas (ILs) como IL-1, IL-1a, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-11, IL-12; un factor de necrosis de tumor como TNF-α o TNF-β; y otros factores polipéptidos incluyendo LIF y kit ligando (KL). Como se utiliza en la presente, el término citocina incluye proteínas de fuentes naturales o de cultivos recombinantes de células y equivalentes biológicamente activos de las citocinas de secuencia natural.

El término "inserto en el paquete" se utiliza para hacer referencia a las instrucciones que se acostumbra incluir en los paquetes comerciales de productos terapéuticos que contiene información sobre las indicaciones, el uso, la dosis, la administración, las contraindicaciones y/o advertencias concernientes al uso de dichos productos terapéuticos.

Tabla 1

```
/*
 *
 * C-C increased from 12 to 15
 * Z is average of EQ
 * B is average of ND
 * match with stop is _M; stop-stop = 0; J (joker) match = 0
 */
#define _M      -8      /* value of a match with a stop */

int     _day[26][26] = {
/*    A B C D E F G H I J K L M N O P Q R S T U V W X Y Z */
/* A */   { 2, 0,-2, 0, 0,-4, 1,-1,-1, 0,-1,-2,-1, 0, _M, 1, 0,-2, 1, 1, 0, 0,-6, 0,-3, 0},
/* B */   { 0, 3,-4, 3, 2,-5, 0, 1,-2, 0, 0,-3,-2, 2, _M,-1, 1, 0, 0, 0, 0,-2,-5, 0,-3, 1},
/* C */   {-2,-4,15,-5,-5,-4,-3,-3,-2, 0,-5,-6,-5,-4, _M,-3,-5,-4, 0,-2, 0,-2,-8, 0, 0,-5},
/* D */   { 0, 3,-5, 4, 3,-6, 1, 1,-2, 0, 0,-4,-3, 2, _M,-1, 2,-1, 0, 0, 0,-2,-7, 0,-4, 2},
/* E */   { 0, 2,-5, 3, 4,-5, 0, 1,-2, 0, 0,-3,-2, 1, _M,-1, 2,-1, 0, 0, 0,-2,-7, 0,-4, 3},
/* F */   {-4,-5,-4,-6,-5, 9,-5,-2, 1, 0,-5, 2, 0,-4, _M,-5,-5,-4,-3,-3, 0,-1, 0, 0, 7,-5},
/* G */   { 1, 0,-3, 1, 0,-5, 5,-2,-3, 0,-2,-4,-3, 0, _M,-1,-1,-3, 1, 0, 0,-1,-7, 0,-5, 0},
/* H */   {-1, 1,-3, 1, 1,-2,-2, 6,-2, 0, 0,-2,-2, 2, _M, 0, 3, 2,-1,-1, 0,-2,-3, 0, 0, 2},
/* I */   {-1,-2,-2,-2,-2, 1,-3,-2, 5, 0,-2, 2, 2,-2, _M,-2,-2,-2,-1, 0, 0, 4,-5, 0,-1,-2},
/* J */   { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */   {-1, 0,-5, 0, 0,-5,-2, 0,-2, 0, 5,-3, 0, 1, _M,-1, 1, 3, 0, 0, 0,-2,-3, 0,-4, 0},
/* L */   {-2,-3,-6,-4,-3, 2,-4,-2, 2, 0,-3, 6, 4,-3, _M,-3,-2,-3,-3,-1, 0, 2,-2, 0,-1,-2},
/* M */   {-1,-2,-5,-3,-2, 0,-3,-2, 2, 0, 0, 4, 6,-2, _M,-2,-1, 0,-2,-1, 0, 2,-4, 0,-2,-1},
/* N */   { 0, 2,-4, 2, 1,-4, 0, 2,-2, 0, 1,-3,-2, 2, _M,-1, 1, 0, 1, 0, 0,-2,-4, 0,-2, 1},
/* O */   {_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M, 0, _M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M},
/* P */   { 1,-1,-3,-1,-1,-5,-1, 0,-2, 0,-1,-3,-2,-1, _M, 6, 0, 0, 1, 0, 0,-1,-6, 0,-5, 0},
/* Q */   { 0, 1,-5, 2, 2,-5,-1, 3,-2, 0, 1,-2,-1, 1, _M, 0, 4, 1,-1,-1, 0,-2,-5, 0,-4, 3},
/* R */   {-2, 0,-4,-1,-1,-4,-3, 2,-2, 0, 3,-3, 0, 0, _M, 0, 1, 6, 0,-1, 0,-2, 2, 0,-4, 0},
/* S */   { 1, 0, 0, 0, 0,-3, 1,-1,-1, 0, 0,-3,-2, 1, _M, 1,-1, 0, 2, 1, 0,-1,-2, 0,-3, 0},
/* T */   { 1, 0,-2, 0, 0,-3, 0,-1, 0, 0, 0,-1,-1, 0, _M, 0,-1,-1, 1, 3, 0, 0,-5, 0,-3, 0},
/* U */   { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */   { 0,-2,-2,-2,-2,-1,-1,-2, 4, 0,-2, 2, 2,-2, _M,-1,-2,-2,-1, 0, 0, 4,-6, 0,-2,-2},
/* W */   {-6,-5,-8,-7,-7, 0,-7,-3,-5, 0,-3,-2,-4,-4, _M,-6,-5, 2,-2,-5, 0,-6,17, 0, 0,-6},
/* X */   { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */   {-3,-3, 0,-4,-4, 7,-5, 0,-1, 0,-4,-1,-2,-2, _M,-5,-4,-4,-3,-3, 0,-2, 0, 0,10,-4},
/* Z */   { 0, 1,-5, 2, 3,-5, 0, 2,-2, 0, 0,-2,-1, 1, _M, 0, 3, 0, 0, 0, 0,-2,-6, 0,-4, 4}
};
```

Tabla 1 (cont.)

```
/*
 */
#include <stdio.h>
#include <ctype.h>

#define  MAXJMP      16      /* max jumps in a diag */
#define  MAXGAP      24      /* don't continue to penalize gaps larger than this */
#define  JMPS        1024    /* max jmps in an path */
#define  MX          4       /* save if there's at least MX-1 bases since last jmp */

#define  DMAT        3       /* value of matching bases */
#define  DMIS        0       /* penalty for mismatched bases */
#define  DINS0       8       /* penalty for a gap */
#define  DINS1       1       /* penalty per base */
#define  PINS0       8       /* penalty for a gap */
#define  PINS1       4       /* penalty per residue */

struct jmp {
        short           n[MAXJMP];      /* size of jmp (neg for dely) */
        unsigned short  x[MAXJMP];      /* base no. of jmp in seq x */
};                                      /* limits seq to 2^16 -1 */

struct diag {
        int             score;          /* score at last jmp */
        long            offset;         /* offset of prev block */
        short           ijmp;           /* current jmp index */
        struct jmp      jp;             /* list of jmps */
};

struct path {
        int     spc;            /* number of leading spaces */
        short   n[JMPS];/* size of jmp (gap) */
        int     x[JMPS];/* loc of jmp (last elem before gap) */
};

char            *ofile;                 /* output file name */
char            *namex[2];              /* seq names: getseqs() */
char            *prog;                  /* prog name for err msgs */
char            *seqx[2];               /* seqs: getseqs() */
int             dmax;                   /* best diag: nw() */
```

```
    int             dmax0;          /* final diag */
    int             dna;            /* set if dna: main() */
    int             endgaps;        /* set if penalizing end gaps */
    int             gapx, gapy;     /* total gaps in seqs */
    int             len0, len1;     /* seq lens */
    int             ngapx, ngapy;   /* total size of gaps */
    int             smax;           /* max score: nw() */
    int             *xbm;           /* bitmap for matching */
    long            offset;         /* current offset in jmp file */
    struct  diag    *dx;            /* holds diagonals */
    struct  path    pp[2];          /* holds path for seqs */

    char            *calloc(), *malloc(), *index(), *strcpy();
    char            *getseq(), *g_calloc();
```

Tabla 1 (cont.)

```
/* Needleman-Wunsch alignment program
 *
 * usage: progs file1 file2
 *  where file1 and file2 are two dna or two protein sequences.
 *  The sequences can be in upper- or lower-case an may contain ambiguity
 *  Any lines beginning with ';', '>' or '<' are ignored
 *  Max file length is 65535 (limited by unsigned short x in the jmp struct)
 *  A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
 *  Output is in the file "align.out"
 *
 * The program may create a tmp file in /tmp to hold info about traceback.
 * Original version developed under BSD 4.3 on a vax 8650
 */
#include "nw.h"
#include "day.h"

static  _dbval[26] = {
        1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};

static  _pbval[26] = {
        1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
        128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
        1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
        1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)                                                    main
        int     ac;
        char    *av[];
{
        prog = av[0];
        if (ac != 3) {
                fprintf(stderr,"usage: %s file1 file2\n", prog);
                fprintf(stderr,"where file1 and file2 are two dna or two protein sequences.\n");
                fprintf(stderr,"The sequences can be in upper- or lower-case\n");
                fprintf(stderr,"Any lines beginning with ';' or '<' are ignored\n");
                fprintf(stderr,"Output is in the file \"align.out\"\n");
                exit(1);
        }
```

```
        namex[0] = av[1];
        namex[1] = av[2];
        seqx[0] = getseq(namex[0], &len0);
        seqx[1] = getseq(namex[1], &len1);
        xbm = (dna)? _dbval : _pbval;

        endgaps = 0;                    /* 1 to penalize endgaps */
        ofile = "align.out";            /* output file */

        nw();           /* fill in the matrix, get the possible jmps */
        readjmps();     /* get the actual jmps */
        print();        /* print stats, alignment */

        cleanup(0);     /* unlink any tmp files */
}
```

Table 1 (cont.)

```
/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */
nw()                                                              nw
{
        char        *px, *py;     /* seqs and ptrs */
        int         *ndely, *dely;  /* keep track of dely */
        int         ndelx, delx;  /* keep track of delx */
        int         *tmp;         /* for swapping row0, row1 */
        int         mis;          /* score for each type */
        int         ins0, ins1;   /* insertion penalties */
        register    id;           /* diagonal index */
        register    ij;           /* jmp index */
        register    *col0, *col1; /* score for curr, last row */
        register    xx, yy;       /* index into seqs */

        dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

        ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
        dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
        col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
        col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
        ins0 = (dna)? DINS0 : PINS0;
        ins1 = (dna)? DINS1 : PINS1;

        smax = -10000;
        if (endgaps) {
                for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
                        col0[yy] = dely[yy] = col0[yy-1] - ins1;
                        ndely[yy] = yy;
                }
                col0[0] = 0;          /* Waterman Bull Math Biol 84 */
        }
        else
                for (yy = 1; yy <= len1; yy++)
                        dely[yy] = -ins0;
```

```
/* fill in match matrix
 */
for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialise first entry in col
         */
        if (endgaps) {
                if (xx == 1)
                        col1[0] = delx = -(ins0+ins1);
                else
                        col1[0] = delx = col0[0] - ins1;
                ndelx = xx;
        }
        else {
                col1[0] = 0;
                delx = -ins0;
                ndelx = 0;
        }
```

Table 1 (cont.)

```
                                                                    ...nW
for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
        mis = col0[yy-1];
        if (dna)
                mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
        else
                mis += _day[*px-'A'][*py-'A'];

        /* update penalty for del in x seq;
         * favor new del over ongong del
         * ignore MAXGAP if weighting endgaps
         */
        if (endgaps || ndely[yy] < MAXGAP) {
                if (col0[yy] - ins0 >= dely[yy]) {
                        dely[yy] = col0[yy] - (ins0+ins1);
                        ndely[yy] = 1;
                } else {
                        dely[yy] -= ins1;
                        ndely[yy]++;
                }
        } else {
                if (col0[yy] - (ins0+ins1) >= dely[yy]) {
                        dely[yy] = col0[yy] - (ins0+ins1);
                        ndely[yy] = 1;
                } else
                        ndely[yy]++;
        }

        /* update penalty for del in y seq;
         * favor new del over ongong del
         */
        if (endgaps || ndelx < MAXGAP) {
                if (col1[yy-1] - ins0 >= delx) {
                        delx = col1[yy-1] -(ins0+ins1);
                        ndelx = 1;
                } else {
                        delx -= ins1;
                        ndelx++;
                }
        } else {
```

```
                if (col1[yy-1] - (ins0+ins1) >= delx) {
                        delx = col1[yy-1] - (ins0+ins1);
                        ndelx = 1;
                } else
                        ndelx++;
        }

        /* pick the maximum score; we're favoring
         * mis over any del and delx over dely
         */
```

Table 1 (cont.)

```
                id = xx - yy + len1 - 1;
                if (mis >= delx && mis >= dely[yy])
                        col1[yy] = mis;
                else if (delx >= dely[yy]) {
                        col1[yy] = delx;
                        ij = dx[id].ijmp;
                        if (dx[id].jp.n[0] && (ldns || (ndelx >= MAXJMP
                        && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
                                dx[id].ijmp++;
                                if (++ij >= MAXJMP) {
                                        writejmps(id);
                                        ij = dx[id].ijmp = 0;
                                        dx[id].offset = offset;
                                        offset += sizeof(struct jmp) + sizeof(offset);
                                }
                        }
                        dx[id].jp.n[ij] = ndelx;
                        dx[id].jp.x[ij] = xx;
                        dx[id].score = delx;
                }
                else {
                        col1[yy] = dely[yy];
                        ij = dx[id].ijmp;
        if (dx[id].jp.n[0] && (ldns || (ndely[yy] >= MAXJMP
                        && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
                                dx[id].ijmp++;
                                if (++ij >= MAXJMP) {
                                        writejmps(id);
                                        ij = dx[id].ijmp = 0;
                                        dx[id].offset = offset;
                                        offset += sizeof(struct jmp) + sizeof(offset);
                                }
                        }
                        dx[id].jp.n[ij] = -ndely[yy];
                        dx[id].jp.x[ij] = xx;
                        dx[id].score = dely[yy];
                }
                if (xx == len0 && yy < len1) {
                        /* last col
```

```
                        */
                        if (endgaps)
                                col1[yy] -= ins0+ins1*(len1-yy);
                        if (col1[yy] > smax) {
                                smax = col1[yy];
                                dmax = id;
                        }
                }
        }
        if (endgaps && xx < len0)
                col1[yy-1] -= ins0+ins1*(len0-xx);
        if (col1[yy-1] > smax) {
                smax = col1[yy-1];
                dmax = id;
        }
        tmp = col0; col0 = col1; col1 = tmp;
}
(void) free((char *)ndely);
(void) free((char *)dely);
(void) free((char *)col0);
(void) free((char *)col1);                }
```

Table 1 (cont.)

```
/*
 *
 * print() -- only routine visible outside this module
 *
 * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[]: print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() -- put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */

#include "nw.h"

#define SPC     3
#define P_LINE  256     /* maximum output line */
#define P_SPC   3       /* space between name or num and seq */

extern  _day[26][26];
int     olen;           /* set output line length */
FILE    *fx;            /* output file */

print()                                                         print
{
        int     lx, ly, firstgap, lastgap;      /* overlap */

        if ((fx = fopen(ofile, "w")) == 0) {
                fprintf(stderr,"%s: can't write %s\n", prog, ofile);
                cleanup(1);
        }
        fprintf(fx, "<first sequence: %s (length = %d)\n", namex[0], len0);
        fprintf(fx, "<second sequence: %s (length = %d)\n", namex[1], len1);
        olen = 60;
        lx = len0;
        ly = len1;
        firstgap = lastgap = 0;
        if (dmax < len1 - 1) {          /* leading gap in x */
                pp[0].spc = firstgap = len1 - dmax - 1;
```

```
            ly -= pp[0].spc;
        }
        else if (dmax > len1 - 1) {     /* leading gap in y */
            pp[1].spc = firstgap = dmax - (len1 - 1);
            lx -= pp[1].spc;
        }
        if (dmax0 < len0 - 1) {         /* trailing gap in x */
            lastgap = len0 - dmax0 -1;
            lx -= lastgap;
        }
        else if (dmax0 > len0 - 1) {    /* trailing gap in y */
            lastgap = dmax0 - (len0 - 1);
            ly -= lastgap;
        }
        getmat(lx, ly, firstgap, lastgap);
        pr_align();
    }
```

Table 1 (cont.)

```
/*
 * trace back the best path, count matches
 */
static
getmat(lx, ly, firstgap, lastgap)                                    getmat
        int     lx, ly;                 /* "core" (minus endgaps) */
        int     firstgap, lastgap;      /* leading trailing overlap */
{
        int             nm, i0, i1, siz0, siz1;
        char            outx[32];
        double          pct;
        register        n0, n1;
        register char   *p0, *p1;

        /* get total matches, score
         */
        i0 = i1 = siz0 = siz1 = 0;
        p0 = seqx[0] + pp[1].spc;
        p1 = seqx[1] + pp[0].spc;
        n0 = pp[1].spc + 1;
        n1 = pp[0].spc + 1;

        nm = 0;
        while ( *p0 && *p1 ) {
                if (siz0) {
                        p1++;
                        n1++;
                        siz0--;
                }
                else if (siz1) {
                        p0++;
                        n0++;
                        siz1--;
                }
                else {
                        if (xbm[*p0-'A']&xbm[*p1-'A'])
                                nm++;
                        if (n0++ == pp[0].x[i0])
                                siz0 = pp[0].n[i0++];
                        if (n1++ == pp[1].x[i1])
```

```
                    siz1 = pp[1].n[l1++];
                p0++;
                p1++;
            }
        }

        /* pct homology:
         * if penalizing endgaps, base is the shorter seq
         * else, knock off overhangs and take shorter core
         */
        if (endgaps)
            lx = (len0 < len1)? len0 : len1;
        else
            lx = (lx < ly)? lx : ly;
        pct = 100.*(double)nm/(double)lx;
        fprintf(fx, "\n");
        fprintf(fx, "<%d match%s in an overlap of %d: %.2f percent similarity\n",
            nm, (nm == 1)? "" : "es", lx, pct);
```

Tabla 1 (cont.)

```
        fprintf(fx, "<gaps in first sequence: %d", gapx);                    ...getmat
        if (gapx) {
                (void) sprintf(outx, " (%d %s%s)",
                        ngapx, (dna)? "base":"residue", (ngapx == 1)? "":"s");
                fprintf(fx,"%s", outx);

        fprintf(fx, ", gaps in second sequence: %d", gapy);
        if (gapy) {
                (void) sprintf(outx, " (%d %s%s)",
                        ngapy, (dna)? "base":"residue", (ngapy == 1)? "":"s");
                fprintf(fx,"%s", outx);
        }
        if (dna)
                fprintf(fx,
                "\n<score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
                smax, DMAT, DMIS, DINS0, DINS1);
        else
                fprintf(fx,
                "\n<score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
                smax, PINS0, PINS1);
        if (endgaps)
                fprintf(fx,
                "<endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
                firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
                lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
        else
                fprintf(fx, "<endgaps not penalized\n");
}

static          nm;             /* matches in core -- for checking */
static          lmax;           /* lengths of stripped file names */
static          ij[2];          /* jmp index for a path */
static          nc[2];          /* number at start of current line */
static          ni[2];          /* current elem number -- for gapping */
static          siz[2];
static char     *ps[2];         /* ptr to current element */
static char     *po[2];         /* ptr to next output char slot */
static char     out[2][P_LINE]; /* output line */
static char     star[P_LINE];   /* set by stars() */
```

603

pr_align

```
/*
 * print alignment of described in struct path pp[]
 */
static
pr_align()
{
        int             nn;     /* char count */
        int             more;
        register        i;

        for (i = 0, lmax = 0; i < 2; i++) {
                nn = stripname(namex[i]);
                if (nn > lmax)
                        lmax = nn;

                nc[i] = 1;
                ni[i] = 1;
                siz[i] = dif[i] = 0;
                pe[i] = seqx[i];
                po[i] = out[i];                 }
```

Tabla 1 (cont.)

...pr_align

```
for (nn = nm = 0, more = 1; more; ) {
        for (i = more = 0; i < 2; i++) {
                /*
                 * do we have more of this sequence?
                 */
                if (!*ps[i])
                        continue;

                more++;

                if (pp[i].spc) {        /* leading space */
                        *po[i]++ = ' ';
                        pp[i].spc--;
                }
                else if (siz[i]) {      /* in a gap */
                        *po[i]++ = '-';
                        siz[i]--;
                }
                else {                  /* we're putting a seq element
                                         */
                        *po[i] = *ps[i];
                        if (islower(*ps[i]))
                                *ps[i] = toupper(*ps[i]);
                        po[i]++;
                        ps[i]++;

                        /*
                         * are we at next gap for this seq?
                         */
                        if (ni[i] == pp[i].x[ij[i]]) {
                                /*
                                 * we need to merge all gaps
                                 * at this location
                                 */
                                siz[i] = pp[i].n[ij[i]++];
                                while (ni[i] == pp[i].x[ij[i]])
                                        siz[i] += pp[i].n[ij[i]++];
                        }
                        ni[i]++;
```

```
                        }
                }
                if (++nn == olen || !more && nn) {
                        dumpblock();
                        for (i = 0; i < 2; i++)
                                po[i] = out[i];
                        nn = 0;
                }
        }
}

/*
 * dump a block of lines, including numbers, stars: pr_align()
 */
static
dumpblock()                                                    dumpblock
{
        register  i;

        for (i = 0; i < 2; i++)
                *po[i]-- = '\0';
```

Tabla 1 (cont.)

...dumpblock

```
        (void) putc('\n', fx);
        for (i = 0; i < 2; i++) {
                if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
                        if (i == 0)
                                nums(i);
                        if (i == 0 && *out[1])
                                stars();
                        putline(i);
                        if (i == 0 && *out[1])
                                fprintf(fx, star);
                        if (i == 1)
                                nums(i);
                }
        }

        /*
         * put out a number line: dumpblock()
         */
        static
        nums(ix)                                                    nnnnn
                int     ix;     /* index in out[] holding seq line */
        {
                char            nline[P_LINE];
                register        i, j;
                register char   *pn, *px, *py;

                for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
                        *pn = ' ';
                for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
                        if (*py == ' ' || *py == '-')
                                *pn = ' ';
                        else {
                                if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                                        j = (i < 0)? -i : i;
                                        for (px = pn; j; j /= 10, px--)
                                                *px = j%10 + '0';
```

```
                        if (i < 0)
                                *px = '-';
                }
                else
                        *px = ' ';
                i++;
        }
    }
    *pn = '\0';
    nc[lx] = i;
    for (pn = nline; *pn; pn++)
        (void) putc(*pn, fx);
    (void) putc('\n', fx);
}

/*
 * put out a line (name, [num], seq, [num]): dumpblock()
 */
static
putline(lx)
    int     lx;                 {
```

putline

Table 1 (cont.)

...putline

```
        int             i;
        register char   *px;

        for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
                (void) putc(*px, fx);
        for (; i < lmax+P_SPC; i++)
                (void) putc(' ', fx);

        /* these count from 1:
         * ni[] is current element (from 1)
         * nc[] is number at start of current line
         */
        for (px = out[ix]; *px; px++)
                (void) putc(*px&0x7F, fx);
        (void) putc('\n', fx);
}

/*
 * put a line of stars (segs always in out[0], out[1]): dumpblock()
 */
static
stars()                                                          stars
{
        int             i;
        register char   *p0, *p1, cx, *px;

        if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
            !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
                return;
        px = star;
        for (i = lmax+P_SPC; i; i--)
                *px++ = ' ';

        for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
                if (isalpha(*p0) && isalpha(*p1)) {
```

```
                if (xbm[*p0-'A']&xbm[*p1-'A']) {
                        cx = '*';
                        nm++;
                }
                else if (!dns && _day[*p0-'A'][*p1-'A'] > 0)
                        cx = '.';
                else
                        cx = ' ';
            }
            else
                cx = ' ';
            *px++ = cx;
        }
        *px++ = '\n';
        *px = '\0';
    }
```

Table 1 (cont.)

```
/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
stripname(pn)                                                    stripname
        char    *pn;    /* file name (may be path) */
{
        register char   *px, *py;

        py = 0;
        for (px = pn; *px; px++)
                if (*px == '/')
                        py = px + 1;
        if (py)
                (void) strcpy(pn, py);
        return(strlen(pn));

}
```

Table 1 (cont.)

```
/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_calloc() -- calloc() with error checkin
 * readjmps() -- get the good jmps, from tmp file if necessary
 * writejmps() -- write a filled array of jmps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>

char    *jname = "/tmp/homgXXXXXX";             /* tmp file for jmps */
FILE    *fj;

int     cleanup();                              /* cleanup tmp file */
long    lseek();

/*
 * remove any tmp file if we blow
 */
cleanup(i)                                                      cleanup
        int     i;
{
        if (fj)
                (void) unlink(jname);
        exit(i);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
char    *
getseq(file, len)                                               getseq
        char    *file;  /* file name */
        int     *len;   /* seq len */
{
        char            line[1024], *pseq;
        register char   *px, *py;
```

```
int             natgc, tlen;
FILE            *fp;

if ((fp = fopen(file,"r")) == 0) {
        fprintf(stderr,"%s: can't read %s\n", prog, file);
        exit(1);
}
tlen = natgc = 0;
while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
                continue;
        for (px = line; *px != '\n'; px++)
                if (isupper(*px) || islower(*px))
                        tlen++;
}
if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr,"%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
}
pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
```

Tabla 1 (cont.)

...getseq

```
        py = pseq + 4;
        *len = tlen;
        rewind(fp);

        while (fgets(line, 1024, fp)) {
                if (*line == ';' || *line == '<' || *line == '>')
                        continue;
                for (px = line; *px != '\n'; px++) {
                        if (isupper(*px))
                                *py++ = *px;
                        else if (islower(*px))
                                *py++ = toupper(*px);
                        if (index("ATGCU",*(py-1)))
                                natgc++;
                }
        }
        *py++ = '\0';
        *py = '\0';
        (void) fclose(fp);
        dna = natgc > (tlen/3);
        return(pseq+4);
}

char    *
g_calloc(msg, nx, sz)                                   g_calloc
        char    *msg;           /* program, calling routine */
        int     nx, sz;         /* number and size of elements */
{
        char            *px, *calloc();

        if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
                if (*msg) {
                        fprintf(stderr, "%s: g_calloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
                        exit(1);
                }
        }
        return(px);
}
```

```
/*
 * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
 */
readjmps()                                                    readjmps
{
        int             fd = -1;
        int             siz, i0, i1;
        register        i, j, xx;

        if (f) {
                (void) fclose(f);
                if ((fd = open(jname, O_RDONLY, 0)) < 0) {
                        fprintf(stderr, "%s: can't open() %s\n", prog, jname);
                        cleanup(1);
                }
        }
        for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {
                while (1) {
                        for (j = dx[dmax].jmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                                ;
```

Table 1 (cont.)

...readjmps

```
                if (j < 0 && dx[dmax].offset && fj) {
                        (void) lseek(fd, dx[dmax].offset, 0);
                        (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
                        (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
                        dx[dmax].jmp = MAXJMP-1;
                }
                else
                        break;
        }
        if (i >= JMPS) {
                fprintf(stderr, "%s: too many gaps in alignment\n", prog);
                cleanup(1);
        }
        if (j >= 0) {
                siz = dx[dmax].jp.n[j];
                xx = dx[dmax].jp.x[j];
                dmax += siz;
                if (siz < 0) {                  /* gap in second seq */
                        pp[1].n[i1] = -siz;
                        xx += siz;
                        /* id = xx - yy + len1 - 1
                         */
                        pp[1].x[i1] = xx - dmax + len1 - 1;
                        gapy++;
                        ngapy -= siz;
/* ignore MAXGAP when doing endgaps */
                        siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
                        i1++;
                }
                else if (siz > 0) {     /* gap in first seq */
                        pp[0].n[i0] = siz;
                        pp[0].x[i0] = xx;
                        gapx++;
                        ngapx += siz;
/* ignore MAXGAP when doing endgaps */
                        siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
                        i0++;
                }
        }
```

```
            else
                    break;
    }

    /* reverse the order of jmps
     */
    for (j = 0, i0--; j < i0; j++, i0--) {
            i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
            i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
    }
    for (j = 0, i1--; j < i1; j++, i1--) {
            i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
            i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
    }
    if (fd >= 0)
            (void) close(fd);
    if (f) {
            (void) unlink(name);
            f = 0;
            offset = 0;
    }                               }
```

Tabla 1 (cont.)

writejmps

```
/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
writejmps(lx)
        int     lx;
{
        char    *mktemp();

        if (!fj) {
                if (mktemp(jname) < 0) {
                        fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
                        cleanup(1);
                }
                if ((fj = fopen(jname, "w")) == 0) {
                        fprintf(stderr, "%s: can't write %s\n", prog, jname);
                        exit(1);
                }
        }
        (void) fwrite((char *)&dx[lx].jp, sizeof(struct jmp), 1, fj);
        (void) fwrite((char *)&dx[lx].offset, sizeof(dx[lx].offset), 1, fj);
}
```

Tabla 2

TAT XXXXXXXXXXXXXXX (Longitud = 15 aminoácidos)

Proteína de

Comparación XXXXXYYYYYYY (Longitud = 12 aminoácidos)

% de identidad de la secuencia de aminoácidos =

(el número de residuos de aminoácidos idénticos entre las dos secuencias de polipéptidos como se determina por ALIGN-2) dividido por (el número total de residuos de aminoácidos del polipéptido TAT) =

5 dividido por 15 = 33.3%

Tabla 3

TAT XXXXXXXXXX (Longitud = 10 aminoácidos)

Proteína de

Comparación XXXXXYYYYYYZZYZ (Longitud = 15 aminoácidos)

% de identidad de la secuencia de aminoácidos =

(el número de residuos de aminoácidos idénticos entre las dos secuencias de polipéptidos como se determina por ALIGN-2) dividido por (el número total de residuos de aminoácidos del polipéptido TAT) =

5 dividido por 10 = 50%

Tabla 4

TAT-ADN NNNNNNNNNNNNN (Longitud = 14 nucleótidos)

ADN de

Comparación NNNNNNLLLLLLLLL (Longitud = 16 nucleótidos)

% de identidad de la secuencia de ácido nucleico =

(el número de nucleótidos idénticos entre las dos secuencias de ácido nucleico como se determina por ALIGN-2) dividido por (el número total de nucleótidos de la secuencia TAT-ADN de ácido nucleico) =

6 dividido por 14 = 42.9%

Tabla 5

TAT-ADN	NNNNNTSINNNNNN	(Longitud = 12 nucleótidos)
ADN de Comparación	NNNNLLLVV	(Longitud = 9 nucleótidos)

% de identidad de la secuencia de ácido nucleico =

(el número de nucleótidos idénticos entre las dos secuencias de ácido nucleico como se determina por ALIGN-2) dividido por (el número total de nucleótidos de la secuencia TAT-ADN de ácido nucleico) =

4 dividido por 12 = 33.3%

II. Composiciones y Métodos de la Invención

A. Anticuerpos anti-TAT

En una incorporación, la presente invención dispone anticuerpos anti-TAT que pueden encontrar un uso en la presente como agentes terapéuticos y/o diagnósticos. Anticuerpos ejemplares incluyen anticuerpos policlonales, monoclonales, humanizados, biespecíficos y heteroconjugados.

1. Anticuerpos Policlonales

Los anticuerpos policlonales preferiblemente se cultivan en animales mediante múltiples inyecciones subcutáneas (sc) o intraperitoneales (ip) del antígeno respectivo y un aditivo. Puede ser útil conjugar el respectivo antígeno (especialmente cuando se utilizan péptidos sintéticos) a una

proteína que es inmunógena en las especies que se van a inmunizar. Por ejemplo, el antígeno se puede conjugar a una hemocianina de lapa californiana (KLH), seroalbúmina, tiroglobulina bovina o inhibidor de tripsina de la soya empleando un agente bifuncional o de derivación, por ejemplo, maleimidobenzoilo sulfosuccinimida éster (conjugación mediante residuos de cisteína), N-hidroxisuccinimida (mediante residuos de lisina), glutaraldehido, anhidruro succínico, $SOCl_2$ o $R^1N=C=NR$, donde R y R^1 son grupos alquilo diferentes.

Los animales se inmunizan contra el antígeno, los conjugados inmunógenos o sus derivados combinando por ejemplo, 100 µg o 5 µg de la proteína o el conjugado (para conejos o una ratone, respectivamente) con 3 volúmenes de aditivo completo de Freund e inyectando la solución de forma intradérmica en múltiples sitios. Un mes más tarde, los animales reciben un refuerzo de 1/5 a 1/10 de la cantidad original del péptido o conjugado en el aditivo completo de Freund mediante inyección subcutánea en múltiples sitios. De siete a 14 días más tarde, se extrae sangre de los animales y el suero se analiza en busca de títulos del anticuerpo. Los animales reciben un refuerzo hasta que los títulos se estabilizan. Igualmente los conjugados se pueden hacer en cultivos de células recombinantes como fusiones de proteínas. De igual manera se acostumbra apropiadamente agregar agentes como sulfato de aluminio para intensificar la respuesta inmune.

2. Anticuerpos Monoclonales

Se pueden hacer anticuerpos monoclonales utilizando el método hibridoma descrito primero por Kohler y colaboradores, Nature, 256:495 (1975) o se pueden hacer mediante métodos de ADN recombinantes (Patente de Estados Unidos No. 4,816,567).

En el método hibridoma, un ratón u otro animal hospedero apropiado, como un hámster, se inmuniza como se describe arriba para obtener linfocitos que producen o son capaces de producir anticuerpos que se van a unir específicamente a la proteína utilizada para la inmunización. Alternativamente, los linfocitos se pueden inmunizar *in vitro*. Después de

la inmunización, los linfocitos se aíslan y luego se fusionan con una estirpe celular de mieloma empleando un agente de fusión apropiado, como polietilenglicol, para formar una célula hibridoma (Goding, Monoclonal Antibodies: Principles and Practice, pp.59-103 (Academic Press, 1986)).

Las células hibridomas preparadas en esta forma se siembran y se cultivan en un medio de cultivo apropiado, medio que preferiblemente contiene una o varias sustancias que inhiben el crecimiento o la supervivencia de las células progenitoras de mieloma sin fusionar (igualmente se hace referencia a estos como compañeras de fusión). Por ejemplo, si las células progenitoras de mieloma carecen de la enzima hipoxantina guanina fosforibosil transferasa (HGPRT o HPRT), el medio de cultivo selectivo para los hibridomas típicamente va a incluir hipoxantina, aminopterina y timidina (medio HAT), sustancias que previenen el crecimiento de las células con deficiencia de HGPRT.

Las células de mieloma compañeras de fusión son las que se fusionan eficientemente, soportan un alto nivel de producción estable del anticuerpo a través de las células que producen el anticuerpo y son sensibles a un medio selectivo que efectúa una selección de células progenitoras sin fusionar. Las estirpes celulares de mieloma preferidas con las estirpes de mieloma murina como las que se derivan de tumores de ratón MOPC-21 y MPC-11 disponibles en el Salk Institute Cell Distribution Center, San Diego, California USA, y SP-2 y derivados, por ejemplo, células X63-Ag8-653 disponibles en la American Type Culture Collection, Manassas, Virginia, USA. Las estirpes celulares de mieloma humano y de heteromieloma también se han descrito para la producción de anticuerpos humanos monoclonales (Kozbor, J. Inmunol., 133:3001 (1984)y Brodeur y colaboradores, Monoclonal Antibody Production Techniques and Applications, pp. 51-63 (Marcel Dekker, Inc., New York, 1987)).

Se analiza el medio de cultivo donde las células hibridomas están creciendo para ver la producción de anticuerpos monoclonales dirigidos contra el antígeno. Preferiblemente, la especificidad de unión de los anticuerpos monoclonales producidos por las células hibridomas se

determina mediante la inmunoprecipitación o por medio de un análisis de fijación *in vitro*, como un radioinmunoanálisis (RIA) o un análisis inmunoabsorbente ligado a enzimas (ELISA).

La afinidad de fijación del anticuerpo monoclonal se puede determinar, por ejemplo, mediante el análisis Scatchard descrito en Munson y colaboradores, Anal. Biochem., 107:220 (1980).

Una vez se identifican las células hibridomas que producen anticuerpos de la especificidad, afinidad y/o actividad deseada, los clones se pueden subclonar al limitar los procedimientos de dilución mediante métodos estándar (Goding, Monoclonal Antibodies: Principles and Practice, pp.59-103 (Academic Press, 1986)). Unos medios de cultivo apropiados para este propósito incluyen por ejemplo, el medio D-MEM o RPMI-1640. Adicionalmente, las células hibridomas se pueden cultivar *in vivo* como tumores ascitis en un animal por ejemplo, por inyección i.p. de las células en ratones.

Los anticuerpos monoclonales secretados por los subclones se separan en forma apropiada del medio de cultivo, de los fluidos de ascitis o del suero mediante unos procedimientos convencionales de purificación de anticuerpos como por ejemplo, mediante cromatografía por afinidad (por ejemplo, utilizando proteína A o proteína G-Sefarosa) o por una cromatografía de intercambio de iones, cromatografía de hidroxilapatita, electroforesis en gel, diálisis, etc.

El ADN codificando los anticuerpos monoclonales es muy sencillo de aislar y secuenciar utilizando procedimientos convencionales (por ejemplo, empleando sondas de oligonucleótidos capaces de fijarse específicamente a los genes que codifican las cadenas pesadas y livianas de anticuerpos murinos). Las células hibridomas sirven como una fuente preferida de dicho ADN. Una vez aislado, el ADN se puede poner en vectores de expresión, que luego se transinfectan a células hospederas como las células *E. coli*, las células COS de simio, las células ováricas de hámster chino (CHO) o células de mieloma que de otra manera no producen la proteína anticuerpo, para obtener la síntesis de los anticuerpos monoclonales en las

células hospederas recombinantes. Artículos de revisión sobre la expresión de ADN recombinante en bacteria codificando el anticuerpo incluyen Skerra y colaboradores, Curr. Opinion in Inmunol., 5:256-262 (1993) y Plückthun, Inmunol. Revs. 130:151-188 (1992).

En otra incorporación, se pueden aislar los anticuerpos monoclonales o fragmentos de anticuerpos de las bibliotecas de anticuerpos fagos generados empleando las técnicas descritas en McCafferty y colaboradores, Nature, 348:552-554 (1990). Clackson y colaboradores, Nature, 352:624-628 (1991) y Marks y colaboradores, J. Mol. Biol., 222:581-597 (1991) describen el aislamiento de anticuerpos murinos y humanos, respectivamente, empleando bibliotecas fago. Publicaciones posteriores describen la producción de anticuerpos humanos de alta afinidad (rango nM) por barajado de cadena (Marks y colaboradores, Bio/Technology, 10:779-783 (1992)), al igual que la infección combinatoria y la recombinación *in vivo* como una estrategia para construir unas bibliotecas muy grandes de fagos (Waterhouse y colaboradores, Nuc. Acids. Res. 21:2265-2266 (1993)). Por lo tanto estas técnicas son alternativas viables para las técnicas convencionales para los anticuerpos monoclonales de hibridoma tradicionales para aislar anticuerpos monoclonales.

El ADN que codifica el anticuerpo se puede modificar para producir anticuerpos polipéptidos quiméricos o de fusión, por ejemplo, al sustituir las secuencias del dominio constante humano de la cadena pesada y de la cadena liviana (CH y C_L) por las secuencias murinas homólogas (Patente de Estados Unidos No. 4,816,567y Morrison, y colaboradores, Proc. Natl Acad. Sci. USA, 81:6851 (1984)) o fusionando la secuencia de codificación de la inmunoglobulina con toda o con parte de la secuencia de codificación de un polipéptido no-inmunoglobulina (polipéptido heterólogo). Las secuencias del polipéptido no-inmunoglobulina se pueden sustituir con los dominios constantes de un anticuerpo o se sustituyen con los dominios variables de un sitio de combinación de antígeno de un anticuerpo para crear un anticuerpo quimérico bivalente que comprende un sitio de combinación de antígeno teniendo una especificidad para un antígeno y otro

sitio de combinación de antígeno teniendo una especificidad para un antígeno diferente.

3. Anticuerpos Humanos y Humanizados

Los anticuerpos anti-TAT de la invención pueden comprender además anticuerpos humanizados o anticuerpos humanos. Formas humanizadas de anticuerpos no-humanos (por ejemplo, murinos) son inmunoglobulinas quiméricas, cadenas de inmunoglobulina o fragmentos de estas (como Fv, Fab, Fab', $F(ab')_2$ u otras subsecuencias de unión de antígeno de anticuerpos) que contienen una secuencia mínima derivada de una inmunoglobulina no-humana. Los anticuerpos humanizados incluyen inmunoglobulinas humanas (anticuerpo receptor) donde los residuos de una región de determinante complementario (CDR) del receptor se reemplazan con residuos de una CDR de especies no-humanas (anticuerpo donante) como un ratón, rata o conejo que tienen la especificidad, afinidad y capacidad deseadas. En algunas instancias, los residuos del soporte Fv de la inmunoglobulina humana se reemplazan con los residuos no humanos correspondientes. Los anticuerpos humanizados también pueden comprender residuos que no se encuentran ni en el anticuerpo receptor ni en las secuencias de las CDR o los soportes importados. En general, el anticuerpo humanizado va a comprender todos de por lo menos uno y típicamente dos dominios variables donde todas o sustancialmente todas las regiones CDR corresponden a las de la inmunoglobulina no-humana y todas o sustancialmente todas las regiones FR son aquellas de una secuencia consensual de inmunoglobulina humana. El anticuerpo humanizado de forma óptima también va a comprender por lo menos una porción de una región constante de inmunoglobulina (Fc), típicamente la de la inmunoglobulina humana [Jones y colaboradores, Nature, 321:522-525 (1986); Riechmann y colaboradores, Nature, 332:323- 329 (1988) y Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

Métodos para humanizar anticuerpos no humanos son bien conocidos en el arte. Generalmente, en un anticuerpo humanizado se ha introducido uno o varios residuos de aminoácidos provenientes de una fuente no

humana. Con frecuencia se hace referencia a estos residuos no humanos, como residuos "importados", que típicamente se toman de un dominio variable "de importación". Esencialmente la humanización se puede realizar siguiendo el método de Winter y colaboradores [Jones y colaboradores, Nature, 321:522-525 (1986); Riechmann y colaboradores, Nature, 332:323-327 (1988); Verhoeyen y colaboradores, Science, 239:1534-1536 (1988)], sustituyendo las CDRs o las secuencias de CDR de roedores por las secuencias correspondientes de un anticuerpo humano. Por consiguiente, dichos anticuerpos "humanizados" son anticuerpos quiméricos (Patente de Estados Unidos No. 4,816,567), donde sustancialmente menos de un dominio variable humano intacto ha sido sustituido por la secuencia correspondiente de una especie no humana. En la práctica, los anticuerpos humanizados típicamente son anticuerpos humanos donde algunos residuos de la CDR y posiblemente algunos residuos de la FR se sustituyen por residuos de sitios análogos en anticuerpos de roedores.

La elección de los dominios variables humanos, tanto livianos como pesados, que se van a emplear para fabricar los anticuerpos humanizados es muy importante para reducir la antígenoicidad y la respuesta HAMA (anticuerpo humano anti-ratón) cuando el anticuerpo se va emplear para un uso terapéutico en humanos. De acuerdo con el método denominado "de mejor ajuste" la secuencia del dominio variable de un anticuerpo de roedor se criba con toda la biblioteca de las secuencias humanas de dominio variable conocidas. Se identifica la secuencia humana de dominio V más cercana a la del roedor y la región de soporte humano (FR) en la misma aceptada para el anticuerpo humanizado (Sims y colaboradores, J. Inmunol. 151:2296 (1993); Chotia y colaboradores, J. Mol. Biol., 196:901 (1987)). Otro método emplea una región de soporte particular derivada de la secuencia consensual de todos los anticuerpos humanos de un subgrupo en particular de cadenas livianas o pesadas. Se puede emplear el mismo soporte para varios anticuerpos humanizados (Carter y colaboradores, Proc.

Natl. Acad. Sci. USA, 89:4285 (1992); Presta y colaboradores, J. Inmunol. 151:2623 (1993)).

Adicionalmente es importante que los anticuerpos se humanicen preservando la alta afinidad de unión para el antígeno y otras propiedades biológicas favorables. Para alcanzar este objetivo, de acuerdo con un método preferido, los anticuerpos humanizados se preparan mediante un proceso de análisis de las secuencias progenitoras y varios productos humanizados conceptuales empleando modelos tridimensionales de las secuencias progenitora y humanizada. Los modelos tridimensionales de inmunoglobulina están disponibles de manera común y aquellos con destreza en el arte están familiarizados con estos. Hay unos programas de computador disponibles que ilustran y presentan estructuras tridimensionales probables de conformación para unas secuencias de inmunoglobulina seleccionadas como candidatas. La inspección de estos muestrarios permite analizar el papel probable de los residuos en el funcionamiento de la secuencia de inmunoglobulina candidata, es decir, el análisis de residuos que tienen una influencia sobre la capacidad de la inmunoglobulina candidata para fijar se antígeno. En esta forma, los residuos FR de las secuencias del receptor y de importación se pueden seleccionar y combinar de manera que se pueden obtener las características del anticuerpo deseado, como una mayor afinidad para el(los) antígeno(s) objetivo. En general, los residuos de la región hipervariable están comprometidos de forma directa y sustancial para ejercer una influencia sobre la fijación del antígeno.

Se contemplan varias formas de un anticuerpo humanizado anti-TAT. Por ejemplo, el anticuerpo humanizado puede ser un fragmento de anticuerpo como un Fab, que opcionalmente se conjuga con uno o varios agentes citotóxicos con el fin de generar un inmunoconjugado. Alternativamente, el anticuerpo humanizado puede ser un anticuerpo intacto, como un anticuerpo IgG1 intacto. Como una alternativa a la humanización, se pueden generar anticuerpos humanos. Por ejemplo, ahora es posible producir animales transgénicos (por ejemplo, ratones) que

pueden, una vez inmunizados, producir todo un repertorio de anticuerpos humanos en ausencia de la producción endógena de inmunoglobulina. Por ejemplo, se ha descrito que la eliminación homocigótica del gen de la región de unión de la cadena pesada del anticuerpo (J_H) en ratones mutantes quiméricos y en la estirpe del microorganismo resulta en la inhibición total de la producción endógena de anticuerpos. La transferencia de la variedad de genes de la estirpe de microorganismos de inmunoglobulina humana en las citadas estirpes de microorganismos de los ratones mutantes va a resultar en la producción de anticuerpos humanos en el momento de la provocación del antígeno. Consultar por ejemplo, Jakobovits y colaboradores, Proc. Natl. Acad. Sci. USA. 90:2551 (1993); Jakobovits y colaboradores, Nature, 362:255-258 (1993); Bruggemann y colaboradores, Year in Immuno. 7:33 (1993); Patente de Estados Unidos Nos. 5,545,806, 5,569,825, 5,591,669 (todas de GenPharm); 5,545,807y WO 97/17852. Alternativamente, se puede utilizar la tecnología de muestrario fago (McCafferty y colaboradores, Nature 348:552-553 [1990]) para producir anticuerpos humanos y fragmentos de anticuerpos *in vitro*, a partir de unos repertorios de genes de dominio variable (V) de inmunoglobulina de donantes sin inmunizar. De acuerdo con esta técnica, los genes del dominio variable V de anticuerpo se clonan sin variación del marco de lectura en un gen de cápside mayor o menor de proteína de un bacteriófago filamentoso, como M13 o fd, y se representan como fragmentos funcionales de anticuerpos en la superficie de la partícula fago. Como la partícula filamentosa contiene una copia de una sola hebra de ADN del genoma fago, las selecciones basadas en las propiedades funcionales del anticuerpo también resultan en la selección del gen codificando el anticuerpo que exhibe esas propiedades. De este modo, el fago imita algunas de las propiedades de la célula B. El muestrario fago se puede completar en una variedad de formatos, revisados por ejemplo en Johnson, Kevin S. and Chiswell, David J., Current Opinion in Structural Biology 3:564-571 (1993). Se pueden emplear diversas fuentes de segmentos de gen V para el muestrario fago. Clackson y colaboradores, Nature. 352:624-628 (1991)

aislaron una gran variedad de anticuerpos anti-oxazolona de una pequeña biblioteca aleatoria de combinación de genes V derivados del bazo de ratones inmunizados. Se puede construir un repertorio de genes V a partir de donantes sin inmunizar y anticuerpos para una amplia variedad de antígenos (incluyendo auto antígenos) esencialmente siguiendo las técnicas de Marks y colaboradores, J. Mol. Biol. 222:581-597 (1991) o Griffith y colaboradores, EMBO J. 12:725-734 (1993). Consultar también las Patentes de Estados Unidos Nos. 5,565,332 y 5,573,905.

Como se discutió arriba, también se pueden generar los anticuerpos humanos mediante células B activadas in vitro (consultar Patentes de Estados Unidos 5,567,610 y 5,229,275).

4. Fragmentos de Anticuerpos

En determinadas circunstancias hay ventajas cuando se emplean fragmentos de anticuerpos, en vez de anticuerpos enteros. El tamaño más pequeño de los fragmentos permite una rápida depuración y puede permitir un mejor acceso a los tumores sólidos.

Se han desarrollado varias técnicas para la producción de fragmentos de anticuerpos. Tradicionalmente estos fragmentos se derivaban mediante la digestión proteolítica de anticuerpos intactos (consultar por ejemplo, Morimoto y colaboradores, Journal of Biochemistry and Biophysical Methods 24:107-117 (1992) y Brennan y colaboradores, Science. 229:81 (1985)). Sin embargo, ahora se pueden producir estos fragmentos directamente mediante células hospederas recombinantes. Todos los fragmentos de anticuerpos Fab, Fv y ScFv se pueden expresar en y secretar de *E. coli*, permitiendo de este modo una fácil producción de grandes cantidades de estos fragmentos. Se pueden aislar los fragmentos de anticuerpos de las bibliotecas anticuerpo fago descritas arriba. Alternativamente, los fragmentos Fab'-SH se pueden recuperar directamente de *E. coli* y se acoplan químicamente para formar fragmentos $F(ab')_2$ (Carter y colaboradores, Bio/Technology 10:163-167 (1992)). De acuerdo con otro enfoque, se pueden aislar fragmentos $F(ab')_2$ directamente del cultivo de células hospederas recombinantes. El fragmento Fab y $F(ab')_2$

con una vida media *in vivo* intensificada que comprende un receptor de reciclaje que se une a residuos de epítopo se describen en Patente de Estados Unidos No. 5,869,046. Otras técnicas para la producción de fragmentos de anticuerpos serán aparentes para el médico tratante con destrezas. En otras incorporaciones, el anticuerpo escogido es un fragmento Fv (scFv) de una sola cadena. Consultar WO 93/16185; Patente de Estados Unidos No. 5,571,894y Patente de Estados Unidos No. 5,587,458. Fv y sFv son las únicas especies con sitios de combinación intactos que están desprovistos de regiones constantes; de este modo, son apropiadas para una unión reducida no específica durante el uso *in vivo*. Las proteínas de fusión sFv se pueden construir para producir la fusión de una proteína efectora en el terminal amino o en el terminal carboxi de un sFv. Consultar Antibody Engineering, ed. Borrebaeck, supra. El fragmento de anticuerpo también puede ser un "anticuerpo linear", por ejemplo, como se describe en la Patente de Estados Unidos 5,641,870. Estos fragmentos de anticuerpos lineares pueden ser monoespecíficos o biespecíficos.

5. Anticuerpos Biespecíficos

Anticuerpos biespecíficos son anticuerpos que tienen unas especificidades de unión por lo menos para dos tipos diferentes de epítopos. Los anticuerpos biespecíficos ejemplares se pueden fijar a dos epítopos diferentes de una proteína TAT como se describe en la presente. Otros anticuerpos de este tipo pueden combinar un sitio de unión TAT con un sitio de unión para otra proteína. Alternativamente, un brazo anti TAT se puede combinar con un brazo que se fija a una molécula de activación en un leucocito como una molécula receptora de célula T (por ejemplo CD3) o receptores Fc para IgG (FcγR), como FcγRI (CD64), FcγRII (CD32) y FcγRIII (CD16), para centrarse y localizar los mecanismos de defensa celular a la célula que expresa el TAT. También se pueden utilizar anticuerpos biespecíficos para localizar agentes citotóxicos para las células que expresan el TAT. Estos anticuerpos poseen un brazo de fijación TAT y un brazo que fija el agente citotóxico (por ejemplo, saporina, anti-interferon-α, vinca alcaloide, cadena de ricina A, metotrexato o isótopo

radioactivo hapteno). Los anticuerpos biespecíficos se pueden preparar como anticuerpos de longitud total o como fragmentos de anticuerpos (por ejemplo, anticuerpos biespecíficos $F(ab')_2$).

WO 96/16673 describe un anticuerpo biespecífico anti-ErbB2/anti-FcγRIII y la Patente de Estados Unidos No. 5,837,234 divulga un anticuerpo biespecífico anti-ErbB2/anti-FcγRI. En WO98/02463 que muestra un anticuerpo biespecífico anti-ErbB2/Fcα. La Patente de Estados Unidos No. 5,821,337 enseña un anticuerpo biespecífico anti-ErbB2/anti-CD3.

Métodos para manufacturar anticuerpos biespecíficos son conocidos en el arte. La producción tradicional de anticuerpos biespecíficos de longitud total se basa en la coexpresión de dos pares de inmunoglobulinas de cadena liviana-cadena pesada, donde las dos cadenas tienen diferentes especificidades (Millstein y colaboradores, Nature 305:537-539 (1983)). Debido a la clasificación aleatoria de las cadenas pesadas y livianas de la inmunoglobulina, estos hibridomas (cuadromas) producen una mezcla potencial de 10 tipos diferentes de moléculas de anticuerpo, de las cuales solamente una tiene la estructura biespecífica correcta. La purificación de la molécula correcta, que normalmente se realiza mediante los pasos de la cromatografía por afinidad, es bastante difícil de manejar y el rendimiento de los productos es bajo. Se divulgan unos procedimientos similares en WO 93/08829 y en Traunecker y colaboradores, EMBO J. 10:3655-3659 (1991).

De acuerdo con un enfoque diferente, los dominios variables de anticuerpo con las especificidades de unión deseadas (sitios de combinación anticuerpo-antígeno) se fusionan a las secuencias de dominio constante de la inmunoglobulina. Preferiblemente, la fusión es con un dominio constante de cadena pesada Ig que comprende por lo menos parte de la bisagra, regiones CH2 y CH3. Se prefiere que una primera región constante de cadena pesada (CH1) contenga el sitio necesario para la unión de la cadena liviana presente por lo menos en una de las fusiones. Los ADNs codificando las fusiones de cadena pesada de la inmunoglobulina, y si se

CRF 631

desea, la cadena liviana de inmunoglobulina, se insertan en vectores de expresión separados y se cotransinfectan en una célula hospedera apropiada. Esto permite una mayor flexibilidad para ajustar las proporciones mutuas de los tres fragmentos polipéptidos en incorporaciones donde unas proporciones desiguales de las tres cadenas de polipéptidos empleadas en la construcción ofrecen una producción óptima del anticuerpo bioespecífico deseado. Sin embargo, es posible insertar las secuencias de codificación de dos o tres cadenas de polipéptidos en un solo vector de expresión cuando la expresión de por lo menos dos cadenas de polipéptido de iguales proporciones resulta en una alta producción o cuando las proporciones no tienen un efecto significativo sobre el rendimiento de la combinación deseada de la cadena.

En una incorporación preferida de este enfoque, los anticuerpos biespecíficos están compuestos por una inmunoglobulina híbrida de cadena pesada con una primera especificidad de unión en un brazo y una inmunoglobulina híbrida con una pareja de cadena pesada-cadena liviana (brindando una segunda especificidad de unión) en el otro brazo. Se encontró que esta estructura asimétrica facilita la separación del compuesto biespecífico deseado de inmunoglobulinas de combinaciones de cadenas indeseadas, puesto que la presencia de una inmunoglobulina de cadena liviana solamente en una mitad de la molécula biespecífica dispone una forma de separación fácil. Este enfoque se divulga en WO 94/04690. Para conocer mayores detalles sobre la generación de anticuerpos biespecíficos, consultar por ejemplo, Suresh y colaboradores, Methods in Enzymology 121:210 (1986).

De acuerdo con otro enfoque descrito en la Patente de Estados Unidos No. 5,731,168, se puede construir la interface entre una pareja de moléculas del anticuerpo para maximizar el porcentaje de heterodímeros que se recuperan del cultivo de células recombinantes. La interface preferida comprende por lo menos una parte del domino CH3. En este método, una o varias cadenas laterales pequeñas de aminoácidos de la interface de la primera molécula del anticuerpo se reemplaza con unas

633
632

cadenas de mayor tamaño (por ejemplo, tirosina o triptofano). Se crean unas "cavidades" compensatorias de tamaño idéntico o similar a la(s) cadena(s) lateral(es) más grande(s) en la interface de la segunda molécula del anticuerpo al reemplazar las grandes cadenas laterales de aminoácidos con unas más pequeñas (por ejemplo, alanina o treonina). Esto dispone un mecanismo para aumentar la producción del heterodímero frente a los productos finales indeseados como los homodímeros.

Los anticuerpos biespecíficos incluyen anticuerpos enlazados en cruz o "heteroconjugados". Por ejemplo uno de los anticuerpos en el heteroconjugado se puede acoplar a la avidita y el otro a la biotina. Por ejemplo se han propuesto dichos anticuerpos para apuntar a las células del sistema inmune de células indeseadas (Patente de Estados Unidos No. 4,676,980) y para el tratamiento de la infección VIH (WO 91/00360, WO 92/200373 y EP 03089). Los anticuerpos heteroconjugados se pueden hacer mediante el uso de cualquier método de enlace cruzado conveniente. Unos agentes de enlace cruzado apropiados son bien conocidos en el arte y se divulgan en la Patente de Estados Unidos No. 4,676,980, junto con un número de técnicas para el enlace cruzado.

Las técnicas para generar anticuerpos biespecíficos a partir de fragmentos de anticuerpos también se han descrito en la literatura. Por ejemplo, se pueden preparar anticuerpos biespecíficos empleando el ligamiento químico. Brennan y colaboradores, Science 229:81 (1985) describen un procedimiento en el cual unos anticuerpos intactos se segmentan en forma proteolítica para generar fragmentos $F(ab')_2$. Estos fragmentos se reducen en presencia del agente acomplejante ditiol, arsenita de sodio, para estabilizar los ditioles circundantes y para prevenir la formación de disúlfido intermolecular. Luego los fragmentos Fab' generados se convierten en derivados de tionitrobenzoato (TNB). Después uno de los derivados de Fab'-TNB se vuelve a convertir al Fab'-tiol mediante la reducción con mercaptoetilamina y se mezcla con una cantidad equimolar del otro derivado de Fab'-TNB para formar el anticuerpo

biespecífico. Los anticuerpos biespecíficos producidos se pueden utilizar como agentes para la inmovilización selectiva de las enzimas.

Los adelantos recientes han facilitado la recuperación directa de fragmentos de Fab'-SH, fragmentos de *E. coli*, que se pueden acoplar químicamente para formar anticuerpos biespecíficos. Shalaby y colaboradores, J. Exp. Med. 175: 217-225 (1992) describen la producción de una molécula de anticuerpo biespecífico totalmente humanizado $F(ab')_2$. Cada fragmento Fab' fue secretado en forma separada del *E. coli* y se sometió a un acoplamiento químico directo *in vitro* para formar el anticuerpo biespecífico. El anticuerpo biespecífico formado de este modo se pudo fijar a las células que sobreexpresan el receptor ErbB2 y las células humanas T normales, al igual que desencadenar la actividad de los linfocitos humanos citotóxicos contra blancos en tumores de mama humanos.

También se han descrito varias técnicas para hacer y aislar fragmentos de anticuerpos biespecíficos directamente de cultivos de células recombinantes. Por ejemplo, se han producido anticuerpos biespecíficos empleando cremalleras de leucina. Kostelny y colaboradores, J. Inmunol. 148(5): 1547-1553 (1992). Los péptidos de la cremallera de leucina de las proteínas Fos y Jun se enlazaron a las porciones Fab' de dos anticuerpos diferentes mediante la fusión de genes. Los homodímeros del anticuerpo se redujeron en la región bisagra para formar monómeros y luego se vuelven a oxidar para formar los heterodímeros del anticuerpo. Este método también se puede utilizar para la producción de homodímeros del anticuerpo. La tecnología de "diacuerpo" descrita por Hollinger y colaboradores, Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) ha suministrado un mecanismo alterno para fabricar fragmentos de anticuerpos biespecífico. Los fragmentos comprenden un VH conectado a un V_L mediante un enlazante demasiado corto para permitir la formación de parejas entre los dos dominios en la misma cadena. Por consiguiente, los dominios VH y V_L de un fragmento se ven forzados a formar pareja con los dominios V_L y VH complementarios de otro fragmento, formando de esta manera dos sitios de

fijación de antígeno. También se ha reportado otra estrategia para fabricar fragmentos de anticuerpos bioespecídicos mediante el uso de dímeros de una sola cadena Fv (sFv). Consultar Gruber y colaboradores, J. Inmunol., 152:5368 (1994).

Se contemplan anticuerpos con más de dos valencias. Por ejemplo se pueden preparar anticuerpos triespecíficos. Tutt y colaboradores, J. Inmunol. 147:60 (1991).

6. Anticuerpos Heteroconjugados

Los anticuerpos heteroconjugados también están dentro del ámbito de la presente invención. Los anticuerpos heteroconjugados están compuestos por dos anticuerpos unidos en forma covalente. Por ejemplo se han propuesto dichos anticuerpos para apuntar a las células del sistema inmune a las células indeseadas [Patente de Estados Unidos No. 4,676,980] y para el tratamiento de la infección VIH [WO 91/00360; WO 92/200373; EP 03089]. Se contempla que los anticuerpos se pueden preparar *in vitro* utilizando métodos conocidos en la química de proteínas sintéticas, incluyendo aquellas que comprenden agentes de enlace cruzado. Por ejemplo, se pueden construir inmunotoxinas empleando una reacción de intercambio de disúlfido o formando un enlace tioéter. Ejemplos de reactivos apropiados para este propósito incluyen iminotiolato y metil-4-mercaptobutirimidato y aquellos divulgados, por ejemplo, en la Patente de Estados Unidos No. 4,676,980.

7. Anticuerpos Multivalentes

Un anticuerpo multivalente se puede internalizar (y/o catabolizar) más rápido que un anticuerpo bivalente mediante una célula que expresa un antígeno al cual se fijan los anticuerpos. Los anticuerpos de la presente invención pueden ser anticuerpos multivalentes (diferentes a los de la clase IgM) con tres o más puntos de unión de antígeno (por ejemplo anticuerpos tetravalentes), que se pueden producir fácilmente mediante la expresión recombinante de ácido nucleico codificando cadenas polipéptido del anticuerpo. El anticuerpo multivalente puede comprender un dominio de dimerización y tres o más sitios de fijación de antígeno. El dominio de

dimerización preferido comprende (o consiste en) una región Fc o una región bisagra. En este escenario, el anticuerpo va a comprender una región Fc y tres o más sitios de fijación de antígeno terminal amino a la región Fc. El anticuerpo multivalente preferido en la presente comprende (o consiste en) tres hasta alrededor de ocho, pero preferiblemente cuatro, sitios de fijación de antígeno. El anticuerpo multivalente comprende por lo menos una cadena polipéptido (y preferiblemente dos cadenas polipéptido), donde la(s) cadena(s) polipéptido comprenden dos o más dominios variables. Por ejemplo la(s) cadena(s) polipéptido pueden comprender VD1-$(X1)_n$- VD2-$(X2)_n$-Fc, donde VD1 es un primer dominio variable, VD2 es un segundo dominio variable, Fc es una cadena polipéptido de una región Fc, X1 y X2 representan un aminoácido o polipéptido y n es 0 o 1. Por ejemplo, la(s) cadena(s) polipéptido pueden comprender: VH-CH1-enlazante flexible-VH-CH1-Fc cadena de región; o VH-CH1 -VH-CH1-Fc cadena de región. El anticuerpo multivalente de la presente preferiblemente comprende además por lo menos dos (y preferiblemente cuatro) polipéptidos de dominio variable de cadena liviana. El anticuerpo multivalente en la presente, puede, por ejemplo, comprender de alrededor de dos hasta alrededor de ocho polipéptidos de dominio variable de cadena liviana. Los polipéptidos de dominio variable de cadena liviana contemplados en la presente comprenden un dominio variable de cadena liviana y, opcionalmente comprenden además un dominio CL.

8. Construcción de la Función Efectora

Puede ser deseable modificar el anticuerpo de la invención con respecto a la función efectora, por ejemplo, con el fin de incrementar la citotoxicidad mediada por célula dependiente de antígeno (ADCC) y/o citotoxicidad complemento dependiente (CDC) del anticuerpo. Esto se puede lograr introduciendo una o varias sustituciones de aminoácido en una región Fc del anticuerpo. Alternativamente o adicionalmente, se pueden introducir residuo(s) de cisteína en la región Fc, permitiendo en esta forma la formación de un enlace disúlfido intercambiable en esta región. El anticuerpo homodimérico de este modo puede tener una mejor capacidad de

internalización y/o una mayor eliminación de células mediada por complemento y la citotoxicidad celular dependiente de anticuerpo (ADCC). Consultar Caron y colaboradores, J. Exp Med. 176:1191-1195 (1992) y Shopes, B. J. Inmunol. 148:2918-2922 (1992). Los anticuerpos homodiméricos con una mayor actividad antitumoral también se pueden preparar empleando enlazantes cruzados heterobifuncionales como se describe en Wolff y colaboradores, Cancer Research 53:2560-2565 (1993). Alternativamente, se puede construir un anticuerpo que tiene regiones Fc duales y por lo tanto tener una lisis de complemento mejorada y unas capacidades ADCC. Consultar Stevenson y colaboradores, Anti-Cancer Drug Design 3:219-230 (1989). Para aumentar la vida media del suero del anticuerpo, se puede incorporar un receptor de reciclaje que se une a los residuos de epítopo en el anticuerpo (especialmente un fragmento de anticuerpo) como se describe en Patente de Estados Unidos 5,739,277, por ejemplo. Como se utiliza en la presente, el término "receptor de reciclaje que se une al epítopo" hace referencia a un epítopo de la región Fc de una molécula IgG (por ejemplo, IgG_1, IgG_2, IgG_3 o IgG_4) que es la responsable de aumentar la vida media del suero *in vivo* de la molécula IgG.

9. Inmunoconjugados

La invención también se relaciona con inmunoconjugados que comprenden un anticuerpo conjugado a un agente citotóxico como un agente quimioterapéutico, un agente inhibidor de crecimiento, una toxina (*por ejemplo*, una toxina enzimáticamente activa de origen bacteriano, fungal, de planta o animal o fragmentos de los mismos) o un isótopo radioactivo (*es decir*, un radioconjugado).

Arriba se han descrito agentes quimioterapéuticos útiles en la generación de dichos inmunoconjugados. Las toxinas enzimáticamente activas y los fragmentos de las mismas que se pueden emplear incluyen una cadena difteria A, unos fragmentos activos no enlazantes de toxina de difteria, una cadena exotoxina A (de *Pseudomonas aeruginosa*), una cadena de ricina A, una cadena de abrina A, una cadena de modeccina A, una cadena de alfa-sarcina, proteínas *Aleurites fordii*, proteínas diantina,

proteínas Phytolaca americana (PAPI, PAPII y PAP-S), inhibidor momordica charantia, curcina, crotina, inhibidor sapaonaria officinalis, gelonina, mitogelina, restrictocina, fenomicina, enomicina y los tricotecenos. Existe una variedad de radionúclidos disponibles para la producción de anticuerpos radioconjugados. Los ejemplos incluyen ^{212}Bi, ^{131}I, ^{131}In, ^{90}Y y ^{186}Re. Los conjugados del anticuerpo y del agente citotóxico se hacen empleando una variedad de agentes de acoplamiento bifuncionales de proteína como N-succinimidil-3-(2-piridilditiol) propionato (SPDP), iminotiolano (IT), derivados bifuncionales de imidoésteres (como dimetil adipimidato HCL), ésteres activos (como disuccinimidil suberato), aldehidos (como glutaraldehido), compuestos bis-azida (como bis (p-azidobenzoil)hexanodiamina), derivados de bis-diazonio (como bis-(p-diazoniobenzoil)-etilendiamina), diisocianatos (como tolieno 2,6-diisocianato) y compuestos de flúor bis-activo (como 1,5-difluoro-2,4-dinitrobenceno). Por ejemplo se puede preparar una ricina inmunotoxina como se describe en Vitetta *y colaboradores*. Science, 238: 1098 (1987). Carbono-14-marcado l-isotiocianatobencil-3-metildietilen triaminapenta ácido acético (MX-DTPA) es un agente quelante ejemplar para la conjugación del radionucleótido al anticuerpo. Consultar WO94/11026.

En la presente también se contemplan los conjugados de un anticuerpo y una o varias toxinas de molécula pequeña, como una calicheamicina, maitansinoides, un tricoteno y CC1065 y los derivados de estas toxinas que tienen una actividad toxina.

Maitansina y maitansinoides

En una incorporación preferida, un anticuerpo anti-TAT (longitud total o fragmentos) de la invención se conjuga a una o varias moléculas es maitansinoides.

Los maitansinoides son inhibidores mitóticos que actúan inhibiendo la polimerización de tubulina. La maitansina se aisló por primera vez del arbusto *Maytenus serrata* del este de África (Patente de Estados Unidos No. 3,896,111). Posteriormente, se descubrió que ciertos microbios también producen maitansinoides, como maitanisol y C-3 ésteres de maitanisol

(Patente de Estados Unidos No. 4,151,042). El maitanisol sintético y los derivados y análogos del mismo se divulgan por ejemplo, en las Patentes de Estados Unidos Nos. 4,137,230; 4,248,870; 4,256,746; 4,260,608; 4,265,814; 4,294,757; 4,307,016; 4,308,268; 4,308,269; 4,309,428; 4,313,946; 4,315,929; 4,317,821; 4,322,348; 4,331,598; 4,361,650; 4,364,866; 4,424,219; 4,450,254; 4,362,663 y 4,371,533, cuyas divulgaciones se incorporan expresamente a la presente por referencia.

Conjugados Maitansinoide-anticuerpo

En una tentativa por mejorar su índice terapéutico, la maitansina y los maitansinoides se han conjugado a anticuerpos específicamente fijándose a los antígenos de células tumorales. Los inmunoconjugados que contienen maitansinoides y su uso terapéutico se divulgan por ejemplo, en las Patentes de Estados Unidos Nos. 5,208,020, 5,416,064 y en la Patente Europea EP 0 425 235 B1, cuyas divulgaciones se incorporan expresamente a la presente por referencia. Liu y colaboradores, Proc. Natl. Acad. Sci. USA 93:8618-8623 (1996) describieron inmunoconjugados que comprenden un maitansinoide designado DM1 enlazado al anticuerpo monoclonal C242 dirigido contra el cáncer colorrectal humano. Se encontró que el conjugado era altamente citotóxico frente a las células de cáncer de colon cultivadas y mostró una actividad antitumoral en un análisis de un crecimiento de tumor *in vivo*. Chari y colaboradores, Cancer Research 52:127-131 (1992) describe inmunoconjugados donde se conjugó un maitansinoide mediante un enlazante disúlfido al anticuerpo murino A7 fijándose a un antígeno en estirpes celulares de cáncer de colon humano o a otro anticuerpo monoclonal murino TA.1 que fija el oncógeno HER-2/*neu*. La citotoxicidad del conjugado maitansonoide TA.1 se analizó *in vitro* en la estirpe celular del cáncer de mama humano SK-BR-3, que expresa 3 x 10^5 HER-2 antígenos de superficie por célula. El fármaco conjugado alcanzó un grado de citotoxicidad similar al fármaco de maitansinoide libre, que se podía intensificar al incrementar el número de moléculas de maitansinoides por molécula de anticuerpo. El conjugado A7-maitansinoide mostró una baja citotoxicidad sistémica en ratones.

Conjugados polipéptido anticuerpo anti TAT-maitansinoide (inmunoconiugados)

Los conjugados polipéptido anticuerpo anti TAT-maitansinoide se preparan al enlazar un anticuerpo anti-TAT químicamente a una molécula maitansinoide sin reducir de forma significativa la actividad biológica ni del anticuerpo o ni de la molécula de maitansinoide. Un promedio de 3-4 moléculas de maitansinoide conjugadas por molécula de anticuerpo ha mostrado una eficacia al mejorar la citotoxicidad de las células objetivo sin afectar la función ni la solubilidad del anticuerpo en forma negativa, aunque se esperaría que una sola molécula de toxina/anticuerpo intensificara la citotoxicidad del uso de un anticuerpo desnudo. Los maitansinoides son bien conocidos en el arte y se pueden sintetizar mediante técnicas conocidas o se pueden aislar de fuentes naturales. Unos maitansinoides apropiados se divulgan por ejemplo, en la Patente de Estados Unidos No. 5,208,020 y en las demás patentes y publicaciones que no son patentes citadas arriba. Los maitansinoides preferidos son maitanisol y análogos de maitanisol modificados en el anillo aromático de la molécula de maitanisol, como varios ésteres de maitanisol.

Existen muchos grupos de enlace conocidos en el arte para preparar conjugados anticuerpo-maitansinoide incluyendo por ejemplo, aquellos descritos en la Patente de Estados Unidos No. 5,208,020 o en la Patente EP 0 425 235 B1 y Chari y colaboradores, Cancer Research 52:127-131 (1992). Los grupos de enlace incluyen grupos disúlfido, grupos tioéter, grupos ácido lábil, grupos fotolábil, grupos peptidasa lábil o grupos esterasa lábil, como se divulga en las patentes identificadas arriba, los grupos disúlfido y tioéter siendo los preferidos.

Los conjugados del anticuerpo y del maitansinoide se pueden fabricar empleando una variedad de agentes de acoplamiento bifuncionales de proteína como N-succinimidil-3-(2-piridilditio) propionato (SPDP), succinimidil-4-(N-maleimidometil) ciclohexano-1-carboxilato, iminotiolano (IT), derivados bifuncionales de imidoésteres (como dimetil adipimidato HCL), ésteres activos (como disuccinimidil suberato), aldehidos (como

glutareldehido), compuestos bis-azida (como bis (p-azidobenzoil) hexanodiamina), derivados de bis-diazonio (como bis-(p-diazoniobenzoil)-etilendiamina), diisocianatos (como tolueno 2,6-diisocianato) y compuestos de flúor bis-activo (como 1,5-difluoro-2,4-dinitrobenceno). Los agentes de acoplamiento particularmente preferidos incluyen N-succinimidil-3-(2-piridilditio) propionato (SPDP) (Carlsson y colaboradores, Biochem. J. 173:723-737 [1978]) y N-succinimidil-4-(2-piridiltio)pentanoato (SPP) para dispone un ligamiento disúlfido.

El enlazante se puede adherir a la molécula de maitansinoide en diversas posiciones, dependiendo del tipo de enlace. Por ejemplo, un ligamiento éster se puede formar mediante la reacción con un grupo hidroxilo utilizando unas técnicas de acoplamiento convencionales. La reacción puede ocurrir en la posición C-3 teniendo un grupo hidroxilo, la posición C-14 modificada con un hidroximetilo, la posición C-15 modificada con un grupo hidroxilo y la posición C-20 teniendo un grupo hidroxilo. En una incorporación preferida, el ligamiento se forma en la posición C-3 del maitanisol o de un análogo de maitanisol.

Calicheamicina

Otro inmunoconjugado de interés comprende un anticuerpo anti-TAT conjugado a una o varias moléculas de calicheamicina. La familia calicheamicina de antibióticos es capaz de producir rupturas bicatenarias de ADN a concentraciones subpicomolares. Para la preparación de conjugados de la familia calicheamicina consultar las Patentes de Estados Unidos 5,712,374, 5,714,586, 5,739,116, 5,767,285, 5,770,701, 5,770,710, 5,773,001, 5,877,296 (todas a American Cianamid Company). Los análogos estructurales de calicheamicina que se pueden utilizar incluyen, pero no se limitan a, γ_1^I, (α_2^I, α_3^I, N-acetil-γ_1^I PSAG y θ^I_1 (Hinman y colaboradores, Cancer Research 53:3336-3342 (1993), Lode y colaboradores, Cancer Research 58:2925-2928 (1998) y la patente de Estados Unidos citada arriba a American Cianamid). Otro fármaco antitumoral al cual se puede conjugar el anticuerpo es QFA que es un antifolato. Tanto la calicheamicina como el QFA tienen sitios de acción

intercelulares y no cruzan la membrana plasmática fácilmente. Por lo tanto, la absorción celular de estos agentes a través de la internalización mediada por el anticuerpo intensifica sus efectos citotóxicos en gran medida.

Otros agentes citotóxicos

Otros agentes antitumorales que se pueden conjugar a los anticuerpos anti-TAT de la invención incluyen BCNU, estreptozoicina, vincristina y 5-fluorouracilo, la familia de agentes conocida conjuntamente como el complejo LL-E33288 descrito en las Patentes de Estados Unidos 5,053,394, 5,770,710, al igual que las esperamicinas (Patente de Estados Unidos 5,877,296).

Las toxinas enzimáticamente activas y los fragmentos de las mismas que se pueden emplear incluyen una cadena difteria A, unos fragmentos activos no enlazantes de toxina de difteria, una cadena exotoxina A (de *Pseudomonas aeruginosa*), una cadena de ricina A, una cadena de abrina A, una cadena de modeccina A, una cadena de alfa-sarcina, proteínas *Aleurites fordii*, proteínas diantina, proteínas *Phytolaca americana* (PAPI, PAPII y PAP-S), inhibidor momordica charantia, curcina, crotina, inhibidor sapaonaria officinalis, gelonina, mitogelina, restrictocina, fenomicina, enomicina y los tricotecenos. Consultar por ejemplo WO 93/21232 publicada el 28 de octubre de 1993.

La presente invención contempla además un inmunoconjugado formado entre un anticuerpo y un compuesto con actividad nucleolítica (por ejemplo, una ribonucleasa o un ADN endonucleasa como una deoxiribonucleasa; DNase).

Para la destrucción selectiva del tumor, el anticuerpo puede comprender un átomo sumamente radioactivo. Existe una variedad de isótopos radioactivos para la producción de anticuerpos radioconjugados anti-TAT. Ejemplos incluyen At^{211}, I^{131}, I^{125}, Y^{90}, Re^{186}, Re^{188}, Sm^{153}, Bi^{212}, P^{32}, Pb^{212} e isótopos radioactivos de Lu. Cuando el conjugado se utiliza para el diagnóstico, puede comprender un átomo radioactivo para estudios gammagráficos, por ejemplo tc^{99m} o I^{123} o un marcador de rotación para imágenes de resonancia magnética (NMR) (también conocida como

imagenología por resonancia magnética, MRI), como yodo-123, yodo-131, indio-111, flúor-19, carbono-13, nitrógeno-15, oxígeno-17, gadolinio, manganeso o hierro.

Los radiomarcadores o los otros marcadores se pueden incorporar al conjugado en formas conocidas. Por ejemplo, el péptido se puede biosintetizar o se puede sintetizar mediante síntesis química de aminoácido empleando precursores apropiados de aminoácido comprendiendo por ejemplo, flúor-19 en lugar de hidrógeno. Marcadores como tc^{99m} o I^{123}, Re^{186}, Re^{188} e In^{111} se pueden adherir mediante un residuo de cisteína en el péptido. Se puede adherir Yttrium-90 mediante un residuo de lisina. El método IODOGEN (Fraker y colaboradores (1978) Biochem. Biophys. Res. Commun. 80: 49-57 se puede emplear para incorporar yodo-123. "Monoclonal Antibodies in Immunoscintigraphy" (Chatal,CRC Press 1989) describe otros métodos en detalle.

Los conjugados del anticuerpo y del agente citotóxico se hacen empleando una variedad de agentes de acoplamiento bifuncionales de proteína como N-succinimidil-3-(2-piridilditiol) propionato (SPDP), iminotiolano (IT), derivados bifuncionales de imidoésteres (como dimetil adipimidato HCL), ésteres activos (como disuccinimidil suberato), aldehidos (como glutareldehido), compuestos bis-azida (como bis (p-azidobenzoil)hexanodiamina), derivados de bis-diazonio (como bis-(p-diazoniobenzoil)-etilendiamina), diisocianatos (como tolieno 2,6-diisocianato) y compuestos de flúor bis-activo (como 1,5-difluoro-2,4-dinitrobenceno). Por ejemplo se puede preparar una ricina inmunotoxina como se describe en Vitetta *y colaboradores*. Science, 238: 1098 (1987). Carbono-14-marcado l-isotiocianatobencil-3-metildietilen triaminapenta ácido acético (MX-DTPA) es un agente quelante ejemplar para la conjugación del radionucleótido al anticuerpo. Consultar WO94/11026. El enlazante puede ser un "enlazante segmentable" que facilita la liberación del fármaco citotóxico en la célula. Por ejemplo, se puede utilizar un enlazante ácido lábil, un enlazante sensible a peptidasa, un enlazante fotolábil, un enlazante conteniendo dimetilo o disúlfido (Chari y

colaboradores, Cancer Research 52:127-131 (1992); Patente de Estados Unidos No. 5,208,020).

Alternativamente, se puede manufacturar una proteína de fusión comprendiendo el anticuerpo anti-TAT y el agente citotóxico, por ejemplo, mediante técnicas recombinantes o por la síntesis de péptidos. La longitud de ADN puede comprender las regiones respectivas codificando las dos porciones del conjugado bien sea adyacente la una a la otra, o separadas por una región codificando un péptido de enlace que no destruye las propiedades deseadas del conjugado.

En aún otra incorporación, el anticuerpo se puede conjugar a un "receptor" (como estreptavidina) para la utilización como blanco previo del tumor donde el conjugado anticuerpo-receptor se administra al paciente, seguido de la extracción del conjugado sin enlazar de la circulación empleando un agente de depuración y luego administrando un "ligando" (por ejemplo, avidina) que se conjuga a un agente citotóxico (por ejemplo, un radionucleótido).

10. Inmunoliposomas

Los anticuerpos anti-TAT divulgados en la presente también se pueden formular como inmunoliposomas. Un "liposoma" es una pequeña vesícula compuesta por varios tipos de lípidos, fosfolípidos y/o surfactantes que es útil para la administración de un fármaco a un mamífero. Los componentes del liposoma comúnmente están dispuestos en una formación de dos capas, similar a la distribución de lípidos en las membranas biológicas. Los liposomas que contienen el anticuerpo se preparan mediante métodos conocidos en el arte, como se describe en Epstein y colaboradores, Proc. Natl. Acad. Sci. USA 82:3688 (1985); Hwang y colaboradores, Proc. Natl Acad. Sci. USA 77:4030 (1980); Patentes de Estados Unidos Nos. 4,485,045 y 4,544,545y WO97/38731 publicada el 23 de octubre 1997. Los liposomas con un tiempo de circulación mejorado se divulgan en la Patente de Estados Unidos No. 5,013,556.

Los liposomas particularmente útiles se pueden generar mediante el método de evaporación de fase inversa con una composición de lípidos

comprendiendo fosfatidilcolina, colesterol y fosfatidiletanolamina derivadas de PEG (PEG-PE). Los liposomas se pasan por unos filtros con un tamaño de poro definido para obtener unos liposomas del diámetro deseado. Se pueden conjugar fragmentos Fab' del anticuerpo de la presente invención con los liposomas como se describe en Martin y colaboradores, J. Biol. Chem. 257:286-288 (1982) mediante una reacción de intercambio de disúlfido. Opcionalmente el liposoma contiene un agente quimioterapéutico. Consultar Gabizon y colaboradores, J. National Cancer Inst 81(19):1484 (1989).

B. Oligopéptidos TAT de unión

Los oligopéptidos TAT de unión de la presente invención son oligopéptidos que se fijan, preferiblemente específicamente, a un polipéptido TAT como se describe en la presente. Los oligopéptidos TAT de unión se pueden sintetizar químicamente utilizando una metodología conocida para la síntesis de oligopéptidos o se pueden preparar y purificar empleando tecnología recombinante. Los oligopéptidos TAT de unión tienen normalmente por lo menos alrededor de 5 aminoácidos de longitud, alternativamente por lo menos alrededor de 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 o 100 aminoácidos de longitud o más, donde dichos oligopéptidos que son capaces de fijarse, preferiblemente específicamente, a un polipéptido TAT como se describe en la presente. Los oligopéptidos TAT de unión se pueden identificar sin una experimentación inapropiada empleando técnicas muy conocidas. En este aspecto, se nota que las técnicas de cribado de bibliotecas oligopéptido en busca de oligopéptidos capaces de fijarse específicamente a un polipéptido blanco son bien conocidas en el arte (consultar por ejemplo, Patentes de Estados Unidos Nos. 5,556,762, 5,750,373, 4,708,871, 4,833,092, 5,223,409, 5,403,484, 5,571,689, 5,663,143; PCT Publicaciones Nos. WO 84/03506 y

WO84/03564; Geysen y colaboradores, Proc. Natl. Acad. Sci. U.S.A., 81:3998-4002 (1984); Geysen y colaboradores, Proc. Natl. Acad. Sci. U.S.A., 82:178-182 (1985); Geysen y colaboradores, en Synthetic Peptides as Antigens, 130-149 (1986); Geysen y colaboradores, J. Inmunol. Meth., 102:259-274 (1987); Schoofs y colaboradores, J. Inmunol, 140:611-616 (1988), Cwirla, S. E. y colaboradores (1990) Proc. Natl. Acad. Sci. USA, 87:6378; Lowman, H.B. y colaboradores (1991) Bioquímica, 30:10832; Clackson, T. y colaboradores (1991) Nature, 352: 624; Marks, J. D. y colaboradores (1991), J. Mol. Biol., 222:581; Kang, A.S. y colaboradores (1991) Proc. Natl. Acad. Sci. USA, 88:8363 y Smith, G. P. (1991) Current Opin. Biotechnol., 2:668).

En este aspecto, el muestrario bacteriófago (fago) es una técnica muy conocida que nos permite cribar grandes bibliotecas oligopéptido para identificar el (los) número(s) de esas bibliotecas que son capaces específicamente de fijarse a un polipéptido objetivo. El muestrario fago es una técnica mediante la cual los variantes de polipéptidos se representan como proteínas de fusión en la cubierta proteínica en la superficie de las partículas bacteriófagas (Scott, J.K. and Smith, G. P. (1990) Science 249: 386). La utilidad de los muestrarios fago yace en el hecho que las grandes bibliotecas de variantes de proteínas seleccionadas aleatoriamente (o ADNc clonados aleatoriamente) se pueden ordenar rápida y eficientemente para aquellas secuencias que se fijan a una molécula objetivo con una alta afinidad. El muestrario de bibliotecas péptido (Cwirla, S. E. y colaboradores (1990) Proc. Natl. Acad. Sci. USA, 87:6378) o de proteína (Lowman, H.B. y colaboradores (1991) Bioquímica, 30:10832; Clackson, T. y colaboradores (1991) Nature, 352: 624; Marks, J. D. y colaboradores (1991), J. Mol. Biol., 222:581; Kang, A.S. y colaboradores (1991) Proc. Natl. Acad. Sci. USA, 88:8363) en fago se ha utilizado para cribar millones de polipéptidos u oligopéptidos para encontrar aquellos con unas propiedades de unión específicas (Smith, G. P. (1991) Current Opin. Biotechnol., 2:668). Ordenar las bibliotecas fago de mutantes aleatorios requiere una estrategia para construir y propagar un gran número de

variantes, un procedimiento para la purificación por afinidad utilizando el receptor objetivo y un medio para evaluar los resultados del enriquecimiento de fijación. Patentes de Estados Unidos Nos. 5,223,409, 5,403,484, 5,571,689 y 5,663,143.

Aunque la mayoría de los métodos de muestrario fago han empleado la fase filamentosa, también se conocen sistemas de muestrario de fago lambdoid (WO 95/34683; U.S.A. 5,627,024), sistemas de muestrario fago T4 (Ren, Z-J. y colaboradores (1998) gen 215:439; Zhu, Z. (1997) CAN 33:534; Jiang, J. y colaboradores (1997) can 128:44380; Ren, Z-J. y colaboradores (1997) CAN 127:215644; Ren, Z-J. (1996) Protein Sci. 5:1833; Efimov, V. P. y colaboradores (1995) Virus Genes 10:173) y sistemas de muestrario fago T7 (Smith, G. P. and Scott, J.K. (1993) Methods in Enzymology, 217, 228-257; U.S.A. 5,766,905).

Ya se han desarrollado muchos otros adelantos y variaciones del concepto básico del muestrario fago. Estos adelantos intensifican la capacidad de los sistemas de muestrario para cribar las bibliotecas péptido para fijarse a las moléculas objetivo seleccionadas y para representar las proteínas funcionales con el potencial de cribar estas proteínas en busca de las propiedades deseadas. Se han desarrollado unos dispositivos de reacción combinatorios para las reacciones de muestrario fago (WO 98/14277) y se han empleado las bibliotecas de muestrario fago para analizar y controlar las interacciones bimoleculares (WO 98/20169; WO 98/20159) y las propiedades de péptidos helicales restringidos (WO 98/20036). WO 97/35196 describe un método para aislar un ligando de afinidad donde una biblioteca de muestrario fago se pone en contacto con una solución en la cual el ligando va a fijarse a una molécula objetivo y una segunda solución donde el ligando de afinidad no se va a fijar a la molécula objetivo, para aislar los ligandos de unión de forma selectiva. WO 97/46251 describe un método para realizar un biopaneo de una biblioteca aleatoria de muestrario fago con un anticuerpo de afinidad purificado y luego aislando el fago de unión, seguido por un proceso de micropaneo utilizando el fago de unión, seguido por un proceso de micropaneo

empleando unos pozos de microplacas para aislar el fago de unión de alta afinidad. También se ha reportado el uso de la proteína A de Staphylococcus aureus como marcador de afinidad (Li y colaboradores (1998) Mol Biotech., 9:187). WO 97/47314 describe el uso de bibliotecas de sustracción de sustrato para distinguir las especificidades enzimáticas utilizando una biblioteca de combinación que puede ser una biblioteca de muestrario fago. Un método para seleccionar las enzimas apropiadas para el uso en detergentes que emplean un muestrario fago se describe en WO 97/09446. Las Patentes de Estados Unidos Nos. 5,498,538, 5,432,018 y WO 98/15833 describen unos métodos adicionales para seleccionar las proteínas de fijación específica.

Los métodos para generar las bibliotecas péptido y para cribar esas bibliotecas se divulgan en las Patentes de Estados Unidos Nos. 5,723,286, 5,432,018, 5,580,717, 5,427,908, 5,498,530, 5,770,434, 5,734,018, 5,698,426, 5,763,192 y 5,723,323.

C. Moléculas orgánicas de unión TAT

Las moléculas orgánicas de unión TAT son moléculas orgánicas diferentes a los oligopéptidos o anticuerpos como se definen en la presente que se fijan, preferiblemente específicamente, a un polipéptido TAT como se describe en la presente. Las moléculas orgánicas de unión TAT se pueden identificar y sintetizar químicamente utilizando una metodología conocida (consultar por ejemplo, PCT Publicaciones Nos. WO00/00823 y WO00/39585). Normalmente las moléculas orgánicas de unión TAT tienen un tamaño inferior a alrededor de 2000 daltones, alternativamente un tamaño de menos de alrededor de 1500, 750, 500, 250 o 200 daltones, donde dichas moléculas orgánicas que son capaces de fijarse, preferiblemente específicamente, a un polipéptido TAT como se describe en la presente, se pueden identificar sin una experimentación indebida empleando unas técnicas muy conocidas. En este aspecto, se nota que las técnicas para cribar las bibliotecas de moléculas orgánicas que son capaces de fijarse a un polipéptido objetivo son bien conocidas en el arte (consultar por ejemplo, PCT Publicaciones Nos. WO00/00823 y WO00/39585). Las

649
648

moléculas orgánicas de unión TAT pueden ser, por ejemplo, aldehidos, cetonas, oximas, hidrazonas, semicarbazonas, carbazidas, aminas primarias, aminas secundarias, aminas terciarias, hidrazinas N-substituidas, hidrazidas, alcoholes, éteres, tioles, tioéteres, disúlfidos, ácidos carboxílicos, ésteres, amidas, ureas, carbamatos, carbonatos, cetales, tiocetales, acetales, tioacetales, aril halidos, aril sulfonatos, alquilo halidos, alquilo sulfonatos, compuestos aromáticos, compuestos heterocíclicos, anilinas, alquenos, alquinos, dioles, amino alcoholes, oxazolidinas, oxazolinas, tiazolidinas, tiazolinas, enaminas, sulfonamidas, epoxidas, aziridinas, isocianatos, sulfonil cloruros, compuestos diazo, cloruros de ácidos o similares.

D. Cribado de Anticuerpos anti-TAT, un oligopéptido de Unión TAT y Moléculas Orgánicas de Unión TAT con las Propiedades Deseadas

Las técnicas para generar anticuerpos, oligopéptidos y moléculas orgánicas que se fijan a polipéptidos TAT se han descrito arriba. Adicionalmente se pueden seleccionar anticuerpos, oligopéptidos u otras moléculas orgánicas con determinadas características biológicas, como se desee.

Los efectos de inhibición de crecimiento de un anticuerpo anti-TAT, un oligopéptido u otra molécula orgánica de la invención se puede evaluar mediante métodos conocidos en el arte, por ejemplo, empleando células que expresan un polipéptido TAT ya sea en forma endógena o después de la transinfección con el gen TAT. Por ejemplo, se pueden tratar unas estirpes celulares de tumor apropiadas y las células transinfectadas con TAT con un anticuerpo monoclonal anti-TAT, un oligopéptido u otra molécula orgánica de la invención a diferentes concentraciones durante unos días (por ejemplo, 2-7) días y se pueden teñir con violeta cristal o MTT o se pueden analizar mediante otro análisis colorimétrico. Otro método para medir la proliferación sería comparando la absorción de ^{3}H-timidina por parte de las células tratadas en presencia o en ausencia de un anticuerpo anti-TAT, un oligopéptido de unión TAT o una molécula orgánica de unión TAT de la invención. Después del tratamiento, las células se recolectan y la cantidad

de radioactividad incorporada en el ADN se cuantifica en un contador de centelleos. Unos controles positivos apropiados incluyen el tratamiento de una estirpe celular con un anticuerpo de inhibición de crecimiento que se sabe inhibe el crecimiento de esa estirpe celular. La inhibición de crecimiento de las células tumorales *in vivo* se puede determinar de varias formas conocidas en el arte. Preferiblemente, la célula tumoral es una que sobreexpresa un polipéptido TAT. Preferiblemente, el anticuerpo anti-TAT, un oligopéptido de unión TAT o la molécula orgánica de unión TAT va a inhibir la proliferación de una célula tumoral expresando el TAT *in vitro* o *in vivo* en alrededor de 25-100% comparado con la célula tumoral sin tratar, más preferiblemente, en alrededor de 30-100% y aún más preferiblemente en alrededor de 50-100% o 70-100%, en una incorporación, con una concentración de anticuerpo de alrededor de 0.5 a 30 μg/ml. La inhibición del crecimiento se puede medir con una concentración de anticuerpo de alrededor de 0.5 a 30 μg/ml o alrededor de 0.5 nM a 200 nM en un cultivo celular, donde la inhibición del crecimiento se determina 1-10 días después de exponer las células tumorales al anticuerpo. El anticuerpo es un inhibidor de crecimiento *in vivo* si la administración del anticuerpo anti-TAT a alrededor de 1 μg/kg hasta alrededor de 100 mg/kg por peso corporal resulta en la reducción del tamaño del tumor o de la proliferación de las células del tumor en alrededor de 5 días a 3 meses a partir de la primera administración del anticuerpo, preferiblemente en alrededor de 5 a 30 días.

Para seleccionar un anticuerpo anti-TAT, un oligopéptido de unión TAT o una molécula orgánica de unión TAT que induce la apoptosis, la pérdida de la integridad de la membrana como se indica por ejemplo, por la absorción de yoduro de propidio (PI), azul triptano o 7AAD se puede evaluar con respecto a un control. Un análisis de absorción de PI se puede realizar en ausencia de células de complemento y efectoras. Las células tumorales que expresan el polipéptido TAT se incuban en un medio solo o en un medio que contiene el anticuerpo anti-TAT apropiado (por ejemplo a alrededor de 10μg/ml), oligopéptido de unión TAT o una molécula orgánica

de unión TAT. Las células se incuban por un periodo de tiempo de 3 días. Después de cada tratamiento, las células se lavan y se distribuyen en porciones iguales en tubos de 35 mm con tapa-filtro de 12 x 75 (1ml por tubo, 3 tubos por grupo de tratamiento) para extraer las masas de células. Luego los tubos reciben PI (lOμg/ml). Las muestras se pueden analizar utilizando un citómetro de flujo FACSCAN® y el software FACSCONVERT® CellQuest (Becton Dickinson). Aquellos anticuerpos anti-TAT, un oligopéptido de unión TAT o las moléculas orgánicas de unión TAT que inducen niveles estadísticamente significativos de apoptosis como se determina por la absorción de PI se pueden seleccionar como anticuerpos anti-TAT, un oligopéptido de unión TAT o moléculas orgánicas de unión TAT que inducen la apoptosis.

Para cribar anticuerpos, oligopéptidos u otras moléculas orgánicas que se fijan a un epítopo en un polipéptido TAT unido por un anticuerpo de interés, se puede realizar un análisis rutinario de bloqueo cruzado como el que se describe en Antibodies, A Laboratory Manual. Cold Spring Harbor Laboratory, Ed Harlow y David Lane (1988). Este análisis se puede utilizar para determinar si un anticuerpo, un oligopéptido u otra molécula orgánica de prueba se fijan en el mismo sitio o epítopo que un anticuerpo anti-TAT conocido. Alternativamente o adicionalmente, se puede realizar un mapeo de epítopo mediante unos métodos conocidos en el arte. Por ejemplo, la secuencia del anticuerpo se puede mutagenizar como mediante escaneo con alanina, para identificar residuos de contacto. Inicialmente se prueba el anticuerpo mutante para ver la fijación con el anticuerpo monoclonal y asegurar el plegamiento adecuado. En un método diferente, se pueden emplear los péptidos que corresponden a diferentes regiones de un polipéptido TAT con los anticuerpos de prueba o con un anticuerpo de prueba y un anticuerpo con un epítopo caracterizado o conocido.

E. Terapia con Profármacos Mediados con Anticuerpos Dependiente de Enzima (ADEPT)

Los anticuerpos de la presente invención también se pueden emplear en ADEPT conjugando el anticuerpo con una enzima activada por

profármaco que convierte un profármaco (por ejemplo, un agente peptidilo quimioterapéutico, consultar WO81/01145) en un fármaco activo contra el cáncer. Consultar por ejemplo, WO 88/07378 y la Patente de Estados Unidos No. 4,975,278.

El componente enzima del inmunoconjugado útil para ADEPT incluye cualquier enzima capaz de actuar sobre un profármaco de tal manera que lo convierte a una forma más activa, citotóxica.

Enzimas útiles en el método de esta invención incluyen, pero no se limitan a, fosfatasa alcalina útil para convertir profármacos que contienen fosfato en fármacos libres; arilsulfatasa útil para convertir profármacos que contienen sulfato en fármacos libres; citosina deaminasa útil para convertir 5-fluorocitosina no tóxica en el fármaco anti-cáncer, 5-fluorouracil; proteasas, como serratia proteasa, termolisina, subtilisina, carboxipeptidasas y catepsinas (como catepsinas B y L), que son útiles para convertir profármacos conteniendo péptidos en fármacos libres; D-alanilcarboxipeptidasas, útiles para convertir profármacos que contienen radicales D-aminoácido; enzimas de segmentación de carbohidrato como β-galactosidasa y neuraminidasa útiles para convertir profármacos glicosilados en fármacos libres; β-lactamasa útil para convertir fármacos derivados con β-lactamos en fármacos libres y penicilina amidasas, como penicilina V amidasa o penicilina G amidasa, útil para convertir fármacos derivados en sus nitrógenos aminos con grupos fenoxiacetilo o fenilacetilo, respectivamente, en fármacos libres. Alternativamente, se pueden emplear anticuerpos con actividad enzimática, también conocidos en el arte como "abzimas", para convertir los profármacos de la invención en fármacos activos libres (consultar por ejemplo, Massey, Nature 328:457-458 (1987)). Se pueden preparar conjugados anticuerpo-abzima como se describe en la presente para administrar la abzima a una población de células tumorales.

Las enzimas de esta invención se pueden ligar de forma covalente a los anticuerpos anti-TAT mediante técnicas bien conocidas en el arte como el uso de reactivos heterobifuncionales de enlace cruzado descritos arriba. Alternativamente, las proteínas de fusión que comprenden por lo menos una

región de unión de antígeno de un anticuerpo de la invención enlazada por lo menos a una porción funcionalmente activa de una enzima de la invención se puede construir empleando técnicas de ADN recombinante bien conocidas en el arte (consultar por ejemplo, Neuberger y colaboradores, Nature 312:604-608 (1984).

F. Polipéptidos TAT de Longitud Total

La presente invención también dispone unas secuencias de nucleótido recientemente identificadas y aisladas que codifican los polipéptidos citados en la presente solicitud como polipéptidos TAT. En particular, se han identificado y aislado los ADNc (de longitud parcial y total) codificando varios polipéptidos TAT, como se divulga con mayor detalle en los Ejemplos abajo.

Como se divulga en los Ejemplos abajo, se han depositado varios clones de ADNc en la ATCC. Las secuencias reales de nucleótido de esos clones fácilmente se pueden determinar por parte de un artesano con destrezas, secuenciando el clon depositado empleando métodos rutinarios en el arte. La secuencia de aminoácidos esperada se puede determinar a partir de la secuencia de nucleótidos empleando una destreza rutinaria. Para los polipéptidos TAT y para codificar los ácidos nucleicos descritos en la presente, en algunos casos, los Solicitantes han identificado lo que se cree es el marco de lectura que mejor se identifica con la información de secuencia disponible en el momento.

G. Variantes de Anticuerpo Anti-TAT y de Polipéptido TAT

Además de los anticuerpos anti-TAT y de los polipéptidos TAT de secuencia natural de longitud total descritos en la presente, se contempla que se pueden preparar variantes de anticuerpo anti TAT y de polipéptido TAT. Se pueden preparar variantes de anticuerpo anti TAT y de polipéptido TAT introduciendo los cambios de nucleótido apropiados en el ADN de codificación y/o mediante la síntesis del anticuerpo o polipéptido deseado. Aquellos con destreza en el arte apreciarán que los cambios en el aminoácido pueden alterar los procesos post traslacionales del anticuerpo anti TAT o del polipéptido TAT, como cambiando el número o la posición

de los sitios de glicosilación o alterando las características de anclaje de la membrana.

Las variaciones en los anticuerpos anti-TAT y en los polipéptidos TAT descritos en la presente, pueden ser por ejemplo, emplear cualquiera de las técnicas y pautas para mutaciones conservadoras y no conservadoras establecidas, por ejemplo en la Patente de Estados Unidos No. 5,364,934. Las variaciones pueden ser una sustitución, eliminación o inserción de uno o varios codones codificando el anticuerpo o el polipéptido que resultan en un cambio en la secuencia de aminoácidos cuando se comparan con un anticuerpo o polipéptido de secuencia natural. Opcionalmente, la variación es mediante la sustitución de por lo menos un aminoácido con cualquier otro aminoácido en uno o varios de los dominios del anticuerpo anti TAT o del polipéptido TAT. Se pueden encontrar las guías para determinar cual residuo de aminoácidos se puede insertar, sustituir o eliminar sin afectar de forma adversa la actividad deseada comparando la secuencia del anticuerpo anti TAT o del polipéptido TAT con las moléculas homólogas conocidas de la proteína, y minimizando el número de cambios en la secuencia de aminoácidos efectuados en las regiones de alta homología. Las sustituciones de aminoácidos pueden ser el resultado de reemplazar un aminoácido con otro aminoácido teniendo unas propiedades estructurales y/o químicas similares, como el reemplazo de una leucina con una serina, es decir, reemplazos conservadores de aminoácidos. Las inserciones o eliminaciones pueden estar opcionalmente en el rango entre 1 a 5 aminoácidos. La variación permitida se puede determinar al efectuar inserciones, eliminaciones o sustituciones de aminoácidos en la secuencia y realizando pruebas con los variantes resultantes para confirmar la actividad exhibida por la secuencia de longitud completa o madura.

En la presente se disponen los fragmentos de anticuerpo anti TAT y de polipéptido TAT. Dichos fragmentos se pueden truncar en el terminal N o en el terminal C o pueden carecer de residuos, por ejemplo cuando se comparan con un anticuerpo o una proteína natural de longitud total. Ciertos fragmentos carecen de residuos de aminoácidos que no son

esenciales para una actividad biológica deseada del anticuerpo anti TAT o del polipéptido TAT.

Se pueden preparar fragmentos de anticuerpo anti TAT y de polipéptido TAT mediante cualquier cantidad de técnicas convencionales. Los fragmentos péptidos deseados se pueden sintetizar químicamente. Un enfoque alternativo comprende generar unos fragmentos de anticuerpo o polipéptido mediante digestión enzimática, por ejemplo, tratando la proteína con una enzima que se sabe segmenta las proteínas en sitios definidos por unos residuos particulares de aminoácidos o digiriendo el ADN con las enzimas de restricción apropiadas y aislando el fragmento deseado. Aún otra técnica apropiada comprende aislar y amplificar un fragmento de ADN codificando un fragmento anticuerpo o de polipéptido deseado, mediante una reacción en cadena de polimerasa (PCR). Los oligonucleótidos que definen los terminales deseados del fragmento de ADN se emplean en los iniciadores 5' 3' en la PCR. Preferiblemente, los fragmentos anticuerpo anti TAT y polipéptido TAT comparten por lo menos una actividad biológica y/o inmunológica con el anticuerpo anti TAT o polipéptido TAT natural divulgados en la presente.

En incorporaciones particulares, las sustituciones de interés se muestran en la Tabla bajo el encabezado de sustituciones preferidas. Si dichas sustituciones resultan en un cambio de la actividad biológica, entonces se introducen otros cambios sustanciales denominados sustituciones ejemplares en la Tabla 6 o como se describe en más detalle más adelante con referencia a las clases de aminoácidos que se introducen y los productos cribados.

Tabla 6

Residuo Original	Sustituciones Ejemplares	Sustituciones Preferidas
Ala (A)	val; leu; ile	val
Arg (R)	lys; gln; asn	lys
Asn (N)	gln; his; lys; arg	gln
Asp (D)	glu	glu

Cys (C)	ser	ser
Gln (Q)	asn	asn
Glu (E)	asp	asp
Gly (G)	pro; ala	ala
His (H)	asn; gln; lys; arg	arg
He (I)	leu; val; met; ala; phe; norleucina	leu
Leu (L)	norleucina; ile; val; met; ala; phe	ile
Lys (K)	arg; gln; asn	arg
Met (M)	leu; phe; ile	leu
Phe (F)	leu; val; ile; ala; tyr	leu
Pro (P)	ala	ala
Ser (S)	thr	thr
Thr (T)	ser	ser
Trp (W)	tyr; phe	tyr
Tyr (Y)	trp; phe; thr; ser	phe
Val (V)	ile; leu; met; phe; ala; norleucina	leu

Las modificaciones sustanciales de la función o de la identidad inmunológica del anticuerpo anti TAT o del polipéptido TAT se logran mediante la selección de las sustituciones que difieren significativamente en su efecto de mantener (a) la estructura de la columna vertebral del polipéptido en el área de la sustitución, por ejemplo, como una conformación de tipo lámina o helical, (b) la carga o la hidrofobia de la molécula en el sitio blanco o (c) el volumen de la cadena lateral. Los residuos que se encuentran en forma natural se dividen con base en las propiedades de las cadenas laterales comunes:

(1) hidrófobo: norleucina, met, ala, val, leu, ile;

(2) hidrofílico neutro: cys, ser, thr;

(3) acídico: asp, glu;

(4) alcalino: asn, gin, his, lys, arg;

(5) residuos que influencian la orientación de la cadena: gly, proy

(6) aromático: trp, tyr, phe.

Las sustituciones no conservadoras van a implicar intercambiar un miembro de una de estas clases por otra clase. Dichos residuos sustituidos también se pueden introducir en los sitios de sustitución conservadora, o más preferiblemente se pueden introducir en los sitios restantes (no conservados).

Las variaciones se pueden realizar empleando métodos conocidos en el arte como mutagénesis mediada por oligonucleótido (dirigido al sitio), escaneo de alanina y mutagénesis PCR. La mutagénesis dirigida al sitio [Carter y colaboradores, Nucl. Acids Res., 13:4331 (1986); Zoller y colaboradores, Nucl. Acids Res., 10:6487 (1987)], la mutagénesis de casete [Wells y colaboradores, Gene, 34:315 (1985)], la mutagénesis de restricción de selección [Wells y colaboradores, Philos. Trans. R. Soc. London SerA, 317:415 (1986)] u otras técnicas conocidas se pueden realizar en el ADN clonado para producir variantes de anticuerpo anti TAT o de polipéptido TAT.

También se puede utilizar el análisis de escaneo de aminoácido para identificar uno o varios aminoácidos a lo largo de una secuencia contigua. Entre los aminoácidos de escaneo preferidos se cuentan los aminoácidos relativamente pequeños, neutros. Dichos aminoácidos incluyen alanina, glicina, serina y cisteína. La alanina es un aminoácido de escaneo típicamente preferido entre este grupo porque elimina la cadena lateral después del beta-carbono y tiene menos probabilidades de alterar la conformación de la cadena principal del variante [Cunningham and Wells, Science, 244:1081-1085 (1989)]. La alanina también es típicamente preferida porque es el aminoácido más común. Además, con frecuencia se encuentra tanto en las posiciones enterradas como en las expuestas [Creighton, The Proteins, (W.H. Freeman & Co., N. Y.); Chotia, J. Mol. Biol., 150:1 (1976)]. Si la sustitución con alanina no produce cantidades apropiadas de variante, se puede utilizar un aminoácido isotérico.

También se puede sustituir cualquier residuo de cisteína no comprometido en la conservación apropiada del anticuerpo anti TAT o polipéptido TAT, generalmente con serina, para mejorar la estabilidad oxidante de la molécula y evitar un enlace en cruz aberrante. Inversamente, se pueden agregar enlace(s) de cisteína al anticuerpo anti TAT o al polipéptido TAT para mejorar su estabilidad (particularmente cuando el anticuerpo es un fragmento de anticuerpo como un fragmento Fv).

Un tipo particularmente preferido de variante de sustitución comprende sustituir una o varias regiones hipervariables de residuos de un anticuerpo progenitor (por ejemplo, un anticuerpo humanizado o un anticuerpo humano). Generalmente, el (los) variante(s) resultante(s) seleccionado(s) para su posterior desarrollo tendrán unas propiedades biológicas relativas al anticuerpo progenitor a partir del cual se generaron. Una forma conveniente para generar las citadas variantes de sustitución comprende una saturación por afinidad empleando un muestrario fago. Brevemente, varios sitios de regiones hipervariables (por ejemplo, los sitios 6-7) se mutan para generar todas las sustituciones amino posibles en cada sitio. Los variantes del anticuerpo generados de esta manera se representan en una forma monovalente a partir de unas partículas fago filamentosas como fusiones al gen III resultado de M13 empacado con cada partícula. Luego se criban las variantes de muestrario fago para ver su actividad biológica (por ejemplo, afinidad de fijación) como se divulga en la presente. Con el fin de identificar los sitios candidatos a región para la modificación, se puede realizar un escaneo de mutagénesis de alanina para identificar los residuos de región hipervariable que contribuyen significativamente a la fijación de antígeno. Alternativamente o adicionalmente, puede ser beneficioso analizar una estructura de cristal del complejo antígeno-anticuerpo para identificar los puntos de contacto entre el anticuerpo y el polipéptido TAT humano. Los citados residuos de contacto y los residuos contiguos son candidatos a sustitución de acuerdo con las técnicas elaboradas en la presente. Una vez las citadas variantes se generan, el panel de variantes se somete a cribado como se describe en la presente y se

pueden seleccionar los anticuerpos con propiedades superiores en uno o varios análisis para su desarrollo posterior.

Las moléculas de ácido nucleico codificando variantes de secuencias de aminoácidos del anticuerpo anti TAT se preparan mediante una variedad de métodos conocidos en el arte. Estos métodos incluyen, pero no se limitan a, el aislamiento de una fuente natural (en el caso de variantes de secuencia de aminoácidos naturales) o la preparación mediante la mutagénesis mediada por oligonucleótido (o dirigida al sitio), mutagénesis PCR y mutagénesis de casete de una versión de variante o no variante del anticuerpo anti TAT.

H. Modificaciones de Anticuerpos anti-TAT y Polipéptidos TAT

Las modificaciones covalentes de los anticuerpos anti-TAT y de los polipéptidos TAT se incluyen en el ámbito de esta invención. Un tipo de modificación covalente incluye hacer reaccionar residuos de aminoácidos objetivo de un anticuerpo anti- TAT anticuerpo o de un polipéptido TAT con un agente derivador orgánico que es capaz de reaccionar con unas cadenas laterales seleccionadas en los residuos del terminal N o C del anticuerpo anti TAT o del polipéptido TAT. La derivación con agentes bifuncionales es útil, por ejemplo, para un enlace en cruz de un anticuerpo anti TAT o de un polipéptido TAT a una matriz o superficie de soporte insoluble en agua para ser utilizado en el método para purificar anticuerpos anti-TAT y vice-versa. Unos agentes de enlace en cruz comúnmente empleados incluyen por ejemplo, 1,1-bis(diazoacetil)-2-feniletano, glutaraldehido, N-hidroxisuccinimida ésteres, por ejemplo, ésteres con 4-azido ácido salicílico, imidoésteres homobifuncionales, incluyendo disuccinimidil ésteres como 3,3'-ditiobis(succinimidilpropionato), maleimidas bifuncionales como bis-N-maleimido-1,8-octano y agentes como metil-3-[(p-azidofenil)ditio]propioimidato.

Otras modificaciones incluyen la deamidación de residuos de glutaminilo y asparaginilo a los residuos correspondientes de glutamilo y aspartilo, respectivamente, la hidroxilación de prolina y lisina, la fosforilación de grupos hidroxilo de residuos de serilo o treonilo, la

metilación de las cadenas laterales de los grupos α-amino de lisina, arginina e histidina [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetilación de la amina de terminal N y la amidación de cualquier grupo carboxilo de terminal C.

Otro tipo de modificación covalente del anticuerpo anti TAT o del polipéptido TAT incluido en el ámbito de esta invención comprende alterar el patrón de glicosilación natural del anticuerpo o polipéptido. "Alterar el patrón de glicosilación natural" para los fines de la presente tiene el propósito de eliminar una o varias fracciones de carbohidrato que se encuentran en el anticuerpo anti TAT o en el polipéptido TAT de secuencia natural (bien sea retirando el sitio subyacente de glicosilación o eliminando la glicosilación por medios químicos y/o enzimáticos) y/o agregando uno o varios sitios de glicosilación que no se encuentran presentes en el anticuerpo anti TAT o polipéptido TAT de secuencia natural. Adicionalmente, la frase incluye los cambios cualitativos en la glicosilación de las proteínas naturales comprendiendo un cambio en la naturaleza y en las proporciones de las diferentes fracciones de carbohidrato presentes.

La glicosilación de anticuerpos y otros polipéptidos típicamente tiene un enlace N o un enlace O. El enlace N hace referencia a la adhesión de la fracción de carbohidrato al lado de la cadena de un residuo de aspargina. Las secuencias tripéptido asparagina-X-serina y asparagina-X-treonina, donde X es un aminoácido excepto prolina, son las secuencias de reconocimiento para la adhesión enzimática de la fracción de carbohidrato a la cadena lateral de asparagina. De este modo, la presencia de cualquiera de estas secuencias tripéptido en un polipéptido crea un sitio de glicosilación potencial. La glicosilación de enlace O hace referencia a la adhesión de uno de los azúcares N-aceilgalactosamina, galactosa o xilosa a un hidroxiaminoácido, más comúnmente serina o treonina, aunque también se puede utilizar 5-hidroxiprolina o 5-hidroxilisina.

La adición de sitios de glicosilación al anticuerpo anti TAT o al polipéptido TAT se logra de forma conveniente alterando la secuencia de

aminoácido de manera que contenga una o varias de las secuencias tripéptido descritas arriba (para sitios de glicosilación de enlace N). La alteración también se puede realizar mediante la adición de o la sustitución por uno o varios residuos de serina o treonina de la secuencia del anticuerpo anti TAT o del polipéptido TAT original (para los sitios de glicosilación de enlace O). La secuencia de aminoácido del anticuerpo anti TAT o del polipéptido TAT opcionalmente se puede alterar mediante cambios a nivel del ADN, particularmente mutando el ADN codificando el anticuerpo anti TAT o el polipéptido TAT en las bases previamente seleccionadas de manera que se generen unos codones que se van a trasladar a los aminoácidos deseados.

Otros medios para incrementar el número de fracciones de carbohidrato en el anticuerpo anti TAT o el polipéptido TAT son los acoplamientos químicos o enzimáticos de glicosidos al polipéptido. Estos métodos se describen en el arte, por ejemplo, en WO 87/05330 publicada el 11 de septiembre de 1987 y en Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

La extracción de las fracciones de carbohidrato presentes en el anticuerpo anti TAT o el polipéptido TAT se puede lograr químicamente o enzimáticamente o mediante la sustitución mutacional de los codones codificando residuos de aminoácidos que sirven como blancos para la glicosilación. Las técnicas químicas de desglicosilación son conocidas en el arte y se describen, por ejemplo, por Hakimuddin, y colaboradores, Arch. Biochem. Biophys., 259:52 (1987) y por Edge y colaboradores, Anal. Biochem., 118:131 (1981). La segmentación enzimática de las fracciones de carbohidrato en los polipéptidos se puede lograr mediante el uso de una variedad de endo- y exo-glicosidasas como se describe por Thotakura y colaboradores, Meth. Enzymol., 138:350 (1987).

Otro tipo de modificación covalente de anticuerpo anti TAT o polipéptido TAT comprende enlazar el anticuerpo o polipéptido a uno de una variedad de polímeros no proteináceos, por ejemplo, polietilenglicol (PEG), polipropilenglicol o polioxialquilenos, en la forma descrita en las

Patentes de Estados Unidos Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 o 4,179,337. Igualmente el anticuerpo o el polipéptido se puede comprimir en microcápsulas preparadas por ejemplo, por medio de técnicas de coacervación o mediante polimerización interfacial (por ejemplo, microcápsulas de hidroximetilcelulosa o microcápsulas de gelatina y poli-(metilmetacilato), respectivamente), en sistemas coloidales de administración de fármacos (por ejemplo, liposomas, microesferas de albúmina, microemulsiones, nano-partículas y nanocápsulas) o en macroemulsiones. Dichas técnicas se divulgan en Remington's Pharmaceutical Sciences, Edición 16, Oslo, A., Ed., (1980).

El anticuerpo anti TAT o el polipéptido TAT de la presente invención también se puede modificar de manera que formen moléculas quiméricas comprendiendo un anticuerpo anti-TAT o un polipéptido TAT fusionado al otro polipéptido heterólogo o secuencia de aminoácido.

En una incorporación, dicha molécula quimérica comprende una fusión del anticuerpo anti TAT o del polipéptido TAT con un polipéptido marcador que dispone un epítopo al cual se puede adherir un anticuerpo anti-marcador en forma selectiva. El epítopo marcador generalmente se ubica en el terminal amino o carboxilo del anticuerpo anti TAT o del polipéptido TAT. La presencia de dichas formas de epítopo-marcado del anticuerpo anti TAT o del polipéptido TAT se pueden detectar utilizando un anticuerpo contra el polipéptido marcador. Igualmente una disposición del epítopo marcador permite que el anticuerpo anti TAT o el polipéptido TAT fácilmente se purifique mediante purificación por afinidad empleando un anticuerpo anti-marcador u otro tipo de matriz de afinidad que se fija al epítopo marcador. Varios polipéptidos marcadores y sus respectivos anticuerpos son bien conocidos en el arte. Ejemplos incluyen marcadores poli-histidina (poli-his) o poli-histidina-glicina (poli-his-gly); el polipéptido marcador flu HA y su anticuerpo 12CA5 [Field y colaboradores, Mol. Cell. Biol., 8:2159- 2165 (1988)]; el marcador c-myc y los anticuerpos 8F9, 3C7, 6E10, G4, B7 y 9E10 del mismo [Evan y colaboradores, Molecular and Cellular Biology, 5:3610-3616 (1985)]y el marcador de la glicoproteína D

(gD) del virus Herpes Simplex y su anticuerpo [Paborsky y colaboradores, Proteína Engineering, 3(6):547-553 (1990)]. Otros polipéptidos marcadores incluyen el péptido Flag- [Hopp y colaboradores, BioTechnology, 6:1204-1210 (1988)]; el epítopo peptídico KT3 [Martin y colaboradores, Science, 255:192-194 (1992)]; un epítopo peptídico o α-tubulina [Skinner y colaboradores, J. Biol. Chem., 266:15163-15166 (1991)]y el gen T7 del marcador de la proteína peptídica 10 [Lutz-Freyermuth y colaboradores, Proc. Natl. Acad. Sci. USA, 87:6393- 6397 (1990)].

En una incorporación alterna, la molécula quimérica puede comprender una fusión del anticuerpo anti TAT o del polipéptido TAT con una inmunoglobulina o una región particular de una inmunoglobulina. Para una forma bivalente de la molécula quimérica (a la cual también se hace referencia como una "inmunoadhesina"), la citada fusión puede ser a la región Fc de una molécula IgG. Las fusiones Ig preferiblemente incluyen la sustitución de una forma soluble (dominio transmembrana eliminado o inactivado) de un anticuerpo anti-TAT o de un polipéptido TAT en el lugar de por lo menos una región variable en una molécula Ig. En una incorporación particularmente preferida, la fusión de inmunoglobulina incluye la bisagra de las regiones CH_2 y CH_3 o la bisagra, CH_1, CH_2 de CH_3 de una molécula IgG1. Para la producción de fusiones de inmunoglobulina consultar también la Patente de Estados Unidos No. 5,428,130 otorgada el 27 de junio de 1995.

I. Preparación de Anticuerpos anti-TAT y Polipéptidos TAT

La siguiente descripción se relaciona básicamente con la producción de anticuerpos anti-TAT y de polipéptidos TAT cultivando células transformadas o transinfectadas con un vector que contiene el anticuerpo anti TAT- y el polipéptido TAT-codificando ácido nucleico. Naturalmente se contempla que métodos alternos, que son bien conocidos en el arte, se pueden emplear para preparar los anticuerpos anti-TAT y los polipéptidos TAT. Por ejemplo, la secuencia apropiada de aminoácidos o porciones de la misma, se pueden producir mediante la síntesis directa de péptidos empleando técnicas de fase sólida [consultar por ejemplo, Stewart y

colaboradores, Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc, 85:2149-2154 (1963)]. La síntesis de proteína *in vitro* se puede realizar empleando técnicas manuales o mediante automatización. La síntesis automatizada se puede realizar por ejemplo empleando un sintetizador *Applied Biosystems Peptide Synthesizer* (Foster City, CA) utilizando las instrucciones del fabricante. Varias porciones del anticuerpo anti TAT o del polipéptido TAT se pueden sintetizar químicamente en forma separada y combinada utilizando métodos químicos o enzimáticos para producir el anticuerpo anti TAT o el polipéptido TAT deseados.

1. Aislamiento de ADN Codificando Anticuerpo anti TAT o polipéptido TAT

Se puede obtener ADN codificando el anticuerpo anti TAT o el polipéptido TAT a través de una biblioteca ADNc preparada a partir de tejido que se creía poseía el anticuerpo anti TAT o el polipéptido ARNM del TAT y para expresarlo a un nivel detectable. Por consiguiente, el ADN del anticuerpo anti-TAT humano o del polipéptido TAT convenientemente se puede obtener de una biblioteca ADNc preparada a partir de tejido humano. El gen de codificación del anticuerpo anti TAT o polipéptido TAT también se puede obtener de una biblioteca genómica o mediante procedimientos sintéticos conocidos (por ejemplo, síntesis automatizada de ácido nucleico).

Las bibliotecas se pueden cribar con sondas (como oligonucleótidos de por lo menos alrededor de 20-80 bases) designadas para identificar el gen de interés o la proteína que lo codifica. Cribar el ADNc o la biblioteca genómica con la sonda seleccionada se puede realizar mediante procedimientos estándar, como se describe en Sambrook y colaboradores, Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). Un medio alterno para aislar el gen codificando el anticuerpo anti TAT o el polipéptido TAT es emplear la metodología PCR [Sambrook y colaboradores, supra: Dieffenbach y colaboradores, PCR

Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

Las técnicas para cribar una biblioteca ADNc son bien conocidas en el arte. Las secuencias de oligonucleótidos seleccionadas como sondas deben tener una longitud suficientemente ambigua para reducir los falsos positivos. El oligonucleótido preferiblemente está marcado de manera que se pueda detectar en el momento de la hibridación al ADN en la biblioteca que se está cribando. Los métodos de marcado son bien conocidos en el arte e incluyen el uso de radiomarcadores como ^{32}P-marcado ATP, marcado por biotinilación o enzima. Las condiciones de hibridación, incluyendo una severidad moderada y una severidad alta se disponen en Sambrook y colaboradores, supra.

Las secuencias identificadas en dichos métodos de cribado de bibliotecas se pueden comparar y alinear con otras secuencias conocidas depositadas y disponibles en las bases de datos públicas como el GenBank u otras bases de datos de secuencias privadas. La identidad de secuencia (bien sea a nivel de aminoácido o nucleótido) en las regiones definidas de la molécula o a lo largo de la secuencia de longitud total se puede determinar utilizando métodos conocidos en el arte y como se describe en la presente.

Se puede obtener ácido nucleico teniendo una secuencia de codificación de proteína cribando el ADNc seleccionado o las bibliotecas genómicas utilizando la secuencia de aminoácidos deducida divulgada en la presente para la primera vez, y de ser necesario, empleando los procedimientos convencionales de extensión de iniciadores como se describe en Sambrook y colaboradores, supra, para detectar precursores y procesar los intermedios de ARNm que posiblemente no se hayan transcrito de forma inversa en el ADNc.

2. Selección y Transformación de Células Hospederas

Las células hospederas se transinfectan o se transforman con vectores de expresión o de clonación descritos en la presente para la producción de anticuerpo anti TAT o de polipéptido TAT y se cultivan en medios convencionales de nutrientes modificados según lo apropiado para

introducir promotores, seleccionar transformantes o amplificar los genes codificando las secuencias deseadas. Las condiciones de cultivo, como el medio, la temperatura, el pH y similares se pueden seleccionar por parte de un artesano con destreza sin una experimentación indebida. En general, los principios, los protocolos y las técnicas prácticas para maximizar la productividad de los cultivos de las células se puede encontrar en Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) y Sambrook y colaboradores, supra.

Los métodos para la transinfección de la célula eucariota y la transformación de la célula procariota son conocidos por el artesano con destrezas ordinarias, por ejemplo, $CaCl_2$, $CaPO_4$, mediado por liposoma y electroporado. Dependiendo de la célula hospedera utilizada, la transformación se realiza empleando unas técnicas estándar apropiadas para dichas células. El tratamiento con calcio que emplea cloruro de calcio, como se describe en Sambrook y colaboradores, supra o la electroporado se utiliza generalmente para las procariotas. La infección con *Agrobacterium tumefaciens* se emplea para la transformación de ciertas células de plantas como se describe por Shaw y colaboradores, Gene, 23:315 (1983) y WO 89/05859 publicada el 29 de junio de 1989. Para las células de mamíferos sin dichas paredes celulares, se puede utilizar el método de precipitación por fosfato de calcio de Graham and van der Eb, Virology, 52:456-457 (1978). Los aspectos generales del sistema de transinfecciones de células hospederas de los mamíferos se han descrito en la Patente de Estados Unidos No. 4,399,216. Las transformaciones en levadura típicamente se han realizado de acuerdo con el método de Van Solingen y colaboradores, J. Bact, 130:946 (1977) y Hsiao y colaboradores, Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). Sin embargo, también se pueden utilizar otros métodos para introducir el ADN en las células, como mediante microinyección nuclear, electroporado, fusión bacteriana protoplasta con células intactas o policaciones, por ejemplo, polibreno, poliornitina. Para ver varias técnicas para transformar las células de mamíferos, consultar

Keown y colaboradores, Methods in Enzymology, 185:527-537 (1990) y Mansour y colaboradores, Nature, 336:348-352 (1988).

Unas células hospederas apropiadas para clonar o expresar el ADN en los vectores en la presente incluyen células eucariotas, levadura o células eucariotas mayores. Las procariotas apropiadas incluyen pero no se limitan a eubacterias, como organismos Gram-negativos o Gram-positivos por ejemplo, Enterobacteriaceae como *E. coli*. Varias cepas de *E. coli* están disponibles para el público, como *E. coli* K12 cepa MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); la cepa *E. coli* W3110 (ATCC 27,325) y K5 772 (ATCC 53,635). Otras células hospederas procariotas incluyen Enterobacteriaceae como *Escherichia*, por ejemplo, *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, por ejemplo, *Salmonella typhimurium*, *Serratia*, por ejemplo, *Serratia marcescans* y *Shigella*, también como *Bacilli* como *B. subtilis* y B. *licheniformis* (por ejemplo, *B. licheniformis* 41P divulgada en DD 266,710 publicada el 12 de abril de 1989), *Pseudomonas* como *P. aeruginosa* y *Streptomyces*. Estos ejemplos son ilustrativos más que limitantes. La cepa W3110 es un hospedero o un hospedero progenitor particularmente preferido porque es una cepa hospedera común para las fermentaciones del producto de ADN recombinante. Preferiblemente, la célula hospedera secreta cantidades mínimas de enzimas proteolíticas. Por ejemplo, la cepa W3110 se puede modificar para efectuar una mutación genética en los genes codificando proteínas endógenas al hospedero, con ejemplos de dichos hospederos incluyendo *E. coli* W3110 cepa 1A2, que tiene el genotipo completo tonA ; *E. coli* W3110 cepa 9E4, que tiene el genotipo completo tonA ptr3; *E. coli* W3110 cepa 27C7 (ATCC 55,244), que tiene el genotipo completo *tonA ptr3 phoA E15 (argF-lac)169 degP ompT kan*^{r}; *E. coli* W3110 cepa 37D6, que tiene el genotipo completo *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 UvG kan*^{r}; *E. coli* W3110 cepa 40B4, que es la cepa 37D6 con una mutación de eliminación no canamicina *degP* resistente; y una cepa *E. coli* teniendo una proteasa periplásmica mutante divulgada en la Patente de Estados Unidos No. 4,946,783 otorgada el 7 de agosto de 1990.

Alternativamente, unos métodos *in vitro* de clonación, por ejemplo, PCR u otras reacciones de polimerasa de ácido nucleico son apropiados.

Los anticuerpos de longitud total, los fragmentos de anticuerpos y las proteínas de fusión se pueden producir en bacterias, en particular cuando la glicosilación y la función efectora Fc no se necesitan, como cuando el anticuerpo terapéutico se conjuga a un agente citotóxico (por ejemplo, una toxina) y el inmunoconjugado por sí mismo muestra efectividad en la destrucción de las células tumorales. Los anticuerpos de longitud total tienen una vida media en circulación mayor. La producción en *E. coli* es más rápida y más eficiente desde la perspectiva de costos. Para la expresión de fragmentos de anticuerpos y polipéptidos en bacterias, consultar por ejemplo, U.S.A. 5,648,237 (Carter y colaboradores), U.S.A. 5,789,199 (JoIy y colaboradores) y U.S.A. 5,840,523 (Simmons y colaboradores) que describen la región de inicio de la traslación (TIR) y las secuencias de señal para optimizar la expresión y la secreción; estas patentes incorporadas a la presente por referencia. Después de la expresión, el anticuerpo se aísla de la pasta celular de *E. coli* en una fracción soluble y se puede purificar mediante, por ejemplo, una columna de proteína A o G dependiendo del isotipo. La purificación final se puede realizar de forma similar al proceso para purificar el anticuerpo expresado por ejemplo en las células CHO.

Además de las procariotas, comúnmente se emplean microbios eucariotas como fungal filamentoso o levadura como hospederos de clonación o expresión apropiados para los vectores codificando el anticuerpo anti TAT o el polipéptido TAT. *Saccharomyces cerevisiae* es un microorganismo menor eucariota comúnmente empleado. Otros incluyen *Schizosaccharomyces pombe* (Beach and Nurse, Nature, 290: 140 [1981]; EP 139,383 publicada el 2 de mayo de 1985); *Kluyveromyces hosts* (Patente de Estados Unidos No. 4,943,529; Fleer y colaboradores, Bio/Technology, 9:968-975 (1991)) como, por ejemplo, K. lactis (MW98-8C, CBS683, CBS4574; Louvencourt y colaboradores, J. Bacteriol., 154(2):737-742 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickeramii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilarum*

(ATCC 36,906; Van den Berg y colaboradores, Bio/Technology, 8:135 (1990)), *K. thermotolerans y K. marxianus yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna y colaboradores, J. Basic Microbiol., 28:265-278 [1988]); *Candida; Trichoderma reesia* (EP 244,234); *Neurospora crassa* (Case y colaboradores, Proc. Natl. Acad. Sci. USA, 76:5259-5263 [1979]); *Schwanniomyces* como *Schwanniomyces occidentalis* (EP 394,538 publicada el 31 de octubre 1990) y fungi filamentoso como por ejemplo, *Neurospora, Penicillium, Tolipocladium* (WO 91/00357 publicada el 10 de enero de 1991) y hospederos *Aspergillus* como *A. nidulans* (Ballance y colaboradores, Biochem. Biophys. Res. Commun., 112:284-289 [1983]; Tilburn y colaboradores, Gene, 26:205-221 [1983] yelton y colaboradores, Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984]) y A. *niger* (Kelly and Hynes, EMBO J.. 4:475-479 [1985]). Las levaduras metilotrópicas son apropiadas en la presente e incluyen, pero no se limitan a, levaduras capaces de crecer sobre metanol seleccionadas entre el género conformado por *Hansenula, Candida, Kloeckera, Pichia, Saccharomyces, Torulopsis* y *Rhodotorula*. Se puede encontrar una lista de especies específicas ejemplares de esta clase de levaduras en C. Anthony, The Biochemistry of Metilotrophs, 269 (1982).

Unas células hospederas apropiadas para la expresión del anticuerpo glicosilado anti TAT o el polipéptido TAT se derivan de organismos multicelulares. Ejemplos de células invertebradas incluyen células de insectos como Drosophila S2 y Spodoptera Sf9, al igual que células de plantas, como cultivos de células de algodón, maíz, papa, soya, petunia, tomate y tabaco. Se han identificado numerosas cepas baculovirales y variantes y las células hospederas permisivas correspondientes de insectos de hospederos como *Spodoptera frugiperda* (oruga), *Aedes aegypti* (mosquito), *Aedes albopictus* (mosquito), *Drosophila melanogaster* (mosca de la fruta) y *Bombyx mori*. Una variedad de cepas virales para la transinfección están disponibles para el público por ejemplo, la variante L-I de *Autographa californica* NPV y la cepa Bm-5 de *Bombyx mori* NPV y estos virus se pueden emplear como el virus descrito en la presente, de

acuerdo con la presente invención, particularmente para la transinfección de células *Spodoptera frugiperda.*

Sin embargo el interés ha sido mayor en las células de vertebrados, y la propagación de células de vertebrados en cultivos (cultivos de tejidos) se ha convertido en un proceso rutinario. Ejemplos de cepas de células hospederas de mamíferos útiles son la cepa de riñón de mono CV1 transformada por SV40 (COS-7, ATCC CRL 1651); la cepa de riñón embriónico humano (293 o 293 células subclonadas para el crecimiento en un cultivo de suspensión, Graham y colaboradores, J. Gen Virol. 36:59 (1977)); células de riñón de hámster bebé (BHK, ATCC CCL 10); células ováricas de hámster chino/-DHFR (CHO, Urlaub y colaboradores, Proc. Natl. Acad. Sci. USA 77:4216 (1980)); células Sertoli de ratón (TM4, Mather, Biol. Reprod. 23:243-251 (1980)); células de riñón de mono (CV1 ATCC CCL 70); células de riñón de mono verde de África (VERO-76, ATCC CRL-1587); células de carcinoma cervical humano (HELA, ATCC CCL 2); células de riñón canino (MDCK, ATCC CCL 34); células de hígado de búfalo una rata (BRL 3A, ATCC CRL 1442); células de pulmón humano (W138, ATCC CCL 75); células de hígado humano (Hep G2, HB 8065); tumor mamario de ratón (MMT 060562, ATCC CCL51); células TRI (Mather y colaboradores, Annals N. Y. Acad. Sci. 383:44-68 (1982)); células MRC 5; células FS4 y una cepa de hepatoma humano. (Hep G2).

Las células hospederas se transforman en los vectores de expresión o clonación descritos arriba para la producción del anticuerpo anti TAT o del polipéptido TAT y se cultivan en un medio de nutriente convencional modificado según sea apropiado para inducir los promotores, seleccionar los transformantes o amplificar los genes codificando las secuencias deseadas.

3. Selección y Uso de un Vector Replicable

El ácido nucleico (por ejemplo, ADNc o ADN genómico) codificando el anticuerpo anti TAT o el polipéptido TAT se puede insertar en un vector replicable para clonación (amplificación del ADN) o para expresión. Varios vectores están disponibles para el público. El vector

puede por ejemplo, ser en forma de un plásmido, un cósmido, una partícula viral o un fago. La secuencia de ácido nucleico apropiada se puede insertar en el vector mediante una variedad de métodos. En general, el ADN se inserta en un(os) sitio(s) apropiado(s) de restricción de endonucleasa utilizando técnicas conocidas en el arte. Generalmente los componentes de los vectores incluyen, pero no se limitan a, una o varias secuencias de señal, un origen de replicación, uno o varios genes marcadores, un elemento intensificador, un promotor y una secuencia de terminación de trascripción. La construcción de unos vectores apropiados conteniendo uno o varios de estos componentes emplea unas técnicas estándar de ligado conocidas por el artesano con destrezas.

El TAT se puede producir de forma recombinante no directamente, pero también como un polipéptido de fusión con un polipéptido heterólogo, que puede ser una secuencia de señal u otro polipéptido teniendo un sitio de segmentación específico en el terminal N de la proteína o del polipéptido maduro. En general, la secuencia de señal puede ser un componente del vector o puede ser una parte del anticuerpo anti TAT o del polipéptido TAT-codificando ADN que se inserta en el vector. La secuencia de señal puede ser una secuencia de señal procariota seleccionada por ejemplo, entre el grupo de fosfatasa alcalina, penicilinasa, lpp o unos líderes enterotoxina estables en el calor. Para la secreción de levadura, la secuencia de señal puede ser, por ejemplo, el líder de invertasa de levadura, el líder del factor alfa (*incluyendo Saccharomyces* y *Kluyveromyces* líderes del α-factor, este último descrito en la Patente de Estados Unidos No. 5,010,182) o el líder de ácido fosfatasa, el líder glucoamilasa *C. albicans* (EP 362,179 publicada el 4 de abril de 1990) o la señal descrita en WO 90/13646 publicada el 15 de noviembre de 1990. En la expresión de células de mamíferos, las señales de secuencia de los mamíferos se pueden utilizar para dirigir la secreción de la proteína, como señales de secuencia de polipéptidos secretados de la misma especie, o de especies relacionadas, al igual que los líderes secretores virales.

Tanto los vectores de expresión como de clonación contienen una secuencia de ácido nucleico que permite que el vector se replique en una o en varias células hospederas. Dichas secuencias son bien conocidas para una variedad de bacterias, levaduras y virus. El origen de replicación del plásmido ρBR322 es apropiado para la mayoría de las bacterias Gram-negativas, el 2μ de origen plásmido es apropiado para la levadura y varios de orígenes virales (SV40, polioma, adenovirus, VSV o BPV) son útiles para los vectores de clonación en células de mamíferos.

Típicamente los vectores de expresión y de clonación van a contener un gen de selección, también denominado un marcador seleccionable. Los genes de selección típicos codifican proteínas que (a) confieren resistencia a los antibióticos y a otras toxinas, por ejemplo, ampicilina, neomicina, metotrexato o tetraciclina, (b) complementan las deficiencias auxotróficas o (c) aportan nutrientes críticos que no están disponibles en el complejo del medio, por ejemplo, el gen codificando D-alanina racemasa para *Bacilli.*

Un ejemplo de marcadores seleccionables apropiados para células de mamíferos son aquellos que permiten la identificación de las células competentes para absorber el anticuerpo anti TAT o el polipéptido TAT codificando ácido nucleico, como DHFR o timidina cinasa. Una célula hospedera apropiada cuando se emplea el DHFR de tipo natural es la estirpe celular CHO con deficiencia de actividad DHFR, preparada y propagada como se describe por Urlaub y colaboradores, Proc. Natl. Acad. Sci. USA, 77:4216 (1980). Un gen de selección apropiado para ser utilizado en la levadura es el gen *trp*1 presente en el plásmido de levadura YRp7 [Stinchcomb y colaboradores, Nature, 282:39 (1979); Kingsman y colaboradores, Gene, 7:141 (1979); Tschemper y colaboradores, Gene, 10:157 (1980)]. El gen *trp*1 dispone un marcador de selección para una cepa mutante de levadura que carece de la habilidad de crecer en triptofano, por ejemplo, ATCC No. 44076 o PEP4-1 [Jones, Genetics, 85:12 (1977)].

Normalmente los vectores de expresión y de clonación contienen un promotor operacionalmente enlazado al anticuerpo anti TAT o polipéptido TAT codificando la secuencia de ácido nucleico para dirigir la síntesis

ARNm. Los promotores reconocidos por una variedad de células hospederas potenciales son bien conocidos. Los promotores apropiados para el uso con los hospederos procariotas incluyen los sistemas de promotores β-lactamasa y lactosa [Chang y colaboradores, Nature, 275:615 (1978); Goeddel y colaboradores, Nature, 281:544 (1979)], fosfatasa alcalina, un sistema promotor triptofano (trp) [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776] y promotores híbridos como el promotor tac [deBoer y colaboradores, Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Los promotores para ser utilizados en los sistemas bacterianos también van a contener una secuencia Shine-Dalgarno (S.D.) enlazada operacionalmente al ADN codificando el anticuerpo anti TAT o el polipéptido TAT.

Ejemplos de secuencias promotoras apropiadas para el uso con hospederos de levadura incluyen los promotores para 3-fosfoglicerato cinasa [Hitzeman y colaboradores, J. Biol. Chem., 255:2073 (1980)] u otras enzimas glicolíticas [Hess y colaboradores, J. Adv. Enzima Reg., 7: 149 (1968); Holland, Biochemistry, 17:4900 (1978)], como enolasa, gliceraldehido-3-fosfato dehidrogenasa, hexocinasa, piruvato decarboxilasa, fosfofructocinasa, glucosa-6-fosfato isomerasa, 3-fosfoglicerato mutasa, piruvato cinasa, triosefosfato isomerasa, fosfoglucosa isomerasa y glucocinasa.

Otros promotores de levadura, que son promotores inducibles teniendo la ventaja adicional de trascripción controlada por condiciones de crecimiento, son las regiones de promotores para alcohol dehidrogenasa 2, isocitocroma C, ácido fosfatasa, enzimas degradantes asociadas con el metabolismo de nitrógeno, metalotioneína, gliceraldehido-3-fosfato dehidrogenasa y enzimas responsables del uso de maltosa y galactosa. Unos vectores y promotores apropiados para el uso en la expresión de la levadura se describen con más detalle en EP 73,657.

La trascripción del anticuerpo anti TAT o del polipéptido TAT a partir de vectores en células hospederas de mamíferos se controla por ejemplo, por medio de promotores obtenidos de los genomas de virus como el virus de polioma, virus de viruela aviar (UK 2,211,504 publicada el 5 de

julio de 1989), adenovirus (como Adenovirus 2), virus de papiloma bovino, virus de sarcoma aviar, citomegalovirus, un retrovirus, virus de hepatitis-B y Virus Simian 40 (SV40), de promotores mamíferos heterólogos, por ejemplo, el promotor actina o un promotor de inmunoglobulina y de los promotores de shock por calor, siempre que dichos promotores sean compatibles con los sistemas de la célula hospedera.

La trascripción de un ADN codificando el anticuerpo anti TAT o el polipéptido TAT por eucariotas más altas se puede incrementar al insertar una secuencia potenciadora en el vector. Los potenciadores son elementos cis-activos de ADN, normalmente de alrededor de 10 a 300 bp, que actúan en un promotor para incrementar su trascripción. Ya se conocen muchas secuencias potenciadoras para los genes de los mamíferos (globina, elastasa, albúmina, α-fetoproteína e insulina). Sin embargo típicamente, uno va a emplear un potenciador de un virus de una célula eucariota. Ejemplos incluyen el potenciador SV40 en el lado tardío del origen de replicación (bp 100-270), el potenciador del promotor temprano de citomegalovirus, el potenciador de polioma en el extremo tardío del origen de replicación y potenciadores de adenovirus. El potenciador se puede cortar y empalmar en el vector en una posición 5' o 3' a la secuencia de codificación del anticuerpo anti TAT o del polipéptido TAT, pero preferiblemente se ubica en el sitio 5' del promotor.

Los vectores de expresión empleados en las células hospederas eucariotas (levadura, fungi, insecto, planta, animal, humano o células nucleadas de otros organismos multicelulares) también van a contener las secuencias necesarias para la terminación de la trascripción y para estabilizar el ARNm. Dichas secuencias están comúnmente disponibles en la región sin trasladar 5' y ocasionalmente en la 3', de ADNs o ADNc eucariotas o virales. Estas regiones contienen segmentos de nucleótidos transcritos como fragmentos poliadenilados en la porción sin trasladar del ARNm codificando el anticuerpo anti TAT o el polipéptido TAT.

Aún otros métodos, vectores y células hospederas apropiados para la adaptación a la síntesis del anticuerpo anti TAT o del polipéptido TAT en

cultivos recombinantes de células de vertebrados se describen en Gething y colaboradores, Nature, 293:620-625 (1981); Mantei y colaboradores, Nature, 281:40-46 (1979); EP 117,060y EP 117,058.

4. Cultivando las Células Hospederas

Las células hospederas utilizadas para producir el anticuerpo anti TAT o el polipéptido TAT de esta invención se pueden cultivar en una variedad de medios. Los medios disponibles en el comercio como Ham's F10 (Sigma), Minimal Essential Medium ((MEM), (Sigma), RPMI-1640 (Sigma) y Dulbecco's Modified Eagle's Medium ((DMEM), Sigma) son apropiados para cultivar las células hospederas. Adicionalmente, cualquiera de los medios descritos en Ham y colaboradores, Meth. Enz. 58:44 (1979), Barnes y colaboradores, Anal. Biochem.102:255 (1980), Patentes de Estados Unidos Nos. 4,767,704; 4,657,866; 4,927,762; 4,560,655; o 5,122,469; WO 90/03430; WO 87/00195; o Patente de Estados Unidos Re. 30,985 se pueden emplear como medios de cultivo para las células hospederas. Cualquiera de estos medios se puede complementar, según sea necesario, con hormonas y/u otros factores de crecimiento (como insulina, transferina o factor de crecimiento epidérmico), sales (como cloruro de sodio, calcio, magnesio y fosfato), soluciones amortiguadoras (como HEPES), nucleótidos (como adenosina y timidina), antibióticos (como GENTAMYCIN™ fármaco), oligoelementos (definidos como compuestos inorgánicos normalmente presentes en concentraciones finales en el rango micromolar) y glucosa o una fuente de energía equivalente. Cualquier otro complemento necesario adicional también se puede incluir a las concentraciones apropiadas que serían conocidos por aquellos con destreza en el arte. Las condiciones de cultivo, como temperatura, pH y similares, son aquellas previamente empleadas con la célula hospedera seleccionada para la expresión y serán aparentes para el artesano con destrezas ordinarias.

5. Detectando la Amplificación/Expresión de un Gen

La amplificación y/o expresión de un gen se puede medir directamente en una muestra, por ejemplo, mediante el método Southern y

el método Northern convencionales para cuantificar la trascripción de ARNm [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], inmunotransferencia (análisis ADN) o hibridación *in situ*, utilizando una sonda correctamente marcada, con base en las secuencias descritas en la presente. Alternativamente, se pueden emplear anticuerpos que pueden reconocer duplicidades específicas, incluyendo duplicidades de ADN, duplicidades de ARN y duplicidades híbridas ADN-ARN o duplicidades ADN-proteína. A su vez, los anticuerpos se pueden marcar y se puede efectuar un análisis donde la duplicidad se una a una superficie, de manera que cuando se forme la duplicidad en la superficie, se pueda detectar la presencia del anticuerpo fijado a la duplicidad.

Alternativamente se puede medir la expresión del gen mediante métodos inmunológicos, como la coloración inmunohistoquímica de secciones de células o tejidos y el análisis de cultivos de células o fluidos corporales, para cuantificar directamente la expresión del producto del gen. Los anticuerpos útiles para la coloración inmunohistoquímica y/o el análisis de fluidos de muestra pueden ser monoclonales o policlonales, y se pueden preparar en cualquier mamífero. Convenientemente, los anticuerpos se pueden preparar contra un polipéptido TAT de secuencia natural o contra un péptido sintético con base en las secuencias de ADN suministradas en la presente o contra secuencia exógenas fusionadas a ADN de TAT y codificando un epítopo específico de anticuerpo.

6. Purificación de Anticuerpo anti TAT y Polipéptido TAT

Se pueden recuperar formas de anticuerpo anti TAT y de polipéptido TAT de un medio de cultivo o de unos lisados de la célula hospedera. Si están limitados por la membrana, se pueden liberar de la membrana utilizando una solución detergente apropiada (por ejemplo Triton-X 100) o mediante segmentación enzimática. Las células empleadas en la expresión del anticuerpo anti TAT y del polipéptido TAT se pueden alterar por diversos medios físicos o químicos, como unos ciclos de congelamiento y descongelamiento, sonicación, alteración mecánica o por medio de agentes de lisado de células.

Se puede desear purificar el anticuerpo anti TAT y el polipéptido TAT a partir de células recombinantes de proteínas o polipéptidos. Los siguientes procedimientos son ejemplares para unos procedimientos de purificación apropiados: fraccionamiento en una columna de intercambio de iones; precipitación con etanol; HPLC de fase inversa; cromatografía en sílice o en una resina de intercambio de cationes como DEAE; cromatoenfoque; SDS-PAGE; precipitación con sulfato de amonio; filtración por gel utilizando, por ejemplo, Sephadex G-75; columnas de proteína A Sepharose para extraer los contaminantes como IgG y columnas de metal quelante para fijar las formas epítopo-marcadas del anticuerpo anti TAT y del polipéptido TAT. Se pueden emplear varios métodos de purificación de proteína y estos métodos son conocidos en el arte y se describen por ejemplo en Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). El (los) paso(s) de purificación seleccionado(s) va(n) a depender por ejemplo, de la naturaleza del proceso de producción y del anticuerpo anti TAT o polipéptido TAT particular producido.

Cuando se emplean técnicas recombinantes, el anticuerpo se puede producir a nivel intracelular, en el espacio periplásmico o directamente secretado en el medio. Si el anticuerpo se produce a nivel intracelular, como un primer paso, los residuos particulados, ya sea células hospederas o fragmentos lisados, se extraen, por ejemplo, mediante centrifugación o ultrafiltración. Carter y colaboradores, Bio/Technology 10:163-167 (1992) describe un procedimiento para aislar los anticuerpos que se secretan al espacio periplásmico de *E. coli*. Brevemente, la pasta celular se descongela en presencia de acetato de sodio (pH 3.5), EDTA y fenilmetilsulfonilfluoruro (PMSF) durante unos 30 min. Los residuos celulares se pueden extraer mediante centrifugado. Cuando el anticuerpo se secreta en el medio, generalmente primero se concentran los sobrenadantes de dichos sistemas de expresión utilizando un filtro de concentración de proteína disponible en el comercio, por ejemplo, una unidad de ultrafiltración Amicon o Millipore Pellicon. Se puede incluir un inhibidor

de proteasa como PMSF en cualquiera de los pasos anteriores para inhibir la proteolisis y se pueden incluir antibióticos para evitar el crecimiento de contaminantes accidentales.

La composición del anticuerpo preparada a partir de las células se puede purificar, utilizando por ejemplo, la cromatografía de hidroxilapatita, electroforesis en gel, diálisis y cromatografía por afinidad, donde la cromatografía por afinidad es la técnica de purificación preferida. La idoneidad de la proteína A como un ligando de afinidad depende de las especies y de los isotipos de cualquier dominio Fc de inmunoglobulina presente en el anticuerpo. Se puede utilizar la proteína A para purificar los anticuerpos que se basan en cadenas pesadas humanas γ1, γ2 o γ4 (Lindmark y colaboradores, J. Inmunol. Meth. 62:1-13 (1983)). Se recomienda la proteína G para todos los isotipos de ratón y para el γ3 humano (Guss y colaboradores, EMBO J. 5:15671575 (1986)). La matriz a la cual se adhiere el ligando de afinidad con mayor frecuencia es agarosa, pero otras matrices están disponibles. Las matrices mecánicamente estables como vidrio de poro controlado o poli(estirendivinilo)benceno permiten unas tasas de flujo más altas y unos periodos de procesamiento más cortos que los que se pueden alcanzar con agarosa. Cuando el anticuerpo comprende un dominio CH3, la resina Bakerbond ABX™ (J. T. Baker, Phillipsburg, NJ) es útil para la purificación. Otras técnicas para la purificación de la proteína como el fraccionamiento en una columna de intercambio de iones, precipitación con etanol, HPLC de fase inversa, cromatografía en sílice, cromatografía en una heparina SEPHAROSA™ en una resina de intercambio de aniones o cationes (como una columna de ácido poliaspártico); cromatoenfoque, SDS-PAGE, precipitación con sulfato de amonio también están disponibles dependiendo del anticuerpo que se va a recuperar.

Después de cualquier paso(s) de purificación preliminar(es), la mezcla conteniendo el anticuerpo de interés y los contaminantes se puede someter a una cromatografía de interacción hidrófoba de pH bajo empleando un amortiguador de levigado con un pH entre alrededor de 2.5-

4.5, preferiblemente realizada con unas concentraciones de sal bajitas (por ejemplo, desde alrededor de 0-0.25M sal).

J. Formulaciones Farmacéuticas

Las formulaciones farmacéuticas de los anticuerpos anti-TAT, de los oligopéptidos de unión TAT, TAT ARNsi, de las moléculas orgánicas de unión TAT y/o los polipéptidos TAT empleados de conformidad con la presente invención se preparan para el almacenamiento mezclando el anticuerpo, el polipéptido, el oligopéptido, ARNsi o la molécula orgánica teniendo el grado de pureza deseada con unos excipientes opcionales farmacéuticamente aceptables o con estabilizantes (Remington's Pharmaceutical Sciences Edición 16, Osol, A. Ed. (1980)), en forma de formulaciones liofilizadas o de soluciones acuosas. Los excipientes aceptables o los estabilizantes no son tóxicos para los receptores a las dosis y concentraciones empleadas, e incluyen soluciones amortiguadoras como acetato, Tris, fosfato, citrato y otros ácidos orgánicos; antioxidantes incluyendo ácido ascórbico y metionina; conservantes (como cloruro de octadecildimetilbencil amonio; cloruro de hexametonio; cloruro de benzalconio, cloruro de bencetonio; fenol, butilo o bencil alcohol; alquilo parabenos como metil o propil parabeno; catechol; resorcinol; ciclohexanol; 3-pentanol y m-cresol); polipéptidos de peso molecular bajo (menos de alrededor de 10 residuos); proteínas como seroalbúmina, gelatina o inmunoglobulinas; polímeros hidrofílicos como polivinilpirrolidona; aminoácidos como glicina, glutamina, asparagina, histidina, arginina o lisina; monosacáridos, disacáridos y otros carbohidratos incluyendo glucosa, manosa o dextrinas; agentes quelantes como EDTA; tonificantes como trehalosa y cloruro de sodio; azúcares como sacarosa, mannitol, trehalosa o sorbitol; surfactantes como polisorbatos; contraiones para la formación de sal como sodio, complejos de metales (por ejemplo, complejos Zn-proteínas) y/o surfactantes no iónicos como TWEEN®, PLURONICS® o polietilenglicol (PEG). El anticuerpo preferiblemente comprende el anticuerpo a una concentración de entre alrededor de 5- 200 mg/ml, preferiblemente entre 10-100 mg/ml.

Las formulaciones en la presente también pueden contener más de un compuesto activo, según sea necesario para la indicación específica que se está tratando, preferiblemente aquellos con actividades complementarias que no afectan adversamente la una a la otra. Por ejemplo, además de un anticuerpo anti-TAT, un oligopéptido de unión TAT, TAT ARNsi o una molécula orgánica de unión TAT, puede ser deseable incluir en la formulación, un anticuerpo, por ejemplo, un segundo anticuerpo anti TAT que fija un epítopo diferente en el polipéptido TAT o un anticuerpo a algún otro blanco como un factor de crecimiento que afecta el crecimiento de ese cáncer específico. Alternativamente o adicionalmente, la composición puede comprender además un agente quimioterapéutico, un agente citotóxico, citosina, un agente inhibidor de crecimiento, agente anti-hormonal y/o cardioprotector. Dichas moléculas se encuentran presentes de forma apropiada y combinada, en cantidades efectivas para el propósito buscado.

Los ingredientes activos también se pueden comprimir en microcápsulas preparadas por ejemplo, por medio de técnicas de coacervación o mediante polimerización interfacial (por ejemplo, microcápsulas de hidroximetilcelulosa o microcápsulas de gelatina y poli-(metilmetacilato), respectivamente), en sistemas coloidales de administración de fármacos (por ejemplo, liposomas, microesferas de albúmina, microemulsiones, nano-partículas y nanocápsulas) o en macroemulsiones. Dichas técnicas se divulgan en Remington's Pharmaceutical Sciences, Edición 16, Oslo, A., Ed., (1980).

Se pueden manufacturar preparaciones de liberación sostenida. Ejemplos apropiados de preparaciones de liberación sostenida incluyen matrices semipermeables de polímeros hidrófobos sólidos que contienen el anticuerpo, matrices que tienen formas de artículos con forma, por ejemplo, películas o microcápsulas. Los ejemplos de matrices de liberación sostenida incluyen poliésteres, hidrogeles (por ejemplo, poli(2-hidroxietil-metacrilato) o poli(vinilalcohol)), polilactidas (Patente de Estados Unidos No. 3,773,919), copolímeros de L-ácido glutámico y γ etil-L-glutamato,

etilen-vinil acetato no-degradable, copolímeros de ácido láctico – ácido glicólico degradables LUPRON DEPOT® (microesferas inyectables compuestas por copolímeros de ácido láctico – ácido glicólico y acetato de leuprolido) y poli-D-(-)-3- ácido hidroxibutírico.

Las formulaciones que se emplearán para la administración *in vivo* deben ser estériles. Esto se logra fácilmente mediante filtración a través de unas membranas de filtrado estériles.

K. Diagnóstico y Tratamiento con Anticuerpos anti-TAT, oligopéptidos de Unión TAT, TAT ARNsi y Moléculas Orgánicas de Unión TAT

Para determinar la expresión TAT en el cáncer, existen varios análisis de diagnóstico disponibles. En una incorporación, la sobreexpresión del polipéptido TAT se puede analizar mediante inmunohistoquímica (IHC). Unas secciones de tejido incluidas en parafina de una biopsia de tumor se pueden someter al análisis IHC y se les pueden asignar unos criterios de intensidad de coloración de proteína TAT como sigue:

Puntaje 0 – no se observa ninguna coloración o se observa una coloración de membrana inferior al 10% de las células de tumor.

Puntaje 1+ - se detecta una leve/casi imperceptible coloración de membrana en más del 10% de las células de tumor. Las células únicamente tienen coloración en parte de su membrana.

Puntaje 2+ - se observa una coloración tenue a moderada de toda la membrana en más del 10% de las células de tumor.

Puntaje 3+ - se observa una coloración moderada a fuerte de toda la membrana en más del 10% de las células de tumor.

Aquellos tumores con puntajes de 0 o 1+ para la expresión del polipéptido TAT se pueden caracterizar como que no sobreexpresan el TAT, mientras que aquellos tumores con puntajes de 2+ o 3+ se pueden caracterizar como sobreexpresando el TAT.

Alternativamente o adicionalmente, los análisis FISH como el INFORM® (vendido por Ventana, Arizona) o PATHVISION® (Vysis, Illinois) se pueden efectuar en tejido de tumor fijado con formalina con

inclusión de parafina para determinar el grado (en caso de existir) de la sobreexpresión TAT en el tumor.

La sobreexpresión o amplificación del TAT se puede evaluar empleando un análisis de diagnóstico *in vivo*, por ejemplo, administrando una molécula (como un anticuerpo, un oligopéptido o la molécula orgánica) que fija la molécula que se va a detectar y se marca con un marcador detectable (por ejemplo, un isótopo radioactivo o un marcador fluorescente) y externamente se escanea el paciente para localizar el marcador.

Como se describe arriba, los anticuerpos anti-TAT, los oligopéptidos y las moléculas orgánicas de la invención tienen varias aplicaciones no terapéuticas. Los anticuerpos anti-TAT, los oligopéptidos y las moléculas orgánicas de la presente invención pueden ser útiles para el diagnóstico y la estadificación de las neoplasias expresando el polipéptido TAT (por ejemplo, en radioimagenología). Los anticuerpos, los oligopéptidos y las moléculas orgánicas también son útiles para la purificación o inmunoprecipitación del polipéptido TAT desde las células, para la detección y cuantificación del polipéptido TAT *in vitro*, por ejemplo, en una prueba de ELISA o de inmunotransferencia, para matar y eliminar las células expresando TAT de una población de células combinadas en un paso en la purificación de otras células.

Actualmente, dependiendo de la estadificación del cáncer, el tratamiento del cáncer comprende una o una combinación de las siguientes terapias: cirugía para extirpar el tejido canceroso, radioterapia y quimioterapia. La terapia de anticuerpo anti TAT, de oligopéptido, ARNsi o de molécula orgánica puede ser especialmente deseable en pacientes mayores que no toleran la toxicidad y los efectos secundarios de la quimioterapia bien, y en enfermedades metastásicas donde la radioterapia tiene un uso limitado. Los anticuerpos anti-TAT, los oligopéptidos, ARNsi y las moléculas orgánicas de la invención que apuntan al tumor son útiles para aliviar las neoplasias que expresan TAT en el momento del diagnóstico inicial o durante la recidiva. Para las aplicaciones terapéuticas, el anticuerpo anti TAT, el oligopéptido, ARNsi o la molécula orgánica se

pueden utilizar solos o en una terapia de combinación por ejemplo con hormonas, antiangiogénicos o con compuestos radiomarcados o con cirugía, crioterapia y/o radioterapia. El tratamiento con anticuerpo anti TAT, oligopéptido, ARNsi o molécula orgánica se puede administrar conjuntamente con otras formas de terapia convencional, bien sea consecutivamente con, pre o post terapia convencional. Los fármacos quimioterapéuticos como TAXOTERE® (docetaxel), TAXOL® (palictaxel), estramustina y mitoxantrona se utilizan para el tratamiento del cáncer, en particular, en pacientes de poco riesgo. En el método actual de la invención para tratar o aliviar el cáncer, se puede administrar al paciente con cáncer el anticuerpo anti TAT, el oligopéptido, ARNsi o la molécula orgánica conjuntamente con el tratamiento con uno o varios de los agentes quimioterapéuticos anteriores. En particular, se contempla la terapia de combinación con palictaxel y derivados modificados (consultar por ejemplo, EP0600517). El anticuerpo anti TAT, el oligopéptido, ARNsi o la molécula orgánica será administrado con una dosis terapéuticamente efectiva del agente quimioterapéutico. En otra incorporación, el anticuerpo anti TAT, el oligopéptido, ARNsi o la molécula orgánica se administra conjuntamente con quimioterapia para intensificar la actividad y la eficacia del agente quimioterapéutico, por ejemplo, paclitaxel. La Guía de Referencia Médica (PDR) divulga dosis de estos agentes que se han empleado para el tratamiento de diversas neoplasias. El régimen efectivo de dosis y las dosis de estos fármacos quimioterapéuticos va a depender del cáncer en particular que se está tratando, del grado de la enfermedad y de otros factores conocidos para un galeno con destreza en el arte y el galeno los puede determinar.

En una incorporación en particular, un conjugado comprendiendo un anticuerpo anti-TAT, un oligopéptido, ARNsi o la molécula orgánica conjugado con un agente citotóxico se administra al paciente. Preferiblemente, la célula internaliza el inmunoconjugado fijado a la proteína TAT, resultando en una mayor eficacia terapéutica del inmunoconjugado para matar la célula cancerosa a la cual se fija. En una

GEN 683

incorporación preferida, el agente citotóxico apunta a o interfiere con el ácido nucleico en la célula cancerosa. Ejemplos de dichos agentes citotóxicos se describen arriba e incluyen maitansinoides, calicheamicinas, ribonucleasas y ADN endonucleasas.

Los anticuerpos anti-TAT, los oligopéptidos, las moléculas orgánicas o los conjugados de toxinas de los mismos se administran a un paciente humano, de acuerdo con métodos conocidos, como mediante la administración intravenosa, por ejemplo, como una inyección rápida o mediante una infusión continua durante un periodo de tiempo, mediante unas vías de administración intramuscular, intraperitoneal, intracerobroespinal, subcutánea, intra-articular, intrasinovial, intratecal, oral, tópica o por inhalación. Se prefiere la administración intravenosa o subcutánea del anticuerpo, oligopéptido o de la molécula orgánica.

Otros regímenes terapéuticos se pueden combinar con la administración del anticuerpo anti TAT, del oligopéptido o de la molécula orgánica. La administración combinada incluye la co-administración, empleando formulaciones separadas o una sola formulación farmacéutica y la administración consecutiva en cualquier orden, donde preferiblemente existe un periodo de tiempo durante el cual ambos (o todos) los agentes activos ejercen sus actividades biológicas en forma simultánea. Preferiblemente la citada terapia combinada resulta en un efecto terapéutico sinérgico.

Igualmente puede ser deseable combinar la administración del anticuerpo anti TAT o los anticuerpos, oligopéptidos o moléculas orgánicas, con la administración de un anticuerpo dirigido contra otro antígeno de tumor asociado con el cáncer específico.

En otra incorporación, los métodos del tratamiento terapéutico de la presente invención comprenden la administración combinada de un anticuerpo (o anticuerpos) anti-TAT, oligopéptidos o moléculas orgánicas y uno o varios agentes quimioterapéuticos o agentes inhibidores de crecimiento, incluyendo la co-administración de cocteles de diferentes agentes quimioterapéuticos. Los agentes quimioterapéuticos incluyen

fosfato de estramustina, prednimustina, cisplatina, 5-fluorouracilo, melfalano, ciclofosfamida, hidroxiurea e hidroxiureataxanos (como paclitaxel y doxetaxel) y/o antibióticos antraciclinos. La preparación y los horarios de dosis de dichos agentes quimioterapéuticos se pueden emplear de acuerdo con las instrucciones del fabricante o según lo determine de forma empírica el médico tratante con destrezas. La preparación y los horarios de dosis de dichos agentes quimioterapéuticos también se describen en Chemotherapy Service Ed., M.C. Perry, Williams & Wilkins, Baltimore, MD (1992).

El anticuerpo, el oligopéptido o la molécula orgánica se puede combinar con un compuesto anti-hormonal; por ejemplo, un compuesto anti-estrógeno como tamoxifeno; un anti-progesterona como onapristona (consultar, EP 616 812); o un anti-andrógeno como flutamida, en dosis conocidas para dichas moléculas. Cuando el cáncer que se va a tratar es un cáncer independiente de andrógeno, previamente se puede someter el paciente a una terapia anti-andrógena y una vez el cáncer se convierte en un cáncer independiente de andrógeno, se puede administrar el anticuerpo anti TAT, oligopéptido o la molécula orgánica (y opcionalmente otros agentes como se describe en la presente) al paciente.

Algunas veces, puede ser benéfico coadministrar también un cardioprotector (para prevenir o evitar una disfunción del miocardio asociada a la terapia) o una o varias citosinas al paciente. Además de los regímenes terapéuticos descritos arriba, se puede someter al paciente a la extracción quirúrgica de las células cancerosas y/o a radioterapia, antes, simultáneamente con o posterior a la terapia con anticuerpo, oligopéptido o molécula orgánica. Unas dosis apropiadas para cualquiera de los agentes anteriores coadministrados son aquellas empleadas en la actualidad y se pueden reducir debido a la acción combinada (sinergia) del agente y anticuerpo anti TAT, oligopéptido o la molécula orgánica.

Para prevenir o tratar la enfermedad, la dosis y el modo de administración serán decididos por el galeno de acuerdo con los criterios conocidos. Las dosis apropiadas del anticuerpo, oligopéptido o de la

molécula orgánica dependerán del tipo de enfermedad que se va a tratar, como se define arriba, de la severidad y del curso de la enfermedad, de si el anticuerpo, oligopéptido o la molécula orgánica se administra con fines preventivos o terapéuticos, como terapia previa, de la historia clínica del paciente y de la respuesta al anticuerpo, oligopéptido o a la molécula orgánica y de la discreción del galeno tratante. El anticuerpo, oligopéptido o la molécula orgánica se administra en forma apropiada al paciente una vez o durante una serie de tratamientos. Preferiblemente, el anticuerpo, oligopéptido o la molécula orgánica se administra mediante infusión intravenosa o por inyecciones subcutáneas. Dependiendo del tipo y de la severidad de la enfermedad, alrededor de 1 μg/kg hasta alrededor de 50 mg/kg peso corporal (por ejemplo, alrededor de 0.1-15mg/kg/dosis) del anticuerpo puede ser un candidato a dosis inicial para la administración al paciente, aunque sea por ejemplo, mediante una o varias administraciones separadas o por infusión contínua. Un régimen de dosis puede comprender administrar una dosis de carga inicial de alrededor de 4 mg/kg, seguida por una dosis semanal de mantenimiento de alrededor de 2 mg/kg del anticuerpo anti TAT. Sin embargo, otros regímenes de dosis pueden ser útiles. Una dosis diaria típica puede estar en el rango de alrededor de 1 μg/kg a 100 mg/kg o más, dependiendo de los factores citados arriba. Para repeticiones de administraciones durante varios días o por periodos más largo, dependiendo de la condición, el tratamiento se sostiene hasta que se presente una supresión deseada de los síntomas de la enfermedad. El progreso de esta terapia fácilmente se puede monitorear mediante métodos convencionales y análisis y con base en los criterios conocidos por el galeno u otras personas con destreza en el arte.

Además de la administración de la proteína anticuerpo al paciente, la presente solicitud contempla la administración del anticuerpo mediante la terapia por genes. Dicha administración de ácido nucleico codificando el anticuerpo se comprende en la expresión "administrando una cantidad terapéuticamente efectiva de un anticuerpo". Consultar por ejemplo,

WO96/07321 publicada el 14 de marzo de 1996 concerniente al uso de la terapia de genes para generar anticuerpos intracelulares.

Existen dos grandes enfoques para llevar el ácido nucleico (opcionalmente contenido en un vector) a las células del paciente; *in vivo* y *ex vivo*. Para la administración *in vivo* el ácido nucleico se inyecta directamente al paciente, normalmente en el sitio donde se requiere el anticuerpo. Para el tratamiento *ex vivo* se extraen las células del paciente, se introduce el ácido nucleico a estas células aisladas y las células modificadas se administran al paciente ya sea directamente o por ejemplo, encapsuladas en membranas porosas que se implantan en el paciente (consultar por ejemplo, Patentes de Estados Unidos Nos. 4,892,538 y5,283,187). Existe una variedad de técnicas disponibles para introducir los ácidos nucleicos en células viables. Las técnicas varían dependiendo de si el ácido nucleico se transfiere a células cultivadas *in vitro* o *in vivo* en las células del hospedero esperado. Las técnicas apropiadas para la transferencia de ácido nucleico a células de mamíferos *in vitro* incluyen el uso de liposomas, electroporado, microinyección, fusión celular, DEAE-dextrano, el método de precipitación de fosfato de calcio, etc. Un vector comúnmente empleado para la administración *ex vivo* del gen es un vector retroviral.

Las técnicas de transferencia de ácido nucleico *in vivo* actualmente preferidas incluyen la transinfección con vectores virales (como adenovirus, virus de Herpes simplex I o virus adeno-asociados) y sistemas a base de lípidos (lípidos útiles para la transferencia mediada por lípido del gen son DOTMA, DOPE y DC-Chol, por ejemplo). Para revisar los protocolos de la marcación de genes conocida en la actualidad y la terapia de genes, consultar Anderson y colaboradores, Science 256:808-813 (1992). Igualmente consultar WO 93/25673 y las referencias citadas en las mismas.

Los anticuerpos anti-TAT de la invención pueden encontrarse en diferentes formas comprendidas en la definición de "anticuerpo" en la presente. De este modo, los anticuerpos incluyen anticuerpos de longitud

total o intactos, fragmentos de anticuerpos, anticuerpos de secuencia natural o variantes de aminoácido, anticuerpos humanizados, quiméricos o de fusión, inmunoconjugados y fragmentos funcionales de los mismos. En los anticuerpos de fusión, una secuencia de anticuerpos se fusiona a una secuencia de polipéptido heterólogo. Los anticuerpos se pueden modificar en la región Fc para proveer las funciones efectoras deseadas. Como se discute en mayor detalle en las secciones de la presente, con las regiones Fc apropiadas, el anticuerpo desnudo fijado en la superficie de la célula puede inducir citotoxicidad, por ejemplo, mediante una citotoxicidad celular dependiente de anticuerpo (ADCC) o mediante el reclutamiento del complemento de la citotoxicidad complemento dependiente o por cualquier otro mecanismo. Alternativamente, cuando es deseable eliminar o reducir la función efectora con el fin de minimizar los efectos secundarios o las complicaciones terapéuticas, se pueden emplear ciertas otras regiones Fc.

En una incorporación, el anticuerpo compite por la fijación o sustancialmente se fija al mismo epítopo que los anticuerpos de la invención. Los anticuerpos teniendo las características biológicas de los actuales anticuerpos anti-TAT de la invención también se contemplan, específicamente incluyendo determinar el tumor como objetivo *in vivo* y cualquier inhibición de proliferación celular o características citotóxicas.

En la presente se describen en detalle los métodos para producir los anticuerpos descritos arriba.

Los actuales anticuerpos anti-TAT, los oligopéptidos, ARNsi y las moléculas orgánicas son útiles para tratar un cáncer que expresa TAT o para aliviar uno o varios síntomas del cáncer en un mamífero. Dicho cáncer incluye cáncer de próstata, cáncer del trato urinario, cáncer de pulmón, cáncer de mama, cáncer de colon, cáncer colorrectal y cáncer ovárico, más específicamente, adenocarcinoma de próstata, carcinomas de células renales, adenocarcinomas colorrectales, adenocarcinomas de pulmón, carcinomas escamosos de pulmón y mesotelioma pleural. Las neoplasias comprenden las neoplasias metastásicas de cualquiera de los anteriores. El anticuerpo, oligopéptido, ARNsi o la molécula orgánica se puede fijar por

lo menos a una porción de las células cancerosas que expresan el polipéptido TAT en el mamífero. En una incorporación preferida, el anticuerpo, oligopéptido, ARNsi o la molécula orgánica es efectivo para destruir o matar las células tumorales expresando TAT o inhibe el crecimiento de dichas células tumorales, *in vitro* o *in vivo*, cuando se fija al polipéptido TAT en la célula. Dicho anticuerpo incluye un anticuerpo anti TAT desnudo (sin conjugar a ningún agente). Los anticuerpos desnudos que tienen propiedades citotóxicas o que inhiben el crecimiento celular se pueden complementar además con un agente citotóxico para volverlos aún más potentes en la destrucción de las células tumorales. Las propiedades citotóxicas se pueden conferir a un anticuerpo anti-TAT, por ejemplo, al conjugar el anticuerpo con un agente citotóxico, para formar un inmunoconjugado como se describe en la presente. El agente citotóxico o un agente inhibidor de crecimiento es preferiblemente una molécula pequeña. Se prefieren las toxinas como calicheamicina o un maitansinoide y análogos o derivados de los mismos.

La invención dispone una composición comprendiendo un anticuerpo anti-TAT, un oligopéptido, ARNsi o la molécula orgánica de la invención y un excipiente. Para el propósito de tratar el cáncer, las composiciones se pueden administrar al paciente que requiera el citado tratamiento, donde la composición puede tener presente uno o varios anticuerpos anti-TAT como un inmunoconjugado o como el anticuerpo desnudo. En otra incorporación, las composiciones pueden comprender estos anticuerpos, oligopéptidos, ARNsi o moléculas orgánicas en combinación con otros agentes terapéuticos como citotóxicos o agentes inhibidores de crecimiento, incluyendo agentes quimioterapéuticos. La invención también dispone formulaciones comprendiendo un anticuerpo anti-TAT, un oligopéptido, ARNsi o la molécula orgánica de la invención y un excipiente. En una incorporación, la formulación es una formulación terapéutica comprendiendo un excipiente farmacéuticamente aceptable.

Otro aspecto de la invención son los ácidos nucleicos aislados codificando los anticuerpos anti-TAT. Los ácidos nucleicos codificando

tanto la cadena H como la cadena L y especialmente los residuos de la región hipervariable, las cadenas que codifican la secuencia natural del anticuerpo al igual que las variantes, las modificaciones y versiones humanizadas del anticuerpo, se comprenden.

La invención igualmente dispone métodos útiles para tratar un cáncer expresando el polipéptido TAT o aliviando uno o varios síntomas del cáncer en un mamífero, que comprende administrar una cantidad terapéuticamente efectiva de un anticuerpo anti-TAT, un oligopéptido, ARNsi o la molécula orgánica al mamífero. Las composiciones terapéuticas del anticuerpo, oligopéptido o la molécula se pueden administrar en corto plazo (agudo) o en forma crónica o intermitente según lo ordene el galeno. Igualmente se disponen métodos para inhibir el crecimiento de y para matar una célula que expresa el polipéptido TAT.

La invención igualmente dispone estuches y artículos de manufactura comprendiendo por lo menos un anticuerpo anti TAT, oligopéptido, ARNsi o la molécula orgánica. Los estuches conteniendo anticuerpos anti-TAT, un oligopéptido, ARNsi o moléculas orgánicas encuentran uso, por ejemplo, en análisis para matar las células TAT, para la purificación o inmunoprecipitación del polipéptido TAT de las células. Por ejemplo, para aislar y purificar el TAT, el estuche puede contener un anticuerpo anti-TAT, un oligopéptido, ARNsi o la molécula orgánica con acoplamiento a glóbulos (por ejemplo, glóbulos de sefarosa). Se pueden disponer estuches que contienen los anticuerpos, oligopéptidos, ARNsi o las moléculas orgánicas para la detección y la cuantificación de TAT in vitro, por ejemplo, en una prueba de ELISA o de inmunotransferencia. Dichos anticuerpos, oligopéptidos, ARNsi o moléculas orgánicas útiles para la detección se pueden disponer con un marcador como un marcador fluorescente o un radiomarcador.

L. Artículos de Manufactura y Estuches

Otra incorporación de la invención es un artículo de manufactura que contiene materiales útiles para el tratamiento de un anti-TAT expresando cáncer. El artículo de manufactura comprende un envase y una etiqueta o

un inserto en el paquete en o asociado con el envase. Unos envases apropiados incluyen, por ejemplo, botellas, ampollas, jeringas, etc. Los envases se pueden formar a partir de una variedad de materiales como vidrio o plástico. El envase contiene una composición que es efectiva para tratar el cáncer y puede tener un puerto de acceso estéril (por ejemplo el envase puede ser una bolsa con una solución intravenosa o una ampolla con un tapón que se puede perforar con una aguja de una jeringa hipodérmica). Por lo menos un agente activo en la composición es un anticuerpo anti TAT, un oligopéptido, ARNsi o la molécula orgánica de la invención. La etiqueta o el inserto indican que la composición se utiliza para tratar el cáncer. La etiqueta o el inserto comprenderán además instrucciones para administrar el anticuerpo, oligopéptido, ARNsi o la molécula orgánica al paciente con cáncer. Adicionalmente, el artículo de manufactura puede comprender un segundo envase que comprende una solución amortiguadora farmacéuticamente aceptable, como agua bacterioestática para inyección (BWFI), solución salina amortiguada con fosfato, solución de Ringer y solución de dextrosa. Adicionalmente puede incluir otros materiales deseables desde el punto de vista comercial y del usuario, incluyendo otras soluciones amortiguadoras, diluyentes, filtros, agujas y jeringas.

También se disponen estuches que son útiles para varios propósitos, por ejemplo, para analizar la muerte de las células expresando TAT, para la purificación o inmunoprecipitación del polipéptido TAT de las células. Para aislar y purificar el polipéptido TAT, el estuche puede contener un anticuerpo anti-TAT, un oligopéptido, ARNsi o la molécula orgánica con acoplamiento a glóbulos (por ejemplo, glóbulos de sefarosa). Se pueden disponer estuches que contienen los anticuerpos, oligopéptidos, ARNsi o las moléculas orgánicas para la detección y cuantificación del polipéptido TAT *in vitro*, por ejemplo, en una prueba de ELISA o de inmunotransferencia. Al igual que con el artículo de manufactura, el estuche comprende un envase y una etiqueta o un inserto de paquete en o asociado al envase. El envase contiene una composición comprendiendo por lo menos un anticuerpo anti TAT, oligopéptido, ARNsi o una molécula orgánica de la

invención. Se pueden incluir envases adicionales que contienen por ejemplo, diluyentes y soluciones amortiguadoras y anticuerpos de control. La etiqueta o el inserto del paquete pueden disponer una descripción además de las instrucciones para el uso propuesto *in vitro* o el uso diagnóstico.

M. Usos de los Polipéptidos TAT y los Ácidos Nucleicos Codificando el Polipéptido TAT

Las secuencias de nucleótido (o su complemento) codificando los polipéptidos TAT tienen varias aplicaciones en el arte de la biología molecular, incluyendo usos como sondas de hibridación, en el mapeo de cromosomas y genes y en la generación de ARN antisentido, ARNsi y sondas ADN. El ácido nucleico codificando TAT también será útil para la preparación de los polipéptidos TAT mediante las técnicas recombinantes descritas en la presente, donde esos polipéptidos TAT pueden encontrar un uso, por ejemplo, en la preparación de anticuerpos anti-TAT como se describe en la presente.

El gen de longitud total de secuencia natural del TAT o porciones del mismo se pueden emplear como sondas de hibridación para una biblioteca ADNc para aislar el ADNc del TAT de longitud total o para aislar otros ADNc (por ejemplo, aquellos codificando variantes naturales de TAT o los TAT de otras especies) que tienen una secuencia de identidad deseada frente a la secuencia natural del TAT divulgados en la presente. Opcionalmente, la longitud de las sondas será desde alrededor de 20 hasta alrededor de 50 bases. Las sondas de hibridación se pueden derivar por lo menos parcialmente de regiones novedosas de la secuencia del nucleótido natural de longitud total donde esas regiones se pueden determinar sin una experimentación indebida o de secuencias genómicas incluyendo promotores, elementos potenciadores e intrones de TAT de secuencia natural. A modo de ejemplo, un método de cribado va a comprender aislar la región de codificación del gen TAT utilizando la secuencia conocida de ADN para sintetizar una sonda seleccionada de alrededor de 40 bases. Las sondas de hibridación se pueden marcar mediante una variedad de

marcadores, incluyendo radionucleótidos como ^{32}P o ^{35}S o marcadores enzimáticos como fosfatasa alcalina con acoplamiento a la sonda mediante sistemas de acoplamiento avidina/biotina. Las sondas marcadas teniendo una complementariedad de secuencia con la del gen TAT de la presente invención se pueden emplear para cribar bibliotecas de ADNc humano, ADN genómico o ARNm para determinar a cuales miembros de dichas bibliotecas se hibrida la sonda. Las técnicas de hibridación se describen con mayor detalle en los Ejemplos abajo. Cualquier secuencia EST divulgada en la presente solicitud se puede utilizar en la presente como sondas empleando los métodos divulgados en la presente.

Otros fragmentos útiles de los ácidos nucleicos codificando el TAT incluyen oligonucleótidos antisentido o sentido comprendiendo una secuencia de ácido nucleico de una sola hebra (bien sea ARN o ADN) capaz de fijarse a las secuencias del ARNM del TAT (sentido) objetivo o del ADN del TAT (antisentido). Los oligonucleótidos sentido o antisentido, de acuerdo con la presente invención, comprenden un fragmento de la región de codificación del ADN del TAT. Dicho fragmento generalmente comprende por lo menos alrededor de 14 nucleótidos, preferiblemente de alrededor de 14 a 30 nucleótidos. La habilidad de derivar un oligonucleótido antisentido o sentido, con base en una secuencia de ADNc codificando determinada proteína se describe en, por ejemplo, Stein and Cohen (Cancer Res. 48:2659, 1988) y van der Krol y colaboradores (BioTechniques 6:958, 1988).

La fijación de oligonucleótidos antisentido o sentido para apuntar a las secuencias de ácido nucleico resulta en la formación de duplicidades que bloquean la transcripción o la traslación de la secuencia objetivo mediante una de varias formas, incluyendo un degradación mejorada de las duplicidades, la terminación prematura de la transcripción o traslación o mediante otras formas. La presente invención abarca estos métodos. De este modo se pueden utilizar los oligonucleótidos antisentido para bloquear la expresión de las proteínas TAT, donde aquellas proteínas TAT pueden jugar un papel en la inducción del cáncer en los mamíferos. Los

oligonucleótidos antisentido o sentido comprenden además los oligonucleótidos que tienen una columna vertebral modificada por azúcar - fosfodiéster (u otros ligamientos de azúcar como los descritos en WO 91/06629) y donde dichos ligamientos de azúcar son resistentes a las nucleasas endógenas. Dichos oligonucleótidos con ligamientos de azúcar resistentes son estables *in vivo* (es decir, son capaces de resistir la degradación enzimática) pero conservan la especificidad de la secuencia de poderse fijar a las secuencias de nucleótidos objetivo.

Los sitios intragénicos preferidos para la fijación antisentido incluyen la región que incorpora el codón de iniciación/comienzo de traslación (5'-AUG / 5'-ATG) o el codón de terminación/parada (5'-UAA, 5'-UAG y 5-UGA / 5'-TAA, 5'-TAG y 5'-TGA) del marco de lectura abierto (ORF) del gen. Estas regiones hacen referencia a una porción del ARNm o del gen que comprende desde alrededor de 25 hasta alrededor de 50 nucleótidos contiguos en cualquier dirección (es decir, 5' o 3') de un codón de iniciación o terminación de traslación. Otras regiones preferidas para la fijación antisentido incluyen: intrones; exones; empalmes intrón-exón; el marco de lectura abierto (ORF) o "región de codificación" que es la región entre el codón de iniciación de la traslación y el codón de terminación de la traslación; la caperuza 5' de un ARNm que comprende un residuo guanosino N7-metilado unido al residuo más orientado al extremo 5' del ARNm mediante un ligamiento 5'-5' trifosfato e incluye una estructura de caperuza 5' al igual que los primeros 50 nucleótidos adyacentes al extremo; la región 5' sin trasladar (5'UTR), la porción de un ARNm en la dirección 5' del codón de iniciación de la traslación incluyendo de este modo los nucleótidos entre el sitio de la caperuza 5' y el codón de iniciación de traslación de un ARNm o los nucleótidos correspondientes en el gen y la región sin trasladar 3' (3'UTR), la porción de un ARNm en la dirección 3' del codón de terminación de la traslación y incluyendo de este modo los nucleótidos entre el codón de terminación de traslación y la caperuza 3' de un ARNm o los nucleótidos correspondientes en el gen.

Ejemplos específicos de compuestos antisentido preferidos útiles para inhibir la expresión de las proteínas TAT incluyen oligonucleótidos conteniendo columnas vertebrales modificadas de ligamientos o ligamientos internucleósidos no naturales. Los oligonucleótidos teniendo columnas vertebrales modificadas incluyen aquellos que retienen un átomo fosfórico en la columna vertebral y aquellos que no tienen un átomo fosfórico en la columna vertebral. Para los fines de esta especificación y como a veces se referencia en el arte, los oligonucleótidos modificados que no tienen un átomo fosfórico en su columna vertebral internucleósida también se pueden considerar oligonucleósidos. Las columnas vertebrales de oligonucleótido preferidos incluyen, por ejemplo, fosforotioatos, fosforotioatos quirales, fosforoditioatos, fosfotriésteres, aminoalquilofosfotri-ésteres, metilo y otros alquilo fosfonatos incluyendo 3'-alquileno fosfonatos, 5'-alquiloeno fosfonatos y fosfonatos quirales, fosfinatos, fosforamidatos incluyendo 3'-amino fosforamidato y aminoalquilofosforamidatos, tionofosforamidatos, tionoalquilofosfonatos, tionoalquilofosfotriésteres, selenofosfatos y borano-fosfatos teniendo ligamientos normales 3'-5', 2'-5' análogos ligados de estos y aquellos teniendo la polaridad invertida donde uno o varios ligamientos de internucleótido son un ligamiento 3' a 3', 5' a 5' o 2' a 2'. Los oligonucleótidos preferidos teniendo una polaridad invertida comprenden un ligamiento sencillo 3' a 3' en el ligamiento más orientado al extremo 3' es decir un residuo nucleósido simple invertido que puede ser abásico (carece de nucleobase o tiene un grupo hidroxilo en lugar de aquél). También se incluyen varias sales, sales combinadas y formas de ácidos libres. Unas patentes representativas de Estados Unidos que enseñan la preparación de ligamientos de contenido fosfórico incluyen, pero no se limitan a, Patentes de Estados Unidos Nos.: 3,687,808; 4,469,863; 4,476,301; 5,023,243; 5,177,196; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361; 5,194,599; 5,565,555;

5,527,899; 5,721,218; 5,672,697 y 5,625,050, cada una de las cuales se incorpora a la presente por referencia.

Las columnas vertebrales modificadas de oligonucleótido preferidas que no incluyen un átomo fosfórico en el mismo, tienen columnas vertebrales formadas por cadenas cortas de ligamientos internucleósidos alquilo o cicloalquilo, ligamientos de heteroátomos combinados e internucleósidos alquilo o cicloalquilo o uno o varios ligamientos internucleósidos heteroatómicos o heterocíclicos de cadena corta. Estos incluyen aquellos que tienen ligamientos morfolinos (formados en parte de la porción de azúcar de un nucleósido); columnas vertebrales de siloxana; columnas vertebrales sulfida, sulfóxida y sulfona; columnas vertebrales formacetilo y tioformacetilo; columnas vertebrales metilen formacetilo y tioformacetilo; columnas vertebrales riboacetilo; columnas vertebrales conteniendo alquenos; columnas vertebrales sulfamato; columnas vertebrales metilenimino y metilenhidracina; columnas vertebrales sulfonato y sulfonamida; columnas vertebrales amida y otras teniendo partes combinadas de N, O, S y CH.sub.2. Unas patentes representativas de Estados Unidos que enseñan la preparación de dichos oligonucleósidos incluyen, pero no se limitan a Patentes de Estados Unidos Nos.: 5,034,506; 5,166,315; 5,185,444; 5,214,134; 5,216,141; 5,235,033; 5,264,562; 5,264,564; 5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677; 5,541,307; 5,561,225; 5,596,086; 5,602,240; 5,610,289; 5,602,240; 5,608,046; 5,610,289; 5,618,704; 5,623,070; 5,663,312; 5,633,360; 5,677,437; 5,792,608; 5,646,269 y 5,677,439, cada una de las cuales se incorpora a la presente por referencia.

En otros oligonucleótidos antisentido preferidos, tanto el ligamiento de azúcar como de internucleósidos, es decir, la columna vertebral de las unidades de nucleótido se reemplazan con grupos novedosos. Las unidades base se preservan para la hibridación con un compuesto de ácido nucleico objetivo apropiado. A uno de los citados compuestos oligoméricos apropiados, un oligonucleótido mimético, que ha mostrado tener unas excelentes propiedades de hibridación, se le conoce como un ácido péptido

nucleico (PNA). En compuestos PNA, la columna vertebral de azúcar de un oligonucleótido se reemplaza con una columna vertebral conteniendo amida, en particular una columna vertebral aminoetilglicina. Las nucleobases se conservan y se fijan directa o indirectamente a átomos de aza nitrógeno de la porción amida de la columna vertebral. Unas patentes representativas de Estados Unidos que enseñan la preparación de compuestos PNA incluyen, pero no se limitan a, Patentes de Estados Unidos Nos.: 5,539,082; 5,714,331 y 5,719,262, cada una de las cuales se incorpora a la presente por referencia. Se pueden encontrar otras enseñanzas sobre los compuestos PNA en Nielsen y colaboradores, Science, 1991, 254, 1497-1500.

Los oligonucleótidos antisentido preferidos incorporan columnas vertebrales de fosforotioato y/o columnas vertebrales de heteroátomo y en particular $-CH_2-NH-O-CH_2-$, $-CH_2-N(CH_3)-O-CH_2-$ [conocidos como un metileno (metilimino) o columna vertebral MMI], $-CH_2-O-N(CH_3)-CH_2-$, $-CH_2-N(CH_3)-N(CH_3)-CH_2-$ y $-O-N(CH_3)-CH_2-CH_2-$ [donde la columna vertebral natural del fosfodiéster se representa como $-O-P-O-CH_2-$] descrito en las Patentes de Estados Unidos referenciadas arriba No. 5,489,677 y las columnas vertebrales amida de las Patentes de Estados Unidos referenciadas arriba No. 5,602,240. Otros oligonucleótidos antisentido también preferidos tienen columnas vertebrales de estructuras morfolinas de las Patentes de Estados Unidos referenciadas arriba No. 5,034,506.

Los oligonucleótidos modificados también pueden contener una o varias fracciones de azúcar sustituidas. Los oligonucleótidos preferidos comprenden uno de los siguientes en la posición 2': OH; F; O-alquilo, S-alquilo o N-alquilo; O-alquenilo, S-alquenilo o N-alquenilo; O-alquinilo, S-alquinilo o N-alquinilo; u O-alquilo-O-alquilo, donde el alquilo, alquenilo y alquinilo se pueden sustituir o no sustituir C_1 a C_{10} alquilo o C_2 a C_{10} alquenilo y alquinilo. Particularmente preferidos son $O[(CH_2)_nO]_mCH_3$, $O(CH_2)_nOCH_3$, $O(CH_2)_nNH_2$, $O(CH_2)_nCH_3$, $O(CH_2)_nONH_2$ y $O(CH_2)_nON[(CH_2)_nCH_3)]_2$, donde n y m son desde 1 hasta alrededor de 10. Otros oligonucleótidos antisentido preferidos comprenden uno de los

siguientes en la posición 2': C1 a C10 alquilo menor, alquilo menor sustituido, alquenilo, alquinilo, alcarilo, aralquilo, O-alcarilo u O-aralquilo, SH, SCH_3, OCN, Cl, Br, CN, CF_3, OCF_3, $SOCH_3$, SO_2 CH_3, ONO_2, NO_2, N_3, NH_2, heterocicloalquilo, heterocicloalcarilo, aminoalquilamino, polialquilamino, sililo sustituido, un grupo de segmentación ARN, un grupo indicador, un intercalador, un grupo para mejorar las propiedades farmacocinéticas de un oligonucleótido o un grupo para mejorar las propiedades farmacodinámicas de un oligonucleótido y otros sustituyentes con propiedades similares. Una modificación preferida incluye 2'-metoxietoxi (2'-O-$CH_2CH_2OCH_3$, también conocido como 2'-O-(2-metoxietilo) o 2'-MOE) (Martin y colaboradores, Helv. Chim. Acta, 1995, 78, 486-504) es decir, un grupo alcoxialcoxi. Otra modificación preferida incluye 2'-dimetilaminooxietoxi, es decir, un grupo $O(CH_2)_2ON(CH_3)_2$, también conocido como 2'-DMAOE, como se describe en los Ejemplos a continuación y 2'-dimetilaminoetoxietoxi (también conocidos en el arte como 2'-O-dimetilaminoetoxietil o 2'-DMAEOE), es decir, 2'-O-CH_2-O-CH_2-N(CH_2).

Otra modificación preferida incluye ácidos nucleicos bloqueados (LNAs) donde el grupo 2'-hidroxilo se liga al átomo de carbono 3' o 4' del anillo de azúcar para formar así una fracción bicíclica de azúcar. El ligamiento es preferiblemente un grupo metileno $(-CH_2-)_n$ haciendo un puente entre el átomo de oxígeno 2' y el átomo de carbono 4' donde n es 1 o 2. Los LNAs y la preparación de los mismos se describen en WO 98/39352 y WO 99/14226.

Otras modificaciones preferidas incluyen 2'-metoxi (2'-O-CH_3), 2'-aminopropoxi (2'-$OCH_2CH_2CH_2$ NH_2), 2'-alilo (2'-CH_2-CH=CH_2), 2'-O-alilo (2'-O-CH_2-CH=CH_2) y 2'-fluoro (2'-F). La modificación 2' puede ser en la posición arabino (arriba) o en la posición ribo (abajo). Una modificación preferida 2'-arabino es 2'-F. También se pueden realizar unas modificaciones similares en otras posiciones en el oligonucleótido, particularmente la posición 3' del azúcar en el nucleótido terminal 3' o en los nucleótidos ligados 2'-5' y la posición 5' del nucleótido terminal 5'.

Igualmente, los oligonucleótidos pueden tener miméticos de azúcar como fracciones de ciclobutilo en lugar del azúcar pentofuranosilo. Unas patentes representativas de Estados Unidos que enseñan la preparación de dichas estructuras modificadas de azúcares incluyen, pero no se limitan a, Patentes de Estados Unidos Nos.: 4,981,957; 5,118,800; 5,319,080; 5,359,044; 5,393,878; 5,446,137; 5,466,786; 5,514,785; 5,519,134; 5,567,811; 5,576,427; 5,591,722; 5,597,909; 5,610,300; 5,627,053; 5,639,873; 5,646,265; 5,658,873; 5,670,633; 5,792,747 y 5,700,920, cada una de las cuales se incorpora a la presente por referencia en su totalidad.

Los oligonucleótidos también pueden incluir unas modificaciones o substituciones nucleobase (en el arte con frecuencia se hace referencia a estos simplemente como "base"). Como se utiliza en la presente, nucleobases "sin modificar" o "naturales" incluyen las bases purinas adenina (A) y guanina (G) y las bases pirimidina timina (T), citosina (C) y uracilo (U). Las nucleobases modificadas incluyen otras nucleobases sintéticas y naturales como 5-metilcitosina (5-me-C), 5-hidroximetil citosina, xantina, hipoxantina, 2-aminoadenina, 6-metilo y otros derivados alquilo de adenina y guanina, 2-propilo y otros derivados alquilo de adenina y guanina, 2-tiouracilo, 2-tiotimina y 2-tiocitosina, 5-halouracilo y citosina, 5-propinilo (-C≡C-CH_3 o -CH_2-C≡CH) uracilo y citosina y otros derivados de alquinilo de bases pirimidina, 6-azo uracilo, citosina y timina, 5-uracilo (pseudouracilo), 4-tiouracilo, 8-halo, 8-amino, 8-tiol, 8-tioalquilo, 8-hidroxilo y otras adeninas and guaninas sustituidas-8 , 5-halo particularmente 5-bromo, 5-trifluorometilo y otros uracilos y citosinas sustituidos-5, 7-metilguanina y 7-metiladenina, 2-F-adenina, 2-amino-adenina, 8-azaguanina y 8-azaadenina, 7-deazaguanina y 7-deazaadenina y 3-deazaguanina y 3-deazaadenina. Otras nucleobases modificadas incluyen pirimidinas tricíclicas como fenoxazina citidina(1H-pirimido[5,4-b][1,4]benzoxazin-2(3H)-ona), fenotiazina citidina (1H-pirimido[5,4-b][1,4]benzotiazin-2(3H)-ona), pinzas-G como una fenoxazina citidina sustituida (por ejemplo 9-(2-aminoetoxi)-H-pirimido[5,4-b][1,4]benzoxazin-2(3H)-ona), carbazole citidina (2H-pirimido[4,5-b]indol-

2-ona), piridoindole citidina (H-pirido[3',2':4,5]pirrolo[2,3-d]pirimidin-2-ona). Unas nucleobases modificadas también pueden incluir aquellas donde la base purina o pirimidina se reemplaza con otros heterociclos, por ejemplo 7-deaza-adenina, 7-deazaguanosina, 2-aminopiridina y 2-piridona. Otras nucleobases incluyen aquellas divulgadas en las Patentes de Estados Unidos No. 3,687,808, aquellas divulgadas en The Concise Enciclopedia Of Polimer Science And Engineering, páginas 858-859, Kroschwitz, J. I, ed. John Wiley & Sons, 1990 y aquellas divulgadas por Englisch y colaboradores, Angewandte Chemie, International Edition, 1991, 30, 613. Algunas de estas nucleobases son particularmente útiles para incrementar la afinidad de fijación de los compuestos oligoméricos de la invención. Estas incluyen pirimidinas' 5-sustituidas, 6-azapirimidinas y N-2, N-6 y O-6 purinas sustituidas, incluyendo 2-aminopropiladenina, 5-propiniluracilo y 5-propinilcitosina. Las sustituciones 5-metilcitosina han demostrado incrementar la estabilidad de la duplicidad de ácido nucleico en 0.6-1.2.grados. C. (Sanghvi y colaboradores, Antisense Research and Applications, CRC Press, Boca Raton, 1993, pp. 276-278) y son las sustituciones preferidas de las bases, aún más particularmente cuando se combinan con modificaciones de azúcar 2'-O-metoxietil. Unas patentes representativas de Estados Unidos que enseñan la preparación de las nucleobases modificadas incluyen, pero no se limitan a: Patentes de Estados Unidos No. 3,687,808, al igual que las Patentes de Estados Unidos Nos.: 4,845,205; 5,130,302; 5,134,066; 5,175,273; 5,367,066; 5,432,272; 5,457,187; 5,459,255; 5,484,908; 5,502,177; 5,525,711; 5,552,540; 5,587,469; 5,594,121, 5,596,091; 5,614,617; 5,645,985; 5,830,653; 5,763,588; 6,005,096; 5,681,941 y 5,750,692, cada una de las cuales se incorpora a la presente por referencia.

Otra modificación de los oligonucleótidos antisentido que se ligan químicamente a una o varias fracciones del oligonucleótido o conjugados que intensifican la actividad, la distribución celular o la absorción celular del oligonucleótido. Los compuestos de la invención pueden incluir grupos de conjugados ligados en forma covalente a grupos funcionales como

grupos hidroxilo primarios o secundarios. Los grupos de conjugados de la invención incluyen intercaladores, moléculas indicadoras, poliaminas, poliamidas, polietilenglicoles, poliéteres, grupos que intensifican las propiedades farmacodinámicas de los oligómeros y grupos que intensifican las propiedades farmacocinéticas de los oligómeros. Los grupos típicos de los conjugados incluyen colesteroles, lípidos, catión lípidos, fosfolípidos, fosfolípidos catiónicos, biotina, fenazina, folato, fenantridina, antraquinona, acridina, fluoresceinas, rodaminas, cumarinas y tintes. Los grupos que intensifican las propiedades farmacodinámicas, en el contexto de esta invención, incluyen grupos que mejoran la absorción de oligómero, intensifican la resistencia del oligómero a la degradación y/o a fortalecer la hibridación secuencia-específica con ARN. Los grupos que intensifican las propiedades farmacocinéticas, en el contexto de esta invención, incluyen grupos que mejoran la absorción de oligómero, la distribución, el metabolismo o la excreción. Las fracciones de conjugado incluyen pero no se limitan a fracciones de lípido como una fracción de colesterol (Letsinger y colaboradores, Proc. Natl. Acad. Sci. USA, 1989, 86, 6553-6556), ácido cólico (Manoharan y colaboradores, Bioorg. Med. Chem. Let., 1994, 4, 1053-1060), un tioéter, por ejemplo, hexil-S-tritiltiol (Manoharan y colaboradores, Ann. N. Y. Acad. Sci., 1992, 660, 306-309; Manoharan y colaboradores, Bioorg. Med. Chem. Let., 1993, 3, 2765-2770), un tiocolesterol (Oberhauser y colaboradores, Nucl. Acids Res., 1992, 20, 533-538), una cadena alifática, por ejemplo, residuos de dodecandiol o undecilo (Saison-Behmoaras y colaboradores, EMBO J., 1991, 10, 1111-1118; Kabanov y colaboradores, FEBS Lett, 1990, 259, 327-330; Svinarchuk y colaboradores, Biochimie, 1993, 75, 49-54), un fosfolípido, por ejemplo, di-hexadecil-rac-glicerol o trietil-amonio 1,2-di-O-hexadecil-rac-glicero-3-H-fosfonato (Manoharan y colaboradores, Tetrahedron Lett., 1995, 36, 3651-3654; Shea y colaboradores, Nucl. Ácidos Res., 1990, 18, 3777-3783), una poliamina o una cadena polietilenglicol (Manoharan y colaboradores, Nucleosides & Nucleotides, 1995, 14, 969-973) o ácido acético adamantano (Manoharan y colaboradores, Tetrahedron Lett., 1995, 36, 3651-3654), una

fracción palmitilo (Mishra y colaboradores, Biochim. Biophys. Acta, 1995, 1264, 229-237) o una fracción octadecilamina o hexilamino-carbonil-oxicolesterol. Los oligonucleótidos de la invención también se pueden conjugar a unas sustancias activas del fármaco, por ejemplo, aspirina, warfarina, fenilbutazona, ibuprofeno, suprofeno, fenbufeno, ketoprofeno, (S)-(+)-pranoprofeno, carprofeno, dansilsarcosina, 2,3,5- ácido triiodobenzoico, ácido flufenámico, ácido folínico, una benzotiadiazida, clorotiazida, una diazepina, indometicina, un barbiturato, una cefalosporina, un fármaco sulfa, un antidiabético, un antibacteriano o un antibiótico. Los conjugados oligonucleótido-fármaco y su preparación se describen en la Solicitud de Patente de Estados Unidos Ser. No. 09/334,130 (presentada el 15 de junio de 1999) y las Patentes de Estados Unidos Nos.: 4,828,979; 4,948,882; 5,218,105; 5,525,465; 5,541,313; 5,545,730; 5,552,538; 5,578,717, 5,580,731; 5,580,731; 5,591,584; 5,109,124; 5,118,802; 5,138,045; 5,414,077; 5,486,603; 5,512,439; 5,578,718; 5,608,046; 4,587,044; 4,605,735; 4,667,025; 4,762,779; 4,789,737; 4,824,941; 4,835,263; 4,876,335; 4,904,582; 4,958,013; 5,082,830; 5,112,963; 5,214,136; 5,082,830; 5,112,963; 5,214,136; 5,245,022; 5,254,469; 5,258,506; 5,262,536; 5,272,250; 5,292,873; 5,317,098; 5,371,241, 5,391,723; 5,416,203, 5,451,463; 5,510,475; 5,512,667; 5,514,785; 5,565,552; 5,567,810; 5,574,142; 5,585,481; 5,587,371; 5,595,726; 5,597,696; 5,599,923; 5,599,928 y 5,688,941, cada una de las cuales se incorpora a la presente por referencia.

No es necesario que todas las posiciones en un compuesto determinado se modifiquen de manera uniforme, y de hecho más de unas de las modificaciones citadas anteriormente se puede incorporar en un solo compuesto o incluso en un solo nucleósido en un oligonucleótido. La presente invención también incluye compuestos antisentido que son compuestos quiméricos. Los compuestos "quiméricos" antisentido o las "quimeras" en el contexto de esta invención, son compuestos antisentido, particularmente oligonucleótidos, que contienen dos o más regiones químicamente diferenciadas, cada una conformada por lo menos por una

unidad monómero, es decir, un nucleótido en el caso de un compuesto oligonucleótido. Estos oligonucleótidos típicamente contienen por lo menos una región donde el oligonucleótido se modifica con el fin de conferir al oligonucleótido una mayor resistencia a la degradación nucleasa, aumentar la absorción celular y/o incrementar la afinidad de fijación para el ácido nucleico objetivo. Una región adicional del oligonucleótido puede servir como un sustrato para las enzimas capaces de segmentación ARN: ADN o híbridos ARN: ARN. A modo de ejemplo, RNase H es una endonucleasa celular que segmenta la hebra ARN de una duplicidad ARN:ADN. La activación de RNase H, por lo tanto, resulta en la segmentación del objetivo ARN, en esta forma intensifica enormemente la eficiencia de la inhibición de oligonucleótido de la expresión del gen. Consecuentemente, con frecuencia se pueden obtener unos resultados comparables con unos oligonucleótidos más cortos cuando se emplean los oligonucleótidos quiméricos, en comparación con fosforotioato deoxioligonucleótidos que se hibrida a la misma región objetivo. Los compuestos quiméricos antisentido de la invención se pueden formar como estructuras compuestas de dos o más oligonucleótidos, oligonucleótidos modificados, oligonucleósidos y/o oligonucleótidos miméticos como se describe arriba. Los oligonucleótidos antisentido quiméricos preferidos incorporan por lo menos un azúcar modificado 2' (preferiblemente 2'-O-$(CH_2)_2$-O-CH_3) en el terminal 3' para conferir una resistencia a nucleasa y una región por lo menos con cuatro azúcares 2'-H para conferir la actividad RNase H. Dichos compuestos también han sido referenciados en el arte como híbridos. Los híbridos preferidos tienen una región 2'de azúcares modificados (preferiblemente 2'-O-$(CH_2)_2$-O-CH_3) en el terminal 3'- y en el terminal 5' separados por lo menos por una región teniendo por lo menos 4 azúcares contiguos 2'-H y preferiblemente incorporan ligamientos de columna vertebral de fosforotioato. Unas patentes representativas de Estados Unidos que enseñan la preparación de dichas estructuras híbridas incluyen, pero no se limitan a, Patentes de Estados Unidos Nos. 5,013,830; 5,149,797; 5,220,007; 5,256,775; 5,366,878; 5,403,711; 5,491,133;

5,565,350; 5,623,065; 5,652,355; 5,652,356 y 5,700,922, cada una de las cuales se incorpora a la presente por referencia en su totalidad.

Los compuestos antisentido empleados de conformidad con esta invención se pueden preparar en forma conveniente y rutinaria mediante una técnica de síntesis de fase sólida muy conocida. Diversos proveedores venden los equipos para estas síntesis como por ejemplo, Applied Biosystems (Foster City, Calif.). Cualquier otro medio empleado para la citada síntesis conocido en el arte adicionalmente o alternativamente se puede utilizar. Es bien conocido emplear técnicas similares para preparar oligonucleótidos de los fosforotioatos y los derivados alquilados. Los compuestos de la invención también se pueden mezclar, encapsular, conjugar o de otra manera asociar con otras moléculas, estructuras de moléculas o mezclas de compuestos como por ejemplo, liposomas, moléculas dirigidas al receptor, formulaciones orales, rectales, tópicas u otras formulaciones, para ayudar en el aporte, la distribución y/o absorción. Unas patentes representativas de Estados Unidos que enseñan la preparación de dichas formulaciones que asisten en el aporte, la distribución y/o absorción incluyen, pero no se limitan a, las Patentes de Estados Unidos Nos. 5,108,921; 5,354,844; 5,416,016; 5,459,127; 5,521,291; 5,543,158; 5,547,932; 5,583,020; 5,591,721; 4,426,330; 4,534,899; 5,013,556; 5,108,921; 5,213,804; 5,227,170; 5,264,221; 5,356,633; 5,395,619; 5,416,016; 5,417,978; 5,462,854; 5,469,854; 5,512,295; 5,527,528; 5,534,259; 5,543,152; 5,556,948; 5,580,575 y 5,595,756, cada una de las cuales se incorpora a la presente por referencia.

Otros ejemplos de oligonucleótidos sentido o antisentido incluyen aquellos oligonucleótidos que se ligan en forma covalente a las fracciones orgánicas, como aquellos descritos en WO 90/10048 y otras fracciones que incrementan la afinidad del oligonucleótido por una secuencia objetivo de ácido nucleico, como poli-(L-lisina). Aún más allá, agentes intercaladores, como elipticina y agentes alquilantes o complejos de metales se pueden adherir a los oligonucleótidos sentido o antisentido para modificar las

especificidades de unión de oligonucleótido antisentido o sentido para la secuencia objetivo del nucleótido.

Se pueden introducir oligonucleótidos antisentido o sentido en una célula que contiene la secuencia objetivo de ácido nucleico mediante cualquier método de transferencia de genes, incluyendo por ejemplo, la transinfección de ADN mediada por $CaPO_4$, electroporado o empleando vectores de transferencia de genes como el virus Epstein-Barr. En un procedimiento preferido, se inserta un oligonucleótido antisentido o sentido en un vector retroviral apropiado. Una célula conteniendo la secuencia objetivo de ácido nucleico entra en contacto con el vector retroviral recombinante, bien sea *in vivo* o *ex vivo*. Unos vectores retrovirales apropiados incluyen, pero no se limitan a, aquellos derivados del retrovirus murino M-MuLV, N2 (un retrovirus derivado de M-MuLV) o los vectores de copia doble designados DCT5A, DCT5B y DCT5C (consultar WO 90/13641).

Los oligonucleótidos sentido o antisentido también se pueden introducir en una célula conteniendo la secuencia objetivo del nucleótido mediante la formación de un conjugado con una molécula de fijación de ligando, como se describe en WO 91/04753. Unas moléculas de fijación de ligando apropiadas incluyen, pero no se limitan a, receptores de superficie celular, factores de crecimiento, otras citosinas u otros ligandos que se fijan a los receptores de superficie celular. Preferiblemente, la conjugación de la molécula de fijación de ligando sustancialmente no interfiere con la habilidad de la molécula de fijación de ligando de fijarse a su molécula correspondiente o receptor o bloquear la entrada del oligonucleótido sentido o antisentido o su versión conjugada en la célula.

Alternativamente, un oligonucleótido sentido o antisentido se puede introducir en una célula conteniendo la secuencia objetivo de ácido nucleico mediante la formación de un complejo oligonucleótido-lípido, como se describe en WO 90/10448. El complejo oligonucleótido sentido o antisentido-lípido preferiblemente se disocia en el interior de la célula mediante una lipasa endógena.

Las moléculas sentido o antisentido de ARN o ADN generalmente tienen por lo menos alrededor de 5 nucleótidos de longitud, alternativamente por lo menos alrededor de 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990 o 1000 nucleótidos de longitud, donde en este contexto, el término "alrededor de" significa que la longitud de una secuencia nucleótida de referencia es más o menos el 10% de la citada longitud de referencia.

Alternativamente, se puede generar un ARN de doble hebra. El ARN de doble hebra que tiene menos de 30 nucleótidos de longitud va a inhibir la expresión de genes específicos cuando se introducen en la célula. Este mecanismo se conoce como interferencia mediada por ARN (RNAi) y ARN pequeño (menos de 30 nucleótidos) empleado como un reactivo se conoce como ARNsi. Los ARN interfiriendo el TAT se pueden identificar y sintetizar utilizando métodos conocidos (Shi *YTrends in Genetics* 19(1):9-12 (2003), WO/2003056012 y WO2003064621). Los ARNsi son útiles para reducir la cantidad de expresión de genes en condiciones donde una reducción de la expresión del gen objetivo aliviaría la enfermedad o el desorden.

Las sondas también se pueden emplear en las técnicas PCR para generar un conjunto de secuencias para la identificación de unas secuencias de codificación TAT estrechamente relacionadas.

Las secuencias nucleótido codificando un TAT también se pueden emplear para construir sondas de hibridación para mapear el gen que codifica ese TAT y para el análisis genético de los individuos con desórdenes genéticos. Las secuencias nucleótido dispuestas en la presente

se pueden mapear a un cromosoma y a regiones específicas de un cromosoma empleando técnicas conocidas como la hibridación *in situ*, el análisis de ligamiento contra marcadores cromosomales conocidos y el cribado de bibliotecas.

Cuando las secuencias de codificación de TAT codifican una proteína que se fija a otra proteína (por ejemplo, cuando el TAT es un receptor), el TAT se puede utilizar en un análisis para identificar las otras proteínas o moléculas comprometidas en la interacción de fijación. Mediante estos métodos, se pueden identificar los inhibidores de la interacción de unión receptor/ligando. Las proteínas comprometidas en dichas interacciones de fijación también se pueden utilizar para cribado en busca de péptido o inhibidores de molécula pequeña o agonistas de la interacción de fijación. Igualmente, se puede emplear el receptor TAT para aislar el(los) ligando(s) correlativo(s). Se pueden diseñar los análisis de cribado para encontrar los compuestos líderes que imitan la actividad biológica de un TAT natural o de un receptor de TAT. Los citados análisis de cribado van incluir unos análisis receptivos a un cribado de alto rendimiento de bibliotecas químicas, haciéndolos particularmente apropiados para identificar candidatos a fármacos de molécula pequeña. Las moléculas pequeñas contempladas incluyen compuestos sintéticos, orgánicos o inorgánicos. Los análisis se pueden realizar bajo diversos formatos, incluyendo análisis de unión proteína-proteína, análisis de cribado bioquímico, inmunoanálisis y análisis a base de células, que están bien caracterizados en el arte.

Los ácidos nucleicos que codifican el TAT o sus formas modificadas también se pueden utilizar para generar animales transgénicos o animales "con inactivación", los cuales a su vez son útiles para el desarrollo y cribado de reactivos terapéuticamente útiles. Un animal transgénico (por ejemplo, un ratón o una rata) es un animal que tiene células que contienen un transgén, transgén este que fue introducido al animal o a un ancestro del animal en una etapa prenatal, por ejemplo, en un estado embriónico. Un transgén es un ADN que se integra al genoma de una célula a partir de la cual se desarrolla un animal transgénico. En una incorporación, ADNc

codificando TAT se puede utilizar para clonar el ADN genómico codificando el TAT de conformidad con técnicas establecidas y las secuencias genómicas empleadas para generar animales transgénicos que contienen células que expresan el ADN codificando el TAT. Los métodos para generar animales transgénicos, particularmente animales como ratones o ratas, se han vuelto convencionales en el arte y se describen por ejemplo, en las Patentes de Estados Unidos Nos. 4,736,866 y 4,870,009. Típicamente, unas células específicas serían consideradas como blanco para la incorporación del transgén TAT con intensificadores específicos para el tejido. Se pueden emplear los animales transgénicos que incluyen una copia de un transgén codificando TAT introducido en la estirpe bacteriana del animal en una etapa embriónica para examinar el efecto de una expresión aumentada del ADN codificando el TAT. Dichos animales se pueden emplear como animales de prueba para reactivos que se considera brindan protección contra, por ejemplo, condiciones patológicas asociadas a su sobreexpresión. De conformidad con esta faceta de la invención, un animal se trata con el reactivo y una incidencia reducida de la condición patológica, comparada con los animales sin tratar que tienen el transgén indicaría una intervención terapéutica potencial de la condición patológica.

Alternativamente, se pueden emplear homólogos TAT no humanos para construir un animal "con inactivación" que tiene un gen defectuoso o alterado codificando TAT como resultado de una recombinación homóloga entre el gen endógeno codificando TAT y el ADN genómico alterado codificando TAT introducido en una célula madre embriónica del animal. Por ejemplo, se pueden utilizar ADNc codificando TAT para clonar el ADN genómico codificando TAT de conformidad con técnicas establecidas. Una porción del ADN genómico codificando TAT se puede eliminar o reemplazar con otro gen, como un gen codificando un marcador seleccionable que se puede emplear para monitorear la integración. Típicamente, se incluyen varios kilobases de ADN de flanco sin alterar (tanto en las caperuzas 5' como 3') en el vector [consultar por ejemplo, Thomas and Capecchi, Cell, 51:503 (1987) para conocer una descripción de

unos vectores de recombinación homóloga]. El vector se introduce en una estirpe de célula madre embriónica (por ejemplo, mediante electroporado) y se escogen las células donde el ADN que se introdujo se recombinó de forma homóloga con el ADN endógeno. [Consultar por ejemplo, Li y colaboradores, Cell, 69:915 (1992)]. Luego las células seleccionadas se inyectan en un blastocito de un animal (por ejemplo, un ratón o una rata) para formar quimeras de agregado [consultar por ejemplo, Bradley, en *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. Luego se puede implantar un embrión quimérico en un animal femenino adoptivo seudoembarazado y el embrión se lleva a término para crear un animal "con inactivación". Se puede identificar la progenie que alberga el ADN recombinado en forma homóloga en sus células bacterianas mediante unas técnicas estándar y se pueden emplear para reproducir animales donde todas las células del animal contienen el ADN recombinado en forma homóloga. Los animales con inactivación se pueden caracterizar, por ejemplo, por su capacidad de defenderse contra determinadas condiciones patológicas debido a la ausencia del polipéptido TAT.

El ácido nucleico codificando los polipéptidos TAT también se puede utilizar en terapia genética. En las aplicaciones de terapia genética, los genes se introducen a las células para lograr la síntesis *in vivo* de un producto genético terapéuticamente efectivo, por ejemplo para reemplazar un gen defectivo. "Terapia genética" incluye tanto la terapia genética convencional donde se alcanza un efecto duradero mediante un solo tratamiento y la administración de agentes terapéuticos genéticos, que comprende la administración por una sola vez o en varias veces repetidas de un ADN o un ARN terapéuticamente efectivo. Se pueden emplear ARNs y ADNs antisentido como agentes terapéuticos para bloquear la expresión de ciertos genes *in vivo*. Ya se ha mostrado que se pueden importar oligonucleótidos antisentido cortos a las células donde actúan como inhibidores, a pesar de sus bajas concentraciones intracelulares causadas por su absorción restringida por parte de la membrana celular. (Zamecnik *y*

colaboradores, Proc. Natl. Acad. Sci. USA 83:4143-4146 [1986]). Se pueden modificar los oligonucleótidos para incrementar la absorción, por ejemplo sustituyendo sus grupos fosfodiéster de carga negativa con grupos sin carga.

Existe una variedad de técnicas disponible para introducir ácidos nucleicos en células viables. Las técnicas varían dependiendo de si el ácido nucleico se transfiere a células cultivadas *in vitro* o *in vivo* en las células del hospedero esperado. Las técnicas apropiadas para la transferencia de ácido nucleico a células mamíferas *in vitro* incluyen el uso de liposomas, electroporado, microinyección, fusión celular, DEAE-dextrano, el método de precipitación calcio fosfato etc. Las técnicas de transferencia de genes *in vivo* preferidas actualmente incluyen la transinfección con vectores virales (típicamente retrovirales) y la transinfección viral mediada por cubierta proteínica-liposoma (Dzau y colaboradores, Trends in Biotechnology 11, 205-210 [1993]). En algunas situaciones es deseable disponer la fuente del ácido nucleico con un agente que apunta a las células objetivo, como un anticuerpo específico para una proteína de superficie de membrana celular o la célula objetivo, un ligando para un receptor en la célula objetivo, etc. Cuando se emplean liposomas, se pueden utilizar las proteínas que se fijan a la proteína de superficie de membrana celular asociados con endocitosis para apuntar a y/o facilitar la absorción de, por ejemplo proteínas cápsidas o fragmentos de las mismas típicas para un tipo de células específicas, anticuerpos para proteínas que sufren una internalización en el ciclo, proteínas que apuntan a la localización intracelular e incrementan la vida media intracelular. La técnica de endocitosis mediada por receptor es descrita por ejemplo, por Wu y colaboradores, J. Biol. Chem. 262, 4429-4432 (1987) y Wagner y colaboradores, Proc. Natl. Acad. Sci. USA 87, 3410-3414 (1990). Para revisar la marcación de genes y los protocolos de terapia genética consultar Anderson y colaboradores, Science 256, 808-813 (1992).

Las moléculas de ácido nucleico codificando los polipéptidos TAT o fragmentos de los mismos descritos en la presente son útiles para la

identificación cromosomal. En este aspecto, existe una necesidad continuada para identificar nuevos marcadores de cromosomas, puesto que relativamente pocos reactivos marcadores de cromosoma, basados en la información actual de las secuencias están disponibles en la actualidad. Cada molécula de ácido nucleico TAT de la presente invención se puede utilizar como un marcador de cromosomas.

Las moléculas de los polipéptidos TAT y del ácido nucleico de la presente invención también se pueden emplear de forma diagnóstica para la tipificación de tejidos, donde los polipéptidos TAT de la presente invención se pueden expresar de manera diferencial en un tejido cuando se compara con otro, preferiblemente en un tejido enfermo cuando se compara con un tejido normal del mismo tipo de tejido. Las moléculas del ácido nucleico TAT serán útiles para generar sondas para PCR, análisis Northern, análisis Southern y análisis Western.

Esta invención abarca los métodos de cribar los compuestos para identificar aquellos que imitan el polipéptido TAT (agonistas) o evitan el efecto del polipéptido TAT (antagonistas). Los análisis de cribado para los candidatos a fármacos antagonistas están designados para identificar los compuestos que fijan o forman un compuesto con los polipéptidos TAT codificados por los genes identificados en la presente, o de otra manera interfieren con la interacción de los polipéptidos codificados con otras proteínas celulares, incluyendo por ejemplo, inhibir la expresión del polipéptido TAT de las células. Estos análisis de cribado van a incluir análisis receptivos a un cribado de alto rendimiento de bibliotecas químicas, haciéndolos particularmente apropiados para identificar candidatos a fármacos de molécula pequeña.

Los análisis se pueden realizar bajo diversos formatos, incluyendo análisis de unión proteína-proteína, análisis de cribado bioquímico, inmunoanálisis y análisis a base de células, que están bien caracterizados en el arte.

Todos los análisis para antagonistas tienen en común que requieren poner en contacto el candidato a fármaco con un polipéptido TAT

codificado por un ácido nucleico identificado en la presente bajo unas condiciones y por un tiempo suficiente para permitir que estos dos componentes interactúen.

En los análisis de unión, la interacción es de unión y el complejo formado se puede aislar o detectar en la mezcla de la reacción. En una incorporación particular, el polipéptido TAT codificado por el gen identificado en la presente, o el candidato a fármaco se inmoviliza en una fase sólida, por ejemplo, en una placa de microtítulo, mediante enlaces covalentes o no covalentes. Generalmente se logra un enlace no covalente al recubrir la superficie sólida con una solución del polipéptido TAT y luego secando. Alternativamente, se puede emplear un anticuerpo inmovilizado, por ejemplo, un anticuerpo monoclonal, específico para el polipéptido TAT que se va a inmovilizar para anclarlo a una superficie sólida. El análisis se completa agregando el componente no inmovilizado, que puede estar marcado con un marcador detectable, al componente inmovilizado, por ejemplo, la superficie recubierta que contiene el componente anclado. Cuando se completa la reacción, se extraen los componentes sin reaccionar, por ejemplo mediante lavado y se detectan los complejos anclados en la superficie sólida. Cuando el componente original no inmovilizado tiene un marcador detectable, la detección del marcador inmovilizado en la superficie indica que se llevó a cabo la aglomeración. Cuando el componente original no inmovilizado no tiene el marcador, se puede detectar la aglomeración, por ejemplo, empleando un anticuerpo marcado que específicamente fija el complejo inmovilizado.

Si el compuesto candidato interactúa con pero no se fija a un polipéptido TAT particular codificado por un gen identificado en la presente, su interacción con ese polipéptido se puede analizar mediante métodos bien conocidos para detectar las interacciones proteína-proteína. Los cromoanálisis incluyen los enfoques tradicionales como, por ejemplo, el enlazamiento cruzado, la co-inmunoprecipitación y la co-purificación mediante gradientes o columnas cromatográficas. Adicionalmente, las interacciones proteína-proteína se pueden monitorear utilizando un sistema

genético a base de levadura descrito por Fields y colegas (Fields and Song, Nature (London), 340:245-246 (1989); Chien y colaboradores, Proc. Natl. Acad. Sci. USA. 88:9578-9582 (1991)) como lo divulga Chevray and Nathans, Proc. Natl. Acad. Sci. USA, 89:5789-5793 (1991). Muchos activadores transcripcionales, como la levadura GAL4, consiste de dos dominios modulares discretos físicamente distintos, uno actúa como el dominio de fijación de ADN y el otro funciona como el dominio de transcripción-activación. El sistema de expresión de levadura descrito en las anteriores publicaciones (generalmente se hace referencia al mismo como "sistema de dos híbridos") aprovecha esta oportunidad y utiliza dos proteínas híbridas, una donde la proteína objetivo se fusiona al dominio de fijación de ADN de GAL4 y la otra, en donde las proteínas de activación candidato se fusionan al dominio de activación. La expresión de un gen indicador GAL1-*lacZ* bajo el control de un promotor activado por GAL4-depende de la reconstitución de la actividad GAL4 mediante la interacción proteína-proteína. Las colonias que contienen polipéptidos que interactúan se detectan con un sustrato cromogénico para β-galactosidasa. Un estuche completo (MATCHMAKER™) para identificar las interacciones proteína-proteína entre dos proteínas específicas empleando la técnica de dos híbridos están disponible en el comercio en Clontech. Este sistema también se puede extender para mapear los dominios de proteínas comprometidas en interacciones específicas de proteínas al igual que para identificar con precisión los residuos de aminoácidos que son cruciales para estas interacciones.

Los compuestos que interfieren con la interacción de un gen codificando un polipéptido TAT identificado en la presente y otros componentes intra- o extracelulares se pueden ensayar como sigue: usualmente se prepara una mezcla de reacción que contiene el producto del gen y el componente intra- o extracelular bajo condiciones y durante un periodo de tiempo que permite la interacción y la unión de los dos productos. Para efectuar ensayos sobre la habilidad de un compuesto candidato para inhibir la unión, la reacción se realiza en ausencia de y en

presencia del compuesto de prueba. Adicionalmente se puede adicionar un placebo a una tercera mezcla de reacción, para que sirva como un control positivo. La unión (formación del complejo) entre el compuesto de prueba y el componente intra- o extracelular presente en la mezcla se monitorea como se describe en la presente arriba. La formación de un complejo en la(s) reacción(es) de control, pero no en la mezcla de reacción que contiene el compuesto de prueba indica que el compuesto de prueba interfiere con la interacción del compuesto de prueba y su pareja de reacción.

Para analizar los antagonistas, se puede agregar el polipéptido TAT a una célula junto con el compuesto que se va a cribar para ver una actividad en particular y la capacidad del compuesto para inhibir la actividad de interés en presencia del polipéptido TAT indica que el compuesto es un antagonista del polipéptido TAT. Alternativamente, se pueden detectar los antagonistas al combinar el polipéptido TAT y un antagonista potencial con unos receptores de polipéptido TAT limitados por la membrana o con unos receptores recombinantes bajo condiciones apropiadas para un análisis de inhibición competitiva. El polipéptido TAT se puede marcar, por ejemplo mediante radioactividad, de manera que el número de moléculas del polipéptido TAT unidas al receptor se puede utilizar para determinar la efectividad del antagonista potencial. El gen codificando el receptor se puede identificar mediante numerosos métodos conocidos por aquellos con destreza en el arte, por ejemplo, el paneo del ligando y la clasificación FACS. Coligan y colaboradores, Current Protocols in Immun., 1(2): Capítulo 5 (1991). Preferiblemente, se emplea la clonación de expresión donde el ARN poliadenilado se prepara partir de una célula que responde al polipéptido TAT y a una biblioteca ADNc creada a partir de este ARN se divide en grupos y se emplea para transinfectar células COS u otras células que no responden al polipéptido TAT. Las células transinfectadas que se cultivan en láminas de vidrio se exponen a un polipéptido TAT marcado. El polipéptido TAT se puede marcar de diversas maneras incluyendo la iodinación o la inclusión de un sitio de reconocimiento de una proteína cinasa de sitio específico. Después de la fijación y la incubación, las

láminas se someten a un análisis autoradiográfico. Los grupos positivos se identifican y se preparan unos subgrupos y se transinfectan nuevamente empleando un proceso interactivo de subgrupos y recribado eventualmente produciendo un solo clon que codifica el receptor putativo.

Como un enfoque alternativo para la identificación del receptor, el polipéptido TAT marcado se puede enlazar mediante fotoafinidad a la membrana celular o se pueden extraer preparaciones que expresan la molécula receptora. El material enlazado en cruz se resuelve mediante PAGE y se expone a una película de rayos X. El complejo marcado que contiene el receptor se puede extirpar y se puede transformar en fragmentos peptídicos y se puede someter a microsecuenciación de proteínas. La secuencia de aminoácidos obtenida de la microsecuenciación se podría emplear para diseñar un conjunto de sondas degeneradas de oligonucleótidos para cribar una biblioteca ADNc para identificar el gen codificando el receptor putativo.

En otro análisis de antagonistas, se podrían incubar células de mamíferos o una preparación de membrana expresando el receptor con polipéptido TAT marcado en presencia del compuesto candidato. Entonces se podría mediar la capacidad del compuesto para intensificar o bloquear esta interacción.

Ejemplos más específicos de antagonistas potenciales incluyen un oligonucleótido que se une a las fusiones de la inmunoglobulina con un polipéptido TAT, y en particular, anticuerpos incluyendo sin limitación, anticuerpos policlonales y monoclonales y fragmentos de anticuerpos, anticuerpos de una sola cadena, anticuerpos anti-idiotípico y versiones quiméricas o humanizadas de dichos anticuerpos o fragmentos, al igual que anticuerpos humanos y fragmentos de anticuerpos. Alternativamente, un antagonista potencial puede ser una proteína muy cercana como por ejemplo, una forma mutada del polipéptido TAT que reconoce el receptor pero no imparte ningún efecto, por lo tanto inhibe la acción del polipéptido TAT en forma competitiva.

Otro antagonista potencial del polipéptido TAT es una construcción de un ARN o de un ADN antisentido preparado mediante tecnología antisentido, donde por ejemplo, una molécula antisentido ARN o ADN actúa para bloquear directamente la traslación de ARNm mediante hibridación al ARNm objetivo y evitando la traslación de la proteína. Se puede utilizar la tecnología antisentido para controlar la expresión del gen a través de la formación de un ADN o un ARN de triple hélice o de antisentido donde ambos métodos se basan en la unión de un polinucleótido a ADN o RNA. Por ejemplo, la porción de codificación 5' de la secuencia de polinucleótido, que codifica los polipéptidos TAT maduros en la presente se utiliza para diseñar un oligonucleótido antisentido de ARN con aproximadamente desde alrededor de 10 a 40 pares de bases de longitud. Se diseña un oligonucleótido ADN para que sea complementario a una región del gen involucrado en la transcripción (triple hélice - consultar Lee y colaboradores, Nucl. Acids Res., 6:3073 (1979); Cooney y colaboradores, Science, 241: 456 (1988); Dervan y colaboradores, Science, 251:1360 (1991)), evitando de esta forma la transcripción y la producción del polipéptido TAT. El oligonucleótido antisentido de ARN se hibrida al ARNm *in vivo* y bloquea la traslación de la molécula de ARNm en el polipéptido TAT (antisense - Okano, Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gen Expression (CRC Press: Boca Raton, FL, 1988). Los oligonucleótidos descritos arriba también se pueden administrar a las células de forma que el ARN o el ADN antisentido se puede expresar *in vivo* para inhibir la producción del polipéptido TAT. Cuando se emplea el ADN antisentido, se prefieren los oligodeoxiribonucleótidos derivados del sitio de traslación-iniciación, por ejemplo, entre alrededor de -10 y +10 posiciones de la secuencia nucleótida del gen objetivo.

Antagonistas potenciales incluyen moléculas pequeñas que se fijan al sitio activo, al sitio de unión del receptor o el factor de crecimiento u otros sitios de unión relevantes del polipéptido TAT, bloqueando de esta forma la actividad biológica normal del polipéptido TAT. Ejemplos de moléculas

pequeñas incluyen, pero no se limitan a, péptidos pequeños o moléculas de tipo peptídico, preferiblemente péptidos solubles y compuestos sintéticos no orgánicos o inorgánicos no peptidilo.

Las ribozimas son moléculas enzimáticas de ARN capaces de catalizar la segmentación específica del ARN. Las ribozimas actúan mediante hibridación específica de secuencias al ARN objetivo complementario, seguido por una segmentación endonucleótica. Los sitios específicos de segmentación de las ribozimas en un ARN objetivo potencial se pueden identificar mediante técnicas conocidas. Para mayores detalles consultar por ejemplo, Rossi, Current Biology, 4:469-471 (1994) y PCT publicación No. WO 97/33551 (publicada el 18 de septiembre de 1997).

Las moléculas de ácido nucleico en formación de hélice triple empleadas para inhibir la transcripción deben ser de una sola hebra y estar compuestas por deoxinucleótidos. La composición base de estos oligonucleótidos se diseña de forma que promueve la formación de la hélice triple mediante las reglas de pares de bases de Hoogsteen que generalmente requieren unos tramos largos de purinas o pirimidinas en una hebra de un duplo. Para mayores detalles, consultar por ejemplo, PCT publicación No. WO 97/33551, *supra*.

Estas moléculas pequeñas se pueden identificar mediante uno o varios de los análisis de cribado descritos arriba y/o por medio de cualquier otra técnica de cribado bien conocida por aquellos con destreza en el arte.

El polipéptido TAT aislado codificando ácido nucleico se puede emplear en la presente para producir polipéptido TAT de forma recombinante utilizando unas técnicas bien conocidas en el arte y como se describe en la presente. A su vez, los polipéptidos TAT producidos se pueden emplear para generar anticuerpos anti-TAT empleando técnicas bien conocidas en el arte y como se describe en la presente.

Los anticuerpos que se fijan específicamente a un polipéptido TAT identificado en la presente, al igual que otras moléculas identificadas por los análisis de cribado divulgados en la presente arriba, se pueden administrar

para el tratamiento de varios desórdenes, incluyendo cáncer, en forma de composiciones farmacéuticas.

Si el polipéptido TAT es intracelular y se emplean anticuerpos enteros como inhibidores, se prefieren los anticuerpos internalizantes. Sin embargo, también se pueden utilizar lipofecciones o liposomas para administrar el anticuerpo o un fragmento de anticuerpo, en las células. Cuando se emplean fragmentos de anticuerpos, se prefiere el fragmento inhibidor más pequeño que específicamente se une al dominio de unión de la proteína objetivo. Por ejemplo, con base en las secuencias de región variable de un anticuerpo, se pueden designar unas moléculas peptídicas que retienen la capacidad de fijar la secuencia de la proteína objetivo. Dichos péptidos se pueden sintetizar químicamente y/o se pueden producir mediante tecnología de ADN recombinante. Consultar por ejemplo, Marasco *y colaboradores*, Proc. Natl. Acad. Sci. USA. 90: 7889-7893 (1993).

La formulación de la presente también puede contener más de un compuesto activo necesario para la indicación en particular que se está tratando, preferiblemente aquellos con actividades complementarias que no tienen un efecto adverso entre sí. Alternativamente o adicionalmente, la composición puede comprender un agente que intensifica su función como, por ejemplo, un agente citotóxico, una citosina, un agente quimioterapéutico o un agente inhibidor de crecimiento. Dichas moléculas se encuentran presentes de forma apropiada en combinación y en cantidades que son efectivas para el propósito buscado.

Los siguientes ejemplos se ofrecen únicamente con fines ilustrativos y su propósito no es limitar el ámbito de la presente invención de ninguna manera.

Todas las referencias de patentes y de literatura citadas en la presente especificación se incorporan a la presente por referencia en su totalidad.

EJEMPLOS

En los ejemplos se emplearon reactivos disponibles en el comercio, de acuerdo con las instrucciones del fabricante, salvo si se indica en contrario. La fuente de las células identificadas en los siguientes ejemplos y en toda la especificación, mediante los números de acceso de ATCC es la American Type Culture Collection, Manassas, VA.

EJEMPLO 1: Perfiles de Expresión de Tejidos Empleando GeneExpress®

Se analizó una base de datos propietaria conteniendo información sobre la expresión de genes (GeneExpress®, Gene Logic Inc., Gaithersburg, MD) en una tentativa por identificar polipéptidos (y los ácidos nucleicos que los codifican) cuya expresión se regula por incremento en uno(s) tejido(s) de tumor particulares de interés cuando se comparan con otro(s) tumor(es) y/o tejidos normales. Específicamente, se efectuó el análisis de la base de datos de GeneExpress® utilizando el software disponible en Gene Logic Inc., Gaithersburg, MD, para ser utilizado con la base de datos de GeneExpress® o con software propietario escrito y desarrollado en Genentech, Inc. para ser empleado con la base de datos de GeneExpress®. La clasificación de coincidencias positivas en el análisis se basa en diversos criterios incluyendo por ejemplo, la especificidad del tejido, la especificidad del tumor y el nivel de expresión en tejidos normales indispensables y/o en tejidos normales de proliferación. A continuación se encuentra una lista de moléculas cuyo perfil de expresión del tejido según se determina por un análisis de la base de datos de GeneExpress® evidencia una alta expresión de tejido y una significativa regulación por incremento de la expresión en uno o varios tumores específicos cuando se compara(n) con otro(s) tumor(es) y/o con tejidos normales y opcionalmente una expresión relativamente baja en tejidos normales indispensables y/o en tejidos normales de proliferación. Como tal, las moléculas enumeradas a

continuación son excelentes objetivos para los polipéptidos para el diagnóstico y la terapia para el cáncer en mamíferos.

Molécula	regulación por incremento en la expresión en:	cuando se compara con:
DNA77507 (TAT161)	tumor de mama	tejido normal de mama
DNA77507 (TAT161)	tumor de colon	tejido normal de colon
DNA77507 (TAT161)	tumor de pulmón	tejido normal de pulmón
DNA77507 (TAT161)	tumor de riñón	tejido normal de riñón
DNA77507 (TAT161)	tumor de hígado	tejido normal de hígado
DNA77507 (TAT161)	tumor ovárico	tejido ovárico normal
DNA77507 (TA161)	tumor pancreático	tejido pancreático normal
DNA77507 (TAT161)	tumor de recto	tejido normal de recto
DNA77507 (TAT161)	tumor de piel	tejido normal de piel
DNA77507 (TAT161)	tumor uterino	tejido uterino normal
DNA77507 (TAT161)	tumor cerebral	tejido cerebral normal
DNA77507 (TAT161)	tumor de tejidos blandos	tejido blando normal
DNA77507 (TAT161)	tumor de hueso	tejido normal de hueso
DNA80894 (TAT101)	tumor de mama	tejido normal de mama
DNA82343 (TAT157)	tumor de colon	tejido normal de colon
DNA82343 (TAT157)	tumor ovárico	tejido ovárico normal
DNA82343 (TAT157)	tumor de estómago	tejido normal de estómago
DNA82343 (TAT157)	tumor de hígado	tejido normal de hígado
DNA82343 (TAT157)	tumor de recto	tejido normal de recto
DNA82343 (TAT157)	tumor de intestino delgado	tejido normal de intestino delgado
DNA82343 (TAT157)	tumor de esófago	tejido normal de esófago
DNA82343 (TAT157)	tumor de testículo	tejido normal de testículo
DNA82343 (TAT157)	tumor de timo	tejido normal de timo
DNA87994 (TAT160)	tumor de mama	tejido normal de mama
DNA87994 (TAT160)	tumor pancreático	tejido pancreático normal
DNA87994 (TAT160)	tumor de recto	tejido normal de recto
DNA87994 (TAT160)	tumor de colon	tejido normal de colon
DNA87994 (TAT160)	tumor de esófago	tejido normal de esófago
DNA87994 (TAT160)	tumor ovárico	tejido ovárico normal

DNA87994 (TAT160)	tumor de pulmón	tejido normal de pulmón
DNA87994 (TAT160)	tumor uterino	tejido uterino normal
DNA88131 (TAT158)	tumor de hueso	tejido normal de hueso
DNA88131 (TAT158)	tumor de mama	tejido normal de mama
DNA88131 (TAT158)	tumor de colon	tejido normal de colon
DNA88131 (TAT158)	tumor uterino	tejido uterino normal
DNA88131 (TAT158)	tumor de esófago	tejido normal de esófago
DNA88131 (TAT158)	tumor de pulmón	tejido normal de pulmón
DNA88131 (TAT158)	tumor ovárico	tejido ovárico normal
DNA88131 (TAT158)	tumor pancreático	tejido pancreático normal
DNA88131 (TAT158)	tumor de próstata	tejido normal de próstata
DNA88131 (TAT158)	tumor de piel	tejido normal de piel
DNA88131 (TAT158)	tumor de tejidos blandos	tejido blando normal
DNA88131 (TAT158)	tumor de estómago	tejido normal de estómago

Molécula	**regulación por incremento en la expresión en:**	**cuando se compara con:**
DNA88131 (TAT158)	tumor de recto	tejido normal de recto
DNA88131 (TAT158)	tumor neuroendocrino	tejido neuroendocrino normal
DNA88131 (TAT158)	tumor cerebral	tejido cerebral normal
DNA95930 (TAT110)	tumor de colon	tejido normal de colon
DNA95930 (TAT110)	tumor uterino	tejido uterino normal
DNA95930 (TAT110)	tumor de endometrio	tejido normal de endometrio
DNA95930 (TAT110)	tumor de recto	tejido normal de recto
DNA95930 (TAT110)	tumor ovárico	tejido ovárico normal
DNA95930 (TAT110)	tumor de mama	tejido normal de mama
DNA95930 (TAT110)	tumor de pulmón	tejido normal de pulmón
DNA95930 (TAT110)	tumor de próstata	tejido normal de próstata
DNA95930-1 (TAT210)	tumor de colon	tejido normal de colon
DNA95930-1 (TAT210)	tumor uterino	tejido uterino normal
DNA95930-1 (TAT210)	tumor de endometrio	tejido normal de endometrio
DNA95930-1 (TAT210)	tumor de recto	tejido normal de recto
DNA95930-1 (TAT210)	tumor ovárico	tejido ovárico normal
DNA95930-1 (TAT210)	tumor de mama	tejido normal de mama
DNA95930-1 (TAT210)	tumor de pulmón	tejido normal de pulmón

DNA95930-1 (TAT210)	tumor de próstata	tejido normal de próstata
DNA96917 (TAT159)	tumor pancreático	tejido pancreático normal
DNA96917 (TAT159)	tumor de pulmón	tejido normal de pulmón
DNA96917 (TAT159)	tumor de hígado	tejido normal de hígado
DNA96930 (TAT112)	tumor de mama	tejido normal de mama
DNA96930 (TAT112)	tumor de colon	tejido normal de colon
DNA96930 (TAT112)	tumor de recto	tejido normal de recto
DNA96930 (TAT112)	tumor uterino	tejido uterino normal
DNA96930 (TAT112)	tumor de pulmón	tejido normal de pulmón
DNA96930 (TAT112)	tumor ovárico	tejido ovárico normal
DNA96930 (TAT112)	tumor pancreático	tejido pancreático normal
DNA96930 (TAT112)	tumor de estómago	tejido normal de estómago
DNA96936 (TAT147)	tumor de mama	tejido normal de mama
DNA96936 (TAT147)	tumor de colon	tejido normal de colon
DNA96936 (TAT147)	tumor de testículo	tejido normal de testículo
DNA96936 (TAT147)	tumor ovárico	tejido ovárico normal
DNA98565 (TAT145)	tumor cerebral	tejido cerebral normal
DNA98565 (TAT145)	glioma	tejido glial normal
DNA246435 (TAT152)	tumor cerebral	tejido cerebral normal
DNA246435 (TAT152)	glioma	tejido glial normal
DNA98591 (TAT162)	tumor de colon	tejido normal de colon
DNA98591 (TAT162)	tumor de recto	tejido normal de recto
DNA98591 (TAT162)	tumor ovárico	tejido ovárico normal
DNA98591 (TAT162)	tumor pancreático	tejido pancreático normal
DNA98591 (TAT162)	tumor de estómago	tejido normal de estómago
DNA108809 (TAT114)	tumor de colon	tejido normal de colon
DNA108809 (TAT114)	tumor de riñón	tejido normal de riñón
DNA119488 (TAT119)	tumor de colon	tejido normal de colon
DNA119488 (TAT119)	tumor de pulmón	tejido normal de pulmón
DNA119488 (TAT119)	tumor de recto	tejido normal de recto
DNA143493 (TAT103)	tumor de mama	tejido normal de mama
DNA167234 (TAT130)	tumor de próstata	tejido normal de próstata
DNA235621 (TAT166)	tumor de próstata	tejido normal de próstata
DNA235621 (TAT166)	tumor de hígado	tejido normal de hígado
DNA176766 (TAT132)	tumor de riñón	tejido normal de riñón

DNA176766 (TAT132)	tumor ovárico	tejido ovárico normal
DNA176766 (TAT132)	tumor uterino	tejido uterino normal

Molécula	**regulación por incremento en la expresión en:**	**cuando se compara con:**
DNA236463 (TAT150)	tumor de riñón	tejido normal de riñón
DNA236463 (TAT150)	tumor ovárico	tejido ovárico normal
DNA236463 (TAT150)	tumor uterino	tejido uterino normal
DNA181162 (TAT129)	tumor de próstata	tejido normal de próstata
DNA188221 (TAT111)	tumor de colon	tejido normal de colon
DNA188221 (TAT111)	tumor de endometrio	tejido normal de endometrio
DNA188221 (TAT111)	tumor de estómago	tejido normal de estómago
DNA233876 (TAT146)	tumor de colon	tejido normal de colon
DNA233876 (TAT146)	tumor de endometrio	tejido normal de endometrio
DNA233876 (TAT146)	tumor de estómago	tejido normal de estómago
DNA193891 (TAT148)	tumor de colon	tejido normal de colon
DNA248170 (TAT187)	tumor de colon	tejido normal de colon
DNA248170 (TAT187)	tumor de mama	tejido normal de mama
DNA194628 (TAT118)	tumor de riñón	tejido normal de riñón
DNA246415 (TAT167)	tumor de riñón	tejido normal de riñón
DNA215609 (TAT113)	tumor de colon	tejido normal de colon
DNA215609 (TAT113)	tumor de recto	tejido normal de recto
DNA220432 (TAT128)	tumor de próstata	tejido normal de próstata
DNA226094 (TAT164)	tumor de mama	tejido normal de mama
DNA226094 (TAT164)	tumor cerebral	tejido cerebral normal
DNA226094 (TAT164)	tumor de pulmón	tejido normal de pulmón
DNA226094 (TAT164)	tumor de piel	tejido normal de piel
DNA226165 (TAT122)	tumor de mama	tejido normal de mama
DNA226165 (TAT122)	tumor de endometrio	tejido normal de endometrio
DNA226165 (TAT122)	tumor de riñón	tejido normal de riñón
DNA226165 (TAT122)	tumor de pulmón	tejido normal de pulmón
DNA226165 (TAT122)	tumor ovárico	tejido ovárico normal
DNA226165 (TAT122)	tumor de colon	tejido normal de colon
DNA226165 (TAT122)	tumor de recto	tejido normal de recto
DNA226165 (TAT122)	tumor de piel	tejido normal de piel

DNA226165 (TAT122)	tumor de tejidos blandos	tejido blando normal
DNA226165 (TAT122)	tumor de vejiga	tejido normal de vejiga
DNA226237 (TAT117)	tumor de riñón	tejido normal de riñón
DNA246450 (TAT168)	tumor de riñón	tejido normal de riñón
DNA226456 (TAT144)	tumor de mama	tejido normal de mama
DNA226456 (TAT144)	tumor de colon	tejido normal de colon
DNA226456 (TAT144)	tumor de recto	tejido normal de recto
DNA226456 (TAT144)	tumor de endometrio	tejido normal de endometrio
DNA226456 (TAT144)	tumor de riñón	tejido normal de riñón
DNA226456 (TAT144)	tumor de pulmón	tejido normal de pulmón
DNA226456 (TAT144)	tumor ovárico	tejido ovárico normal
DNA226456 (TAT144)	tumor de piel	tejido normal de piel
DNA237637 (TAT188)	tumor de mama	tejido normal de mama
DNA237637 (TAT188)	tumor de colon	tejido normal de colon
DNA237637 (TAT188)	tumor de recto	tejido normal de recto
DNA237637 (TAT188)	tumor de endometrio	tejido normal de endometrio
DNA237637 (TAT188)	tumor de riñón	tejido normal de riñón
DNA237637 (TAT188)	tumor de pulmón	tejido normal de pulmón
DNA237637 (TAT188)	tumor ovárico	tejido ovárico normal
DNA237637 (TAT188)	tumor de piel	tejido normal de piel
DNA237637 (TAT188)	tumor de hígado	tejido normal de hígado
DNA237637 (TAT188)	tumor de pulmón	tejido normal de pulmón
DNA226539 (TAT126)	tumor de mama	tejido normal de mama
DNA226539 (TAT126)	tumor de colon	tejido normal de colon
DNA226539 (TAT126)	tumor de recto	tejido normal de recto

Molécula	**regulación por incremento en la expresión en:**	**cuando se compara con:**
DNA226539 (TAT126)	tumor de endometrio	tejido normal de endometrio
DNA226539 (TAT126)	tumor de pulmón	tejido normal de pulmón
DNA226539 (TAT126)	tumor ovárico	tejido ovárico normal
DNA226539 (TAT126)	tumor pancreático	tejido pancreático normal
DNA236511 (TAT151)	tumor de mama	tejido normal de mama
DNA236511 (TAT151)	tumor de colon	tejido normal de colon
DNA236511 TAT151)	tumor de recto	tejido normal de recto

DNA236511 (TAT151)	tumor de endometrio	tejido normal de endometrio
DNA236511 (TAT151)	tumor de pulmón	tejido normal de pulmón
DNA236511 (TAT151)	tumor ovárico	tejido ovárico normal
DNA236511 (TAT151)	tumor pancreático	tejido pancreático normal
DNA226771 (TAT115)	tumor de mama	tejido normal de mama
DNA226771 (TAT115)	tumor de colon	tejido normal de colon
DNA227087 (TAT163)	tumor de mama	tejido normal de mama
DNA227087 (TAT163)	tumor de colon	tejido normal de colon
DNA227087 (TAT163)	tumor de recto	tejido normal de recto
DNA227087 (TAT163)	tumor de pulmón	tejido normal de pulmón
DNA227087 (TAT163)	tumor ovárico	tejido ovárico normal
DNA227087 (TAT163)	tumor de próstata	tejido normal de próstata
DNA227087 (TAT163)	tumor endocrino	tejido endocrino normal
DNA227087 (TAT163)	tumor de riñón	tejido normal de riñón
DNA227087 (TAT163)	tumor de hígado	tejido normal de hígado
DNA227087 (TAT163)	tumor del sistema nervioso	tejido normal del sistema nervioso
DNA227087 (TAT163)	tumor pancreático	tejido pancreático normal
DNA227087 (TAT163)	tumor uterino	tejido uterino normal
DNA227087 (TAT163)	tumor de intestino delgado	tejido normal de intestino delgado
DNA227087 (TAT163)	tumor linfático	tejido linfático normal
DNA266307 (TAT227)	tumor de mama	tejido normal de mama
DNA266307 (TAT227)	tumor de colon	tejido normal de colon
DNA266307 (TAT227)	tumor de recto	tejido normal de recto
DNA266307 (TAT227)	tumor de pulmón	tejido normal de pulmón
DNA266307 (TAT227)	tumor ovárico	tejido ovárico normal
DNA266307 (TAT227)	tumor de próstata	tejido normal de próstata
DNA266307 (TAT227)	tumor endocrino	tejido endocrino normal
DNA266307 (TAT227)	tumor de riñón	tejido normal de riñón
DNA266307 (TAT227)	tumor de hígado	tejido normal de hígado
DNA266307 (TAT227)	tumor del sistema nervioso	tejido normal del sistema nervioso
DNA266307 (TAT227)	tumor pancreático	tejido pancreático normal
DNA266307 (TAT227)	tumor uterino	tejido uterino normal
DNA266307 (TAT227)	tumor de intestino delgado	tejido normal de intestino delgado
DNA266307 (TAT227)	tumor linfático	tejido linfático normal
DNA266311 (TAT228)	tumor de mama	tejido normal de mama

DNA266311 (TAT228)	tumor de colon	tejido normal de colon
DNA266311 (TAT228)	tumor de recto	tejido normal de recto
DNA266311 (TAT228)	tumor de pulmón	tejido normal de pulmón
DNA266311 (TAT228)	tumor ovárico	tejido ovárico normal
DNA266311 (TAT228)	tumor de próstata	tejido normal de próstata
DNA266311 (TAT228)	tumor endocrino	tejido endocrino normal
DNA266311 (TAT228)	tumor de riñón	tejido normal de riñón
DNA266311 (TAT228)	tumor de hígado	tejido normal de hígado
DNA266311 (TAT228)	tumor del sistema nervioso	tejido normal del sistema nervioso
DNA266311 (TAT228)	tumor pancreático	tejido pancreático normal
DNA266311 (TAT228)	tumor uterino	tejido uterino normal
DNA266311(TAT228)	tumor de intestino delgado	tejido normal de intestino delgado
DNA266311 (TAT228)	tumor linfático	tejido linfático normal

Molécula	**regulación por incremento en la expresión en:**	**cuando se compara con:**
DNA266312 (TAT229)	tumor de mama	tejido normal de mama
DNA266312 (TAT229)	tumor de colon	tejido normal de colon
DNA266312 (TAT229)	tumor de recto	tejido normal de recto
DNA266312 (TAT229)	tumor de pulmón	tejido normal de pulmón
DNA266312(TAT229)	tumor ovárico	tejido ovárico normal
DNA266312 (TAT229)	tumor de próstata	tejido normal de próstata
DNA266312 (TAT229)	tumor endocrino	tejido endocrino normal
DNA266312 (TAT229)	tumor de riñón	tejido normal de riñón
DNA266312 (TAT229)	tumor de hígado	tejido normal de hígado
DNA266312 (TAT229)	tumor del sistema nervioso	tejido normal del sistema nervioso
DNA266312 (TAT229)	tumor pancreático	tejido pancreático normal
DNA266312 (TAT229)	tumor uterino	tejido uterino normal
DNA266312 (TAT229)	tumor de intestino delgado	tejido normal de intestino delgado
DNA266312 (TAT229)	tumor linfático	tejido linfático normal
DNA266313 (TAT230)	tumor de mama	tejido normal de mama
DNA266313 (TAT230)	tumor de colon	tejido normal de colon
DNA266313 (TAT230)	tumor de recto	tejido normal de recto
DNA266313 (TAT230)	tumor de pulmón	tejido normal de pulmón
DNA266313 (TAT230)	tumor ovárico	tejido ovárico normal

DNA266313 (TAT230)	tumor de próstata	tejido normal de próstata
DNA266313 (TAT230)	tumor endocrino	tejido endocrino normal
DNA266313 (TAT230)	tumor de riñón	tejido normal de riñón
DNA266313 (TAT230)	tumor de hígado	tejido normal de hígado
DNA266313 (TAT230)	tumor del sistema nervioso	tejido normal del sistema nervioso
DNA266313 (TAT230)	tumor pancreático	tejido pancreático normal
DNA266313 (TAT230)	tumor uterino	tejido uterino normal
DNA266313 (TAT230)	tumor de intestino delgado	tejido normal de intestino delgado
DNA266313 (TAT230)	tumor linfático	tejido linfático normal
DNA227224 (TAT121)	tumor de mama	tejido normal de mama
DNA227224 (TAT121)	tumor de colon	tejido normal de colon
DNA227224 (TAT121)	tumor de recto	tejido normal de recto
DNA227224 (TAT121)	tumor de endometrio	tejido normal de endometrio
DNA227224 (TAT121)	tumor de riñón	tejido normal de riñón
DNA227224 (TAT121)	tumor de pulmón	tejido normal de pulmón
DNA227224 (TAT121)	tumor ovárico	tejido ovárico normal
DNA227224 (TAT121)	tumor de piel	tejido normal de piel
DNA227224 (TAT121)	tumor de testículo	tejido normal de testículo
DNA227224 (TAT121)	tumor de vejiga	tejido normal de vejiga
DNA247486 (TAT183)	tumor de mama	tejido normal de mama
DNA247486 (TAT183)	tumor de colon	tejido normal de colon
DNA247486 (TAT183)	tumor de recto	tejido normal de recto
DNA247486 (TAT183)	tumor de endometrio	tejido normal de endometrio
DNA247486 (TAT183)	tumor de riñón	tejido normal de riñón
DNA247486 (TAT183)	tumor de pulmón	tejido normal de pulmón
DNA247486 (TAT183)	tumor ovárico	tejido ovárico normal
DNA247486 (TAT183)	tumor de piel	tejido normal de piel
DNA247486 (TAT183)	tumor de testículo	tejido normal de testículo
DNA247486 (TAT183)	tumor de vejiga	tejido normal de vejiga
DNA227800(TAT131)	tumor de próstata	tejido normal de próstata
DNA228199 (TAT127)	tumor de mama	tejido normal de mama
DNA228199 (TAT127)	tumor de endometrio	tejido normal de endometrio
DNA228199 (TAT127)	tumor ovárico	tejido ovárico normal
DNA228199 (TAT127)	tumor pancreático	tejido pancreático normal
DNA228199 (TAT127)	tumor de pulmón	tejido normal de pulmón

DNA228201 (TAT116)	tumor de colon	tejido normal de colon

Molécula	regulación por incremento en la expresión en:	cuando se compara con:
DNA228201 (TAT116)	tumor de recto	tejido normal de recto
DNA247488 (TAT189)	tumor de colon	tejido normal de colon
DNA247488 (TAT189)	tumor de recto	tejido normal de recto
DNA236538 (TAT190)	tumor de colon	tejido normal de colon
DNA236538 (TAT190)	tumor de recto	tejido normal de recto
DNA247489 (TAT191)	tumor de colon	tejido normal de colon
DNA247489 (TAT191)	tumor de recto	tejido normal de recto
DNA228211 (TAT133)	tumor uterino	tejido uterino normal
DNA233937 (TAT186)	tumor uterino	tejido uterino normal
DNA233937 (TAT186)	tumor ovárico	tejido ovárico normal
DNA228994 (TAT124)	tumor de pulmón	tejido normal de pulmón
DNA228994 (TAT124)	tumor ovárico	tejido ovárico normal
DNA228994 (TAT124)	tumor de piel	tejido normal de piel
DNA228994 (TAT124)	tumor de mama	tejido normal de mama
DNA229410 (TAT105)	tumor de mama	tejido normal de mama
DNA229411 (TAT107)	tumor de mama	tejido normal de mama
DNA229413 (TAT108)	tumor de mama	tejido normal de mama
DNA229700 (TAT139)	tumor de mama	tejido normal de mama
DNA231312 (TAT143)	tumor de mama	tejido normal de mama
DNA231312 (TAT143)	tumor de colon	tejido normal de colon
DNA231542 (TAT100)	tumor cerebral	tejido cerebral normal
DNA231542 (TAT100)	glioma	tejido glial normal
DNA231542-1 (TAT284)	tumor cerebral	tejido cerebral normal
DNA231542-1 (TAT284)	glioma	tejido glial normal
DNA231542-2 (TAT285)	tumor cerebral	tejido cerebral normal
DNA231542-2 (TAT285)	glioma	tejido glial normal
DNA297393 (TAT285-1)	tumor cerebral	tejido cerebral normal
DNA297393 (TAT285-1)	glioma	tejido glial normal
DNA234833 (TAT149)	tumor de colon	tejido normal de colon
DNA268022 (TAT231)	tumor de colon	tejido normal de colon
DNA268022 (TAT231)	tumor de mama	tejido normal de mama

DNA236534 (TAT102)	tumor de endometrio	tejido normal de endometrio
DNA236534 (TAT102)	tumor de pulmón	tejido normal de pulmón
DNA236534 (TAT102)	tumor ovárico	tejido ovárico normal
DNA236534 (TAT102)	tumor pancreático	tejido pancreático normal
DNA236534 (TAT102)	tumor de próstata	tejido normal de próstata
DNA236534 (TAT102)	tumor de estómago	tejido normal de estómago
DNA236534 (TAT102)	tumor de vejiga	tejido normal de vejiga
DNA246430 (TAT109)	tumor de mama	tejido normal de mama
DNA246430 (TAT109)	tumor de próstata	tejido normal de próstata
DNA247480 (TAT142)	tumor de mama	tejido normal de mama
DNA247480 (TAT142)	tumor de pulmón	tejido normal de pulmón
DNA264454 (TAT106)	tumor de mama	tejido normal de mama

EJEMPLO 2: Análisis con *Microarrays* (pequeñas placas de cristal) para Detectar la Regulación por Incremento de Polipéptidos TAT en Tumores Cancerosos

Unos *microarrays* (pequeñas placas de cristal) de ácido nucleico que con frecuencia contienen miles de secuencias de genes son útiles para identificar genes expresados de forma diferencial en tejidos enfermos cuando se comparan con sus contrapartes normales. Empleando *microarrays* de ácido nucleico, se transcriben y se marcan muestras de prueba y de control de ARNm de forma inversa para generar las sondas de ADNc. Luego se hibridan las sondas de ADNc a un arreglo de ácidos nucleicos inmovilizados en un soporte sólido. El arreglo se configura de manera que se conoce la secuencia y la posición de cada miembro del arreglo. Por ejemplo, se puede arreglar una selección de genes que se sabe que se expresan en determinados estados de la enfermedad en un soporte sólido. La hibridación de una sonda marcada con determinado miembro del arreglo indica que la muestra de la cual se derivó la sonda expresa ese gen. Si la señal de hibridación de una sonda de una muestra de prueba (tejido con la enfermedad) es mayor que la señal de hibridación de una sonda de una muestra de control (tejido normal), se identifica el o los genes sobreexpresado(s) en el tejido de la enfermedad. La implicación de este

resultado es que una proteína sobreexpresada en un tejido con la enfermedad es útil tanto como un marcador de diagnóstico que identifica la presencia de la patología y también como un blanco terapéutico para el tratamiento de la patología.

La metodología de hibridación de ácidos nucleicos y la tecnología de *microarray* son muy conocidas en el arte. En un ejemplo, la preparación específica de ácidos nucleicos para hibridación y sondas, las condiciones de las láminas y la hibridación se describen en detalle en el PCT de la Solicitud de Patente Serial No. PCT/US01/10482, presentada el 30 de marzo de 2001 y la cual se incorpora a la presente por referencia.

En el presente ejemplo, se estudiaron los tumores cancerosos derivados de varios tejidos humanos para identificar la expresión del gen regulado por incremento relativo a tumores cancerosos de diferentes tipos de tejido y/o tejidos humanos no cancerosos en una tentativa por identificar aquellos polipéptidos sobreexpresados en un(os) tumor(es) particular(es) de cáncer. En ciertos experimentos, se obtuvo y se analizó tejido humano de tumor canceroso y tejido humano de tumor no canceroso del mismo tipo de tejido (con frecuencia del mismo paciente) en busca de la expresión del polipéptido TAT. Adicionalmente, se obtuvo tejido humano de tumor canceroso de cualquiera entre una variedad de diferentes tumores humanos y se comparó con una muestra de control epitelial "universal" que se preparó al combinar tejidos humanos no cancerosos de origen epitelial, incluyendo hígado, riñón y pulmón. El ARNm aislado de los tejidos combinados representa una mezcla de productos de genes expresados de diferentes tejidos. Los experimentos de hibridación en *microarray* empleando las muestras de control combinadas generaron un análisis en un plano linear de dos colores. Luego se utilizó la pendiente de la línea generada en un análisis de 2 colores para normalizar las proporciones (detección de prueba:control) en cada experimento. Posteriormente se compararon las proporciones normalizadas de varios experimentos y se utilizaron para identificar los agrupamientos de la expresión del gen. De

este modo, la muestra "de control universal" combinada no solamente permitió una determinación efectiva relativa a la expresión del gen en una simple comparación de 2 muestras sino que también permitió efectuar comparaciones múltiples con diversos experimentos.

En estos experimentos, se utilizaron las sondas de ácido nucleico derivadas de sus secuencias de polipéptido TAT codificando ácido nucleico en la creación del *microarray* y se empleó ARN de varios tejidos de tumores para la hibridación a las mismas. Abajo se muestran los resultados de estos experimentos, comprobando que varios polipéptidos TAT de la presente invención están significativamente sobreexpresados en varios tejidos humanos de tumor cuando se comparan con sus tejidos normales de la contraparte. Adicionalmente, todas las moléculas que se muestran abajo están significativamente sobreexpresadas en su(s) tejido(s) específico(s) de tumor cuando se comparan con el control epitelial "universal". Como se describe arriba, esta información muestra que los polipéptidos TAT de la presente invención son útiles como marcadores de diagnóstico para identificar la presencia de uno o varios tumores cancerosos, y también sirven como agentes terapéuticos para el tratamiento de esos tumores.

Molécula	**regulación por incremento en la expresión en:**	**cuando se compara con:**
DNA95930 (TAT110)	tumor de colon	tejido normal de colon
DNA95930 (TAT110)	tumor de pulmón	tejido normal de pulmón
DNA95930 (TAT110)	tumor de próstata	tejido normal de próstata
DNA95930 (TAT110)	tumor de endometrio	tejido normal de endometrio
DNA95930 (TAT110)	tumor ovárico	tejido ovárico normal
DNA95930-1 (TAT210)	tumor de colon	tejido normal de colon
DNA95930-1 (TAT210)	tumor de pulmón	tejido normal de pulmón
DNA95930-1 (TAT210)	tumor de próstata	tejido normal de próstata
DNA95930-1 (TAT210)	tumor de endometrio	tejido normal de endometrio
DNA95930-1 (TAT210)	tumor ovárico	tejido ovárico normal
DNA96930 (TAT112)	tumor de colon normal	tejido de colon
DNA96930 (TAT112)	tumor de mama	tejido normal de mama

731

DNA96930 (TAT112)	tumor de pulmón	tejido normal de pulmón
DNA96936 (TAT147)	tumor de mama	tejido normal de mama
DNA96936 (TAT147)	tumor de colon	tejido normal de colon
DNA96936 (TAT147)	tumor ovárico	tejido ovárico normal
DNA96936 (TAT147)	tumor de próstata	tejido normal de próstata
DNA108809 (TAT114)	tumor de colon	tejido normal de colon
DNA119488 (TAT119)	tumor de colon	tejido normal de colon
DNA119488 (TAT119)	tumor de pulmón	tejido normal de pulmón
DNA143493 (TAT103)	tumor de mama	tejido normal de mama
DNA181162 (TAT129)	tumor de próstata	tejido normal de próstata
DNA188221 (TAT111)	tumor de colon	tejido normal de colon
DNA188221 (TAT111)	tumor de pulmón	tejido normal de pulmón
DNA188221 (TAT111)	tumor ovárico	tejido ovárico normal
DNA233876 (TAT146)	tumor de colon	tejido normal de colon
DNA233876 (TAT146)	tumor de pulmón	tejido normal de pulmón
DNA233876 (TAT146)	tumor ovárico	tejido ovárico normal
DNA210499 (TAT123)	tumor ovárico	tejido ovárico normal
DNA210499 (TAT123)	tumor de pulmón	tejido normal de pulmón
DNA219894 (TAT211)	tumor ovárico	tejido ovárico normal
DNA219894 (TAT211)	tumor de pulmón	tejido normal de pulmón
DNA215609 (TAT113)	tumor de colon	tejido normal de colon
DNA220432 (TAT128)	tumor de próstata	tejido normal de próstata
DNA226165 (TAT122)	tumor de mama	tejido normal de mama
DNA226165 (TAT122)	tumor de colon	tejido normal de colon
DNA226165 (TAT122)	tumor de recto	tejido normal de recto
DNA226165 (TAT122)	tumor de pulmón	tejido normal de pulmón
DNA226165 (TAT122)	tumor ovárico	tejido ovárico normal
DNA226165 (TAT122)	tumor de próstata	tejido normal de próstata
DNA226456 (TAT144)	tumor de mama	tejido normal de mama
DNA226456 (TAT144)	tumor de colon	tejido normal de colon
DNA237637 (TAT188)	tumor de mama	tejido normal de mama
DNA237637 (TAT188)	tumor de colon	tejido normal de colon
DNA226539 (TAT126)	tumor de recto	tejido normal de recto
DNA226539 (TAT126)	tumor de colon	tejido normal de colon
DNA226539 (TAT126)	tumor de pulmón	tejido normal de pulmón
DNA226539 (TAT126)	tumor ovárico	tejido ovárico normal

DNA236511 (TAT151)	tumor de recto	tejido normal de recto
DNA236511 (TAT151)	tumor de colon	tejido normal de colon
DNA236511 (TAT151)	tumor de pulmón	tejido normal de pulmón

Molécula	**regulación por incremento en la expresión en:**	**cuando se compara con:**
DNA236511 (TAT151)	tumor ovárico	tejido ovárico normal
DNA226771 (TAT115)	tumor de colon	tejido normal de colon
DNA227224 (TAT121)	tumor ovárico	tejido ovárico normal
DNA227224 (TAT121)	tumor de recto	tejido normal de recto
DNA227224 (TAT121)	tumor de colon	tejido normal de colon
DNA227224 (TAT121)	tumor de pulmón	tejido normal de pulmón
DNA227224 (TAT121)	tumor de mama	tejido normal de mama
DNA227224 (TAT121)	tumor de próstata	tejido normal de próstata
DNA247486 (TAT183)	tumor ovárico	tejido ovárico normal
DNA247486 (TAT183)	tumor de recto	tejido normal de recto
DNA247486 (TAT183)	tumor de colon	tejido normal de colon
DNA247486 (TAT183)	tumor de pulmón	tejido normal de pulmón
DNA247486(TAT183)	tumor de mama	tejido normal de mama
DNA247486(TAT183)	tumor de próstata	tejido normal de próstata
DNA228199 (TAT127)	tumor ovárico	tejido ovárico normal
DNA228199 (TAT127)	tumor de pulmón	tejido normal de pulmón
DNA228201 (TAT116)	tumor de colon	tejido normal de colon
DNA247488 (TAT189)	tumor de colon	tejido normal de colon
DNA236538 (TAT190)	tumor de colon	tejido normal de colon
DNA247489 (TAT191)	tumor de colon	tejido normal de colon
DNA228994 (TAT124)	tumor de pulmón	tejido normal de pulmón
DNA228994 (TAT124)	tumor de mama	tejido normal de mama
DNA228994 (TAT124)	tumor ovárico	tejido ovárico normal
DNA231312 (TAT143)	tumor de colon	tejido normal de colon
DNA231542 (TAT100)	tumor cerebral	tejido cerebral normal
DNA231542 (TAT100)	glioma	tejido glial normal
DNA231542-1 (TAT284)	tumor cerebral	tejido cerebral normal
DNA231542-1 (TAT284)	glioma	tejido glial normal
DNA231542-2 (TAT285)	tumor cerebral	tejido cerebral normal
DNA231542-2 (TAT285)	glioma	tejido glial normal

DNA297393 (TAT285-1)	tumor cerebral	tejido cerebral normal
DNA297393 (TAT285-1)	glioma	tejido glial normal
DNA236246 (TAT153)	tumor de mama	tejido normal de mama
DNA236343 (TAT104)	tumor de mama	tejido normal de mama
DNA236534 (TAT102)	tumor de mama	tejido normal de mama
DNA236534 (TAT102)	tumor de colon	tejido normal de colon
DNA246430 (TAT109)	tumor de próstata	tejido normal de próstata
DNA264454 (TAT106)	tumor de mama	tejido normal de mama
DNA98565 (TAT145)	glioma	tejido cerebral normal
DNA246435 (TAT152)	glioma	tejido cerebral normal
DNA226094 (TAT164)	glioma	tejido cerebral normal

EJEMPLO 3: Análisis Cuantitativo de la Expresión del ARNm del TAT

En este análisis, se empleó un análisis 5' nucleasa (por ejemplo, TaqMan®) PCR cuantitativo en tiempo real (por ejemplo, ABI Prizm 7700 Sequence Detection System® (Perkin Elmer, Applied Biosystems Division, Foster City, CA)), para hallar los genes que se encuentran significativamente sobreexpresados en un tumor o unos tumores cancerosos cuando se comparan con otros tumores cancerosos o con tejido normal no canceroso. La reacción del análisis 5' es una técnica fluorescente a base de PCR que utiliza la actividad 5' exonucleasa de la enzima Taq ADN polimerasa para monitorear la expresión del gen en tiempo real. Se utilizan dos oligonucleótidos iniciadores (cuyas secuencias se basan en el gen o la secuencia EST de interés) para generar un amplicón típico de una reacción PCR. Se designa un tercer oligonucleótido o sonda, para detectar la secuencia de nucleótido ubicada entre los dos iniciadores PCR. La sonda no es extensible por la enzima Taq ADN polimerasa y está marcada con un tinte fluorescente indicador y un tinte extinguidor. Cualquier emisión inducida por láser del tinte indicador se extingue mediante el tinte para extinguir cuando los dos tintes se encuentran cerca entre sí pues no están en la sonda. Durante la reacción de amplificación de la PCR, la enzima Taq

ADN polimerasa segmenta la sonda en una forma dependiente de plantilla. Los fragmentos resultantes de la sonda se disocian en la solución y la señal del tinte indicador liberada se libera del efecto de extinción del segundo fluoróforo. Se libera una molécula de tinte indicador por cada nueva molécula sintetizada y la detección del tinte indicador sin extinguir dispone la base de la interpretación cuantitativa de los datos.

El procedimiento 5' nucleasa se realiza en un dispositivo de PCR cuantitativo en tiempo real como el ABI Prism 7700TM Sequence Detection. El sistema consiste de un termociclador, un láser, un dispositivo de carga acoplada (CCD), una cámara y un computador. El sistema amplifica las muestras en un formato de 96 pozos en un termociclador. Durante la amplificación, la señal fluorescente inducida por el láser se recolecta en tiempo real mediante unas fibras de cable óptico de los 96 pozos y se detecta en el CCD. El sistema incluye un software para activar el instrumento y para analizar los datos.

El material de partida para el cribado fue ARNm aislado de una variedad de diferentes tejidos cancerosos. El ARNm se cuantifica de forma precisa, por ejemplo, fluorométricamente. Como un control negativo, se aisló ARN de varios tejidos normales del mismo tipo de tejido que el de los tejidos cancerosos que se estaban analizando.

Inicialmente los datos del análisis de 5' nucleasa se expresan como Ct o el ciclo umbral. Esto se define como el ciclo donde la señal indicadora se acumula por encima del nivel de entorno de la fluorescencia. Los valores ΔCt se utilizan como mediciones cuantitativas del número relativo de las copias de partida de una secuencia objetivo particular en una secuencia de ácido nucleico cuando se comparan los resultados del ARNm con cáncer con los resultados del ARNm normal. Puesto que una unidad Ct corresponde a 1 ciclo de la PCR o aproximadamente a un incremento relativo de 2 veces el normal, dos unidades corresponden a un incremento relativo de 4 veces, 3 unidades corresponden a un incremento relativo de 8 veces y así sucesivamente, cuantitativamente se puede medir el

multiplicador del incremento en la expresión del ARNm entre dos o más tejidos diferentes. Utilizando esta técnica, las moléculas enumeradas abajo se han identificado como significativamente sobreexpresadas en un(os) tumor(es) particular(es) cuando se comparan con su contraparte normal no cancerosa, (tanto del mismo donante como de diferentes donantes) y de este modo, representan excelentes blancos de polipéptidos para el diagnóstico y la terapia de cáncer en mamíferos.

Molécula	**regulación por incremento en la expresión en:**	**cuando se compara con:**
DNA77507 (TAT161)	tumor de mama	tejido normal de mama
DNA82343 (TAT157)	tumor de colon	tejido normal de colon
DNA88131 (TAT158)	tumor de mama	tejido normal de mama
DNA88131 (TAT158)	tumor de colon	tejido normal de colon
DNA95930 (TAT110)	tumor de colon	tejido normal de colon
DNA95930 (TAT110)	tumor de pulmón	tejido normal de pulmón
DNA95930 (TAT110)	tumor de próstata	tejido normal de próstata
DNA95930 (TAT110)	tumor de endometrio	tejido normal de endometrio
DNA95930 (TAT110)	tumor ovárico	tejido ovárico normal
DNA95930-1 (TAT210)	tumor de colon	tejido normal de colon
DNA95930-1 (TAT210)	tumor de pulmón	tejido normal de pulmón
DNA95930-1 (TAT210)	tumor de próstata	tejido normal de próstata
DNA95930-1 (TAT210)	tumor de endometrio	tejido normal de endometrio

Molécula	**regulación por incremento en la expresión en:**	**cuando se compara con:**
DNA95930-1 (TAT210)	tumor ovárico	tejido ovárico normal
DNA96930 (TAT112)	tumor de colon	tejido normal de colon
DNA96936 (TAT147)	tumor de colon	tejido normal de colon
DNA98591 (TAT162)	tumor de colon	tejido normal de colon
DNA108809 (TAT114)	tumor de riñón	tejido normal de riñón
DNA119488 (TAT119)	tumor de pulmón	tejido normal de pulmón
DNA188221 (TAT111)	tumor de colon	tejido normal de colon
DNA233876 (TAT146)	tumor de colon	tejido normal de colon

DNA193891 (TAT148)	tumor de colon	tejido normal de colon
DNA248170 (TAT187)	tumor de colon	tejido normal de colon
DNA194628 (TAT118)	tumor de riñón	tejido normal de riñón
DNA246415 (TAT167)	tumor de riñón	tejido normal de riñón
DNA210499 (TAT123)	tumor de pulmón	tejido normal de pulmón
DNA219894 (TAT211)	tumor de pulmón	tejido normal de pulmón
DNA215609 (TAT113)	tumor de colon	tejido normal de colon
DNA220432 (TAT128)	tumor de próstata	tejido normal de próstata
DNA226165 (TAT122)	tumor de pulmón	tejido normal de pulmón
DNA226237 (TAT117)	tumor de riñón	tejido normal de riñón
DNA246450 (TAT168)	tumor de riñón	tejido normal de riñón
DNA226456 (TAT144)	tumor de mama	tejido normal de mama
DNA237637 (TAT188)	tumor de mama	tejido normal de mama
DNA226539 (TAT126)	tumor ovárico	tejido ovárico normal
DNA236511 (TAT151)	tumor ovárico	tejido ovárico normal
DNA227224 (TAT121)	tumor de pulmón	tejido normal de pulmón
DNA247486 (TAT183)	tumor de pulmón	tejido normal de pulmón
DNA227800 (TAT131)	tumor de próstata	tejido normal de próstata
DNA228199 (TAT127)	tumor ovárico	tejido ovárico normal
DNA228199 (TAT127)	tumor de pulmón	tejido normal de pulmón
DNA228201 (TAT116)	tumor de colon	tejido normal de colon
DNA247488 (TAT189)	tumor de colon	tejido normal de colon
DNA236538 (TAT190)	tumor de colon	tejido normal de colon
DNA247489 (TAT191)	tumor de colon	tejido normal de colon
DNA228993 (TAT120)	tumor de pulmón	tejido normal de pulmón
DNA228994 (TAT124)	tumor de pulmón	tejido normal de pulmón
DNA236343 (TAT104)	tumor de mama	tejido normal de mama
DNA236534 (TAT102)	tumor ovárico	tejido ovárico normal
DNA246430 (TAT109)	tumor de mama	tejido normal de mama
DNA247480 (TAT142)	tumor de pulmón	tejido normal de pulmón
DNA98565 (TAT145)	glioma	tejido cerebral normal
DNA246435 (TAT152)	glioma	tejido cerebral normal
DNA226094 (TAT164)	glioma	tejido cerebral normal
DNA227578 (TAT165)	glioma	tejido cerebral normal
DNA231542 (TAT100)	glioma	tejido cerebral normal
DNA231542-1 (TAT284)	glioma	tejido cerebral normal

DNA231542-2 (TAT285)	glioma	tejido cerebral normal
DNA297393 (TAT285-1)	glioma	tejido cerebral normal

EJEMPLO 4: Hibridación *In situ*

La hibridación *in situ* es una técnica poderosa y versátil para la detección y localización de las secuencias de ácido nucleico en una célula o en las preparaciones de tejidos. Puede ser útil, por ejemplo, para identificar sitios de expresión de un gen, analizar la distribución del tejido de la transcripción, identificar y localizar una infección viral, seguir los cambios en la síntesis de un ARNm específico y ayudar al mapeo de los cromosomas.

La hibridación *in situ* se completó siguiendo una versión optimizada del protocolo por Lu and Gillett, Cell Vision 1:169-176 (1994), utilizando ribosondas generadas por PCR de marcador ^{33}P teniendo una homología con la secuencia objetivo que se va a detectar. Brevemente, los tejidos humanos fijados con formalina con inclusión de parafina se seccionaron, se retiró la parafina, se desproteinizó en proteinasa K (20 g/ml) durante 15 minutos a 37°C y luego se procesaron para la hibridación *in situ* como se describe por Lu and Gillett, *supra*. Se generó una ribosonda antisentido con marcación A[^{33}-P]UTP a partir de un producto de la PCR y se hibridó a 55°C durante la noche. Las láminas se sumergieron en la emulsión de rastreo nuclear Kodak NTB2 y se dejaron expuestas a la misma durante 4 semanas.

Síntesis de ^{33}P-ribosonda

Mediante *speed vac* se secaron 6.0 µl (125 mCi) de ^{33}P-UTP (Amersham BF 1002, SA<2000 Ci/mmol). A cada tubo conteniendo ^{33}P-UTP seco, se agregaron los siguientes ingredientes:

2.0 µl de 5x solución amortiguadora de transcripción

1.0 µl de DTT (100 mM)

2.0 µl de NTP mezcla (2.5 mM : 10 µ; cada uno de 10 mM GTP, CTP & ATP + 10 µl H_2O)

1.0 µl de UTP (50 µM)

1.0 µl de ARNsin

1.0 µl de plantilla de ADN (1µg)

1.0 µl de H_2O

1.0 µl de ARN polimerasa (para productos PCR T3 = AS, T7 = S, normalmente).

Los tubos se incubaron a 37°C durante una hora. Se añadieron 1.0 µl de RQ1 DNasa seguido por incubación a 37°C durante 15 minutos. Se agregaron 90 µl de TE (10 mM Tris pH 7.6/1mM EDTA pH 8.0) y la mezcla se dispensó mediante pipeta sobre papel DE81. El resto de la solución se cargó en una unidad de ultrafiltración Microcon-50 y se centrifugó utilizando el programa 10 (6 minutos). La unidad de filtración se invirtió sobre un segundo tubo y se centrifugó empleando el programa 2(3 minutos). Después de la centrifugada de recuperación final, se añadieron 100 µl de TE. Se dispensó 1 µl del producto mediante pipeta sobre papel DE81 y se contó en 6 ml de Biofluor II.

La sonda se corrió en un gel de TBE/urea. Se agregaron 1-3 µl de la sonda o 5 µl de ARN Mrk III a 3 µl de solución amortiguadora de carga. Después de calentarla en un bloque de calor a 95°C durante tres minutos, inmediatamente se colocó la sonda sobre hielo. Los pozos de gel se lavaron con chorro, la muestra se cargó y se operó a 180-250 voltios durante 45 minutos. El gel se envolvió en sarán y se expuso a una película XAR con una pantalla intensificadora en un congelador a -70°C desde una hora hasta toda la noche.

^{33}P-Hibridación

A. Pretratamiento de secciones congeladas

Las láminas se retiraron del congelador, se colocaron sobre bandejas de aluminio y se dejaron descongelar a temperatura ambiente durante 5 minutos. Las bandejas se colocaron en una incubadora a 55°C durante cinco minutos para reducir la condensación. Las láminas se fijaron durante 10 minutos en paraformaldehido al 4% sobre hielo en la campana de humos

y se lavaron en 0.5 x SSC durante 5 minutos, a temperatura ambiente (25 ml 20 x SSC + 975 ml SQ H_2O). Después de la desproteinación en 0.5 µg/ml de proteinasa K durante 10 minutos a 37°C (12.5 µl de solución matriz 10 mg/ml en 250 ml de solución amortiguadora precalentada ARNsa-sin ARNsa), las secciones se lavaron en 0.5 x SSC durante 10 minutos a temperatura ambiente. Las secciones se deshidrataron con etanol al 70%, 95%, 100%, 2 minutos cada una.

B. Pretratamiento de las secciones con inclusión de parafina

Las láminas se desparafinaron, se colocaron en SQ H_2O y se enjuagaron dos veces en 2 x SSC a temperatura ambiente, durante 5 minutos cada vez. Las secciones se desproteinizaron en 20 µg/ml de proteinasa K (500 µl de 10 mg/ml en 250 ml de solución amortiguadora RNasa - sinRNasa; 37°C, 15 minutos) - embrión humano u 8 x proteinasa K (100 µl en 250 ml de solución amortiguadora RNasa, 37°C, 30 minutos) – tejidos de formalina. Se completaron enjuagues posteriores en 0.5 x SSC y deshidratación como se describe arriba.

C. Prehibridación

Las láminas se colocaron en una caja plástica forrada con solución amortiguadora Box (4 x SSC, formamida al 50%) – papel filtro saturado.

D. Hibridación

Se calentaron 1.0 x 10^6 cpm sonda y 1.0 µl de ARNt (50 mg/ml de solución matriz) por lámina a 95°C durante 3 minutos. Las láminas se enfriaron sobre hielo y se agregaron 48 µl de solución amortiguadora de hibridación por lámina. Después de girar en espiral, se añadieron 50µl de mezcla ^{33}P a 50 µl de prehibridación en la lámina. Las láminas se incubaron durante la noche a 55°C.

E. Lavados

Se realizaron lavados 2 x 10 minutos con 2xSSC, EDTA a temperatura ambiente (400 ml 20 x SSC + 16 ml 0.25M EDTA, V_f=4L), seguido por tratamiento con RNasaA a 37°C durante 30 minutos (500 µl de 10 mg/ml en 250 ml de solución amortiguadora Rnasa = 20 µg/ml). Las

láminas se lavaron 2 x 10 minutos con 2 x SSC, EDTA a temperatura ambiente. Las severas condiciones de lavado fueron como sigue: 2 horas a 55°C, 0.1 x SSC, EDTA (20 ml 20 x SSC + 16 ml EDTA, V_f=4L).

F. Oligonucleótidos

Los análisis *in situ* se realizaron en una variedad de secuencias de ADN divulgadas en la presente. Los oligonucleótidos empleados para estos análisis se obtuvieron de manera que fueran complementarios de los ácidos nucleicos (o los complementos de los mismos) como se muestra en las figuras adjuntas.

G. Resultados

Los análisis *in situ* se realizaron en una variedad de secuencias de ADN divulgadas en la presente. Los resultados de estos análisis son los siguientes.

(1) DNA95930 (TAT110)

En un análisis, se observa una expresión significativa en 3/3 tumores de pulmón, 3/3 adenocarcinomas colorrectales, 1/1 neoplasias de próstata, 3/3 carcinomas de células transicionales y 3/3 adenocarcinomas de endometrio, donde el nivel de expresión en los tejidos normales de la contraparte es significativamente inferior.

En un segundo análisis independiente, se observa una expresión significativa en 7/7 adenocarcinomas de endometrio y en 12/15 adenocarcinomas de ovario, donde el nivel de expresión en los tejidos normales de la contraparte es significativamente inferior.

En un tercer análisis independiente, se observa una expresión significativa en 24/26 muestras de tumor colorrectal, donde el nivel de expresión en el tejido normal de la contraparte es significativamente inferior.

Finalmente, en un cuarto análisis independiente, se observa la expresión en 8/26 muestras de tejido prostático no maligno, 55/82 muestras de cáncer prostático primario y en 5/23 muestras de cáncer prostático metastásico.

(2) DNA95930-1 (TAT210)

En un análisis, se observa una expresión significativa en 3/3 tumores de pulmón, 3/3 adenocarcinomas colorrectales, 1/1 neoplasias de próstata, 3/3 carcinomas de células transicionales y 3/3 adenocarcinomas de endometrio, donde el nivel de expresión en los tejidos normales de la contraparte es significativamente inferior.

En un segundo análisis independiente, se observa una expresión significativa en 7/7 adenocarcinomas de endometrio y 12/15 adenocarcinomas de ovario, donde el nivel de expresión en los tejidos normales de la contraparte es significativamente inferior.

En un tercer análisis independiente, se observa una expresión significativa en 24/26 muestras de tumor colorrectal, donde el nivel de expresión en el tejido normal de la contraparte es significativamente inferior.

Finalmente, en un cuarto análisis independiente, se observa la expresión en 8/26 muestras de tejido prostático no maligno, 55/82 muestras de cáncer prostático primario y en 5/23 muestras de cáncer prostático metastásico.

(3) DNA96930 (TAT112)

Fuerte expresión en las neoplasias colorrectales. La expresión en el epitelio maligno aparece significativamente más fuerte que en el epitelio benigno adyacente. Adicionalmente, se observa la expresión marcada en todas las 23 de las 23 muestras de adenocarcinoma pancreático probado, donde la expresión en el tejido pancreático normal no es detectable.

(4) DNA96936 (TAT147)

En un análisis, se observó una señal fuertemente positiva en 6/6 tumores de mama. En otro análisis independiente, se observó una señal positiva en 4/4 carcinomas no microcíticos de pulmón, donde los tumores parecen tener una mayor expresión comparada con un pulmón normal. 1/1

adenocarcinomas de endometrio muestran una fuerte expresión y 3/3 adenocarcinomas colorrectales muestran una expresión variable.

(5) NA10809 (TAT114)

Una señal positiva en todos los carcinomas de célula renal probados (n=3) aunque no se observa una expresión en el tejido normal de riñón. Adicionalmente, se observa una expresión positiva en 5/12 tumores de estómago, 5/24 tumores colorrectales, 3/8 tumores pancreáticos y 1/3 tumores de pulmón. La expresión de tejido normal no canceroso se limita al estómago y al intestino delgado.

(6) DNA176766 (TAT132)

Una señal positiva en todos los adenocarcinomas de endometrio probados (n=3) aunque no se observa ninguna expresión en el tejido normal de endometrio.

(7) DNA236463 (TAT150)

Una señal positiva en todos los adenocarcinomas de endometrio probados (n=3) aunque no se observa ninguna expresión en el tejido normal de endometrio.

(8) DNA181162 (TAT129)

Generalmente el epitelio de próstata neoplásico es positivo, y las intensidades de señal varían de débil a fuerte entre casos. Los tejidos no prostáticos son negativos.

(9) DNA188221 (TAT111)

Se ha visto una señal fuerte en arreglos colónicos multitumorales en el epitelio maligno. En los tejidos normales, determinada sonda produjo una señal específica en las células epiteliales recubriendo los 2/3 de las criptas colónicas, la intensidad de la señal pareció ser significativamente inferior que en los carcinomas colónicos. Se observa una expresión positiva en 12/18 adenocarcinomas colorrectales, 6/8 adenocarcinomas metastásicos y 2/9 adenocarcinomas gástricos.

(10) DNA233876 (TAT146)

Se ha visto una señal fuerte en arreglos colónicos multitumorales en el epitelio maligno. En los tejidos normales, determinada sonda produjo una señal específica en las células epiteliales recubriendo los 2/3 de las criptas colónicas, la intensidad de la señal pareció ser significativamente inferior que en los carcinomas colónicos. Se observa una expresión positiva en 12/18 adenocarcinomas colorrectales, 6/8 adenocarcinomas metastásicos y 2/9 adenocarcinomas gástricos.

(11) DNA210499 (TAT123)

En un análisis, 12/14 adenocarcinomas de ovario son positivos y 8/9 adenocarcinomas de endometrio son positivos. El estroma ovárico normal es negativo como lo es el miometrio uterino. Otros tejidos ováricos y uterinos normales son negativos.

En un análisis independiente, 16/27 carcinomas no microcíticos de pulmón son positivos, donde la señal es moderada o fuerte.

(12) DNA219894 (TAT211)

En un análisis, 12/14 los adenocarcinomas de ovario son positivos y 8/9 adenocarcinomas de endometrio son positivos. El estroma ovárico normal es negativo como lo es el miometrio uterino. Otros tejidos ováricos y uterinos normales son negativos.

En un análisis independiente, 16/27 carcinomas no microcíticos de pulmón son positivos, donde la señal es moderada o fuerte.

(13) DNA215609 (TAT113)

Se ve una señal fuerte en los carcinomas colónicos, y solo un nivel de señal muy bajo en colon normal. Los carcinomas de pulmón y de mama fueron negativos.

(14) DNA220432 (TAT128)

El único tejido adulto normal expresando este gen es el epitelio prostático. La expresión tiene una intensidad de moderada a fuerte y focal, y es más prevalente en el epitelio hiperplásico.

En un análisis donde están disponibles 50 casos de cáncer prostático primario para su revisión, 29 casos (58%) son positivos, 18 casos (36%) son

negativos y 3 casos (6%) son dudosos. En otro análisis donde están disponibles 37 casos de cáncer de próstata primario para revisión, 33 casos (89%) son positivos, 4 casos (11%) son negativos. Finalmente, en otro análisis independiente donde 27 casos de cáncer de próstata metastásicos están disponibles para revisión, 14 casos (52%) son positivos, 11 casos (41%) son negativos y 2 casos (7%) son dudosos.

(15) DNA226237 (TAT117)

En un análisis, dos de 3 carcinomas de células renales son positivos, donde la expresión del riñón normal es negativa.

(16) DNA246450 (TAT168)

En un análisis, dos de 3 carcinomas de células renales son positivos, donde la expresión del riñón normal es negativa.

(17) DNA227087 (TAT163)

Una sonda para esta molécula mostró una señal positiva en una subpoblación de células estromales asociadas a tumor en todos los casos probados de carcinomas de pulmón, mama, colon, páncreas y endometrio. Con frecuencia, la intensidad del marcado fue bastante fuerte. En un caso de adenocarcinoma de colon con colon benigno adyacente, el marcado se restringía al estroma asociado al tumor y el tejido normal benigno fue negativo. Un fibroadenoma de mama también mostró el marcado de células estromales subepiteliales.

(18) DNA266307 (TAT227)

Una sonda para esta molécula mostró una señal positiva en una subpoblación de células estromales asociadas a tumor en todos los casos probados de carcinomas de pulmón, mama, colon, páncreas y endometrio. Con frecuencia, la intensidad del marcado fue bastante fuerte. En un caso de adenocarcinoma de colon con colon benigno adyacente, el marcado se restringía al estroma asociado al tumor y el tejido normal benigno fue negativo. Un fibroadenoma de mama también mostró el marcado de células estromales subepiteliales.

(19) DNA266311 (TAT228)

Una sonda para esta molécula mostró una señal positiva en una subpoblación de células estromales asociadas a tumor en todos los casos probados de carcinomas de pulmón, mama, colon, páncreas y endometrio. Con frecuencia, la intensidad del marcado fue bastante fuerte. En un caso de adenocarcinoma de colon con colon benigno adyacente, el marcado se restringía al estroma asociado al tumor y el tejido normal benigno fue negativo. Un fibroadenoma de mama también mostró el marcado de células estromales subepiteliales.

(20) DNA266312 (TAT229)

Una sonda para esta molécula mostró una señal positiva en una subpoblación de células estromales asociadas a tumor en todos los casos probados de carcinomas de pulmón, mama, colon, páncreas y endometrio. Con frecuencia, la intensidad del marcado fue bastante fuerte. En un caso de adenocarcinoma de colon con colon benigno adyacente, el marcado se restringía al estroma asociado al tumor y el tejido normal benigno fue negativo. Un fibroadenoma de mama también mostró el marcado de células estromales subepiteliales.

(21) DNA266313 (TAT230)

Una sonda para esta molécula mostró una señal positiva en una subpoblación de células estromales asociadas a tumor en todos los casos probados de carcinomas de pulmón, mama, colon, páncreas y endometrio. Con frecuencia, la intensidad del marcado fue bastante fuerte. En un caso de adenocarcinoma de colon con colon benigno adyacente, el marcado se restringía al estroma asociado al tumor y el tejido normal benigno fue negativo. Un fibroadenoma de mama también mostró el marcado de células estromales subepiteliales.

(22) DNA227224 (TAT121)

Se aprecia la expresión en 2 de 3 adenocarcinomas de endometrio.

(23) DNA247486 (TAT183)

Se aprecia la expresión en 2 de 3 adenocarcinomas de endometrio.

(24) DNA227800 (TAT131)

En un análisis, 46/64 neoplasias primarias de próstata son positivas y 6/14 neoplasias metastásicas de próstata son positivas. Débil a moderada. Se aprecia la expresión en el epitelio de la próstata.

(25) DNA228199 (TAT127)

Se observa la expresión en 13 de 15 tumores ováricos (adenocarcinoma y tumores de superficie epitelial). El epitelio benigno de la superficie ovárica también es positivo. El nivel de expresión en la mayoría de los tumores positivos es fuerte o moderado y bastante uniforme. Igualmente se observa la expresión en 8 de 9 adenocarcinomas uterinos. Siete de 23 carcinomas no microcíticos de pulmón son positivos.

(26) DNA228201 (TAT116)

Las células malignas de 13/16 adenocarcinomas colorrectales son positivas para la expresión del TAT116. Adicionalmente, 9/10 adenocarcinomas metastásicos son positivos para la expresión. Igualmente se observa la expresión en las porciones basales de criptas colónicas normales.

(27) DNA247488 (TAT189)

Las células malignas de 13/16 adenocarcinomas colorrectales son positivas para la expresión del TAT189. Adicionalmente, 9/10 adenocarcinomas metastásicos son positivos para la expresión. Igualmente se observa la expresión en las porciones basales de criptas colónicas normales.

(28) DNA236538 (TAT190)

Las células malignas de 13/16 adenocarcinomas colorrectales son positivas para la expresión del TAT 190. Adicionalmente, 9/10 adenocarcinomas metastásicos son positivos para la expresión. Igualmente se observa la expresión en las porciones basales de criptas colónicas normales.

(29) DNA247489 (TAT191)

Las células malignas de 13/16 adenocarcinomas colorrectales son positivas para la expresión del TAT191. Adicionalmente, 9/10

adenocarcinomas metastásicos son positivos para la expresión. Igualmente se observa la expresión en las porciones basales de criptas colónicas normales.

(30) DNA228994 (TAT124)

Trece de 61 casos de carcinomas no microcíticos de pulmón son positivos para la expresión del TAT124. El nivel de expresión en estas muestras positivas de tumor es significativamente más alto que en tejidos adultos normales.

(31) DNA231542 (TAT100)

Un análisis *in situ* realizado como se describe arriba evidencia una expresión regulada por incremento de forma significativa en el glioma humano y en tejidos de glioblastoma cuando se compara con tejidos normales de cerebro (y otros tejidos).

(32) DNA231542-1 (TAT284)

Un análisis *in situ* realizado como se describe arriba evidencia una expresión regulada por incremento de forma significativa en el glioma humano y en tejidos de glioblastoma cuando se compara con tejidos normales de cerebro (y otros tejidos).

(33) DNA231542-2 (TAT285)

Un análisis *in situ* realizado como se describe arriba evidencia una expresión regulada por incremento de forma significativa en el glioma humano y en tejidos de glioblastoma cuando se compara con tejidos normales de cerebro (y otros tejidos).

(34) DNA297393 (TAT285-1)

Un análisis *in situ* realizado como se describe arriba evidencia una expresión regulada por incremento de forma significativa en el glioma humano y en tejidos de glioblastoma cuando se compara con tejidos normales de cerebro (y otros tejidos).

(35) DNA236534 (TAT102)

La expresión del TAT102 se aprecia en 14 de 15 neoplasias epiteliales de ovario (adenocarcinoma, tumores de superficie epitelial, Ca

endometrioide). También, 8 de 9 adenocarcinomas de endometrio del útero expresan TAT102. Además, la expresión del TAT102 se aprecia en 24 de 27 neoplasias no microcíticas de pulmón, los casos positivos incluyen los escamosos y los adenocarcinomas. La expresión en estos tejidos de tumor es significativamente mayor que en sus contrapartes de tejidos normales.

(36) DNA246430 (TAT109)

Catorce de las 92 muestras de tumor de mama son positivos para la expresión del TAT109. La expresión en todos los tejidos normales no es detectable.

(37) DNA264454 (TAT106)

Se observa la expresión del TAT106 en 38/88 tumores de mama. La expresión en el tejido normal de mama es débil o no es detectable.

(38) DNA98565 (TAT145)

Se observaron señales positivas de TAT145 en la mayoría de los gliomas, glioblastomas, en algunos melanomas y en cerebro normal (localizado fundamentalmente en los astrocitos). La intensidad de la señal en los glioblastomas pareció ser mayor que en los astrocitos normales. Aunque la mayoría de las muestras de glioma y glioblastoma probadas fueron positivas para la expresión del TAT145, la mayoría de las muestras normales de cerebro probadas fueron negativas para dicha expresión.

(39) DNA246435 (TAT152)

Se observaron señales positivas de TAT145 en la mayoría de los gliomas, glioblastomas, en algunos melanomas y en cerebro normal (localizado fundamentalmente en los astrocitos). La intensidad de la señal en los glioblastomas pareció ser mayor que en los astrocitos normales. Aunque la mayoría de las muestras de glioma y glioblastoma probadas fueron positivas para la expresión del TAT152, la mayoría de las muestras normales de cerebro probadas fueron negativas para dicha expresión.

(40) DNA167234 (TAT130)

Setenta casos de adenocarcinoma primario de próstata estuvieron disponibles para revisión. De estos 70 casos, 56 casos (80%) son positivos

para la expresión del TAT130. La expresión del TAT130 en tejidos no prostáticos es débil o no es detectable.

(41) DNA235621 (TAT166)

Setenta casos de adenocarcinoma primario de próstata estuvieron disponibles para revisión. De estos 70 casos, 56 casos (80%) son positivos para la expresión del TAT166. La expresión del TAT166 en tejidos no prostáticos es débil o no es detectable.

(42) DNA236493 (TAT141)

Se observa una expresión positiva en 70/148 carcinomas de mama, 2/63 adenocarcinomas colorrectales, 4/42 tumores ováricos, 9/69 carcinomas no microcíticos de pulmón, 9/67 adenocarcinomas de próstata y 5/25 gliomas. La expresión en tejidos normales no cancerosos parece estar restringida al epitelio de próstata y de mama.

(43) DNA226094 (TAT164)

Veintiuno de 37 muestras de glioblastoma y 8 u 8 muestras de glioma fueron positivas para la expresión del TAT164 mientras que todos los otros tejidos de tumor y normales examinados (incluyendo tejido cerebral normal) fueron negativos.

(44) DNA227578 (TAT165)

Quince de 25 muestras de glioblastoma probadas fueron positivas para la expresión mientras que se observó una expresión significativamente más débil en las muestras de tejido normal probadas.

EJEMPLO 5: Análisis de Inmunohistoquímica

Se prepararon anticuerpos contra ciertos polipéptidos TAT divulgados en la presente y se efectuó un análisis de inmunohistoquímica como sigue. Primero se fijaron unas secciones de tejido durante 5 minutos en acetona/etanol (congelados o con inclusión de parafina). Luego se lavaron las secciones en PBS y después se bloquearon con avidina y biotina (Estuche Vector) durante 10 minutos cada una seguido por lavado en PBS. Después las secciones se bloquearon con suero al 10% durante 20 minutos y

luego se sometió a transferencia para retirar el excedente. Posteriormente se añadió un anticuerpo primario a las secciones a una concentración de 10µg/ml durante 1 hora y luego las secciones se lavaron en PBS. Después se añadió un anticuerpo secundario biotinilado (anticuerpo anti-primario) a las secciones durante 30 minutos y después se lavaron las secciones con PBS. Después se expusieron las secciones a los reactivos del estuche Vector ABC durante 30 minutos y después se lavaron las secciones en PBS. Luego se expusieron las secciones a Diaminobencidina (Pierce) durante 5 minutos y después se lavaron en PBS. Luego las secciones se contramarcaron con hematoxilina de Mayers, se recubrieron con un cubreobjetos y se visualizaron. Igualmente se pueden realizar análisis de inmunohistoquímica como se describe en Sambrook y colaboradores, Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989 y Ausubel y colaboradores, Current Protocols of Molecular Biology, Unidad 3.16, John Wiley and Sons (1997). Los resultados de estos análisis se muestran abajo.

(1) DNA96930 (TAT112)

Se detectó una expresión significativamente mayor en la superficie apical de las criptas colónicas de los tumores de colon que en la superficie apical de las criptas colónicas normales. Adicionalmente, se encontró que el TAT112 estaba significativamente sobreexpresado en las células del adenocarcinoma pancreático cuando se compararon con células pancreáticas normales. Finalmente, el análisis IHC efectuado como se describe arriba evidenció que el TAT112 está significativamente sobreexpresado en el carcinoma de pulmón cuando se compara con tejido normal de pulmón, carcinomas no microcíticos de pulmón cuando se comparan con tejido normal de pulmón y carcinoma de estómago cuando se compara con tejido normal de estómago.

(2) DNA226539 (TAT126)

Se observa una expresión positiva en 2/10 adenocarcinomas uterinos, 9/17 adenocarcinomas de ovario y 2/20 carcinomas no microcíticos de

pulmón. Utilizando este procedimiento, la expresión del TAT126 no fue detectable en ningún tejido normal.

(3) DNA236511 (TAT151)

Se observa una expresión positiva en 2/10 adenocarcinomas uterinos, 9/17 adenocarcinomas de ovario y 2/20 carcinomas no microcíticos de pulmón. Utilizando este procedimiento, la expresión del TAT151 no fue detectable en ningún tejido normal.

EJEMPLO 6: Verificación y Análisis de la Expresión Diferencial del Polipéptido TAT Mediante GEPIS

Los polipéptidos TAT que se pueden haber identificado como un antígeno de tumor como se describe en uno o varios de los Ejemplos anteriores se analizaron y verificaron como sigue. Se buscó un marcador de secuencia expresado (EST) base de datos de ADN (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) y se identificaron unas secuencias EST interesantes mediante GEPIS. Un perfil de expresión de gen en *sílice* (GEPIS) es una herramienta bioinformática desarrollada en Genentech, Inc. que caracteriza los genes de interés para nuevos blancos terapéuticos para el cáncer. GEPIS aprovecha las grandes cantidades de secuencias EST y la información de la biblioteca para determinar los perfiles de expresión de los genes. GEPIS es capaz de determinar el perfil de expresión de un gen con base en su correlación proporcional con el número de sus ocurrencias en las bases de datos EST, y funciona integrando la base de datos relacional LIFESEQ® EST y la información propietaria de Genentech en una forma rigurosa y estadísticamente significativa. En este ejemplo, se utiliza GEPIS para identificar y covalidar novedosos antígenos de tumores aunque se puede configurar el GEPIS para realizar unos análisis muy específicos o unas tareas de amplio cribado. Para el cribado inicial, se utiliza el GEPIS para identificar las secuencias EST en la base de datos LIFESEQ® que se correlacionan con la expresión en un tejido en particular o en unos tejidos de interés (con frecuencia un tejido de interés de un tumor). Las secuencias

EST identificadas en este cribado inicial (o las secuencias consensuales obtenidas al alinear múltiples secuencias EST relacionadas y sobrepuestas obtenidas del cribado inicial) se sometieron a un cribado con el fin de identificar la presencia de por lo menos un dominio de transmembrana en la proteína codificada. Finalmente, se empleó GEPIS para generar un perfil de expresión completo de un tejido para las diferentes secuencias de interés. Utilizando este tipo de bioinformática de cribado, se identificaron varios polipéptidos TAT (y sus moléculas codificando ácido nucleico) como significativamente sobreexpresadas en un tipo de cáncer en particular o en ciertas neoplasias comparadas con otras neoplasias y/o tejidos normales no cancerosos. La clasificación de las coincidencias GEPIS se basa en diversos criterios incluyendo por ejemplo, la especificidad del tejido, la especificidad del tumor y el nivel de expresión en tejido normal indispensable y/o tejidos normales de proliferación. A continuación se encuentra una lista de moléculas cuyo perfil de expresión del tejido según lo determina GEPIS evidencia una alta expresión de tejido y una significativa regulación por incremento de expresión en un tumor o en unos tumores específico(s) cuando se compara(n) con otro(s) tumor(es) y/o tejidos normales y una expresión opcionalmente relativamente baja en unos tejidos normales indispensables y/o tejidos normales de proliferación. Como tal, las moléculas enumeradas abajo son excelentes blancos de polipéptidos para el diagnóstico y la terapia en el cáncer de los mamíferos.

Molécula	regulación por incremento en la expresión en:	cuando se compara con:
DNA77507 (TAT161)	tumor de mama	tejido normal de mama
DNA77507 (TAT161)	tumor de colon	tejido normal de colon
DNA77507 (TAT161)	tumor de pulmón	tejido normal de pulmón
DNA77507 (TAT161)	tumor de riñón	tejido normal de riñón
DNA77507 (TAT161)	tumor de hígado	tejido normal de hígado
DNA77507 (TAT161)	tumor ovárico	tejido ovárico normal
DNA77507 (TAT161)	tumor pancreático	tejido pancreático normal
DNA77507 (TAT161)	tumor de recto	tejido normal de recto
DNA77507 (TAT161)	tumor de piel	tejido normal de piel
DNA77507 (TAT161)	tumor uterino	tejido uterino normal
DNA77507 (TAT161)	tumor cerebral	tejido cerebral normal
DNA77507 (TAT161)	tumor de tejidos blandos	tejido blando normal
DNA77507 (TAT161)	tumor de hueso	tejido normal de hueso
DNA82343 (TAT157)	tumor de colon	tejido normal de colon
DNA82343 (TAT157)	tumor ovárico	tejido ovárico normal
DNA82343 (TAT157)	tumor de estómago	tejido normal de estómago
DNA82343 (TAT157)	tumor de timo	tejido normal de timo
DNA82343 (TAT157)	tumor de intestino delgado	tejido normal de intestino delgado
DNA87994 (TAT160)	tumor de mama	tejido normal de mama
DNA87994 (TAT160)	tumor pancreático	tejido pancreático normal
DNA87994 (TAT160)	tumor de colon	tejido normal de colon
DNA87994 (TAT160)	tumor de esófago	tejido normal de esófago
DNA87994 (TAT160)	tumor ovárico	tejido ovárico normal
DNA87994 (TAT160)	tumor de próstata	tejido normal de próstata
DNA88131 (TAT158)	tumor de mama	tejido normal de mama
DNA88131 (TAT158)	tumor de colon	tejido normal de colon
DNA88131 (TAT158)	tumor de pulmón	tejido normal de pulmón

Molécula	regulación por incremento en la expresión en:	cuando se compara con:
DNA88131 (TAT158)	tumor pancreático	tejido pancreático normal
DNA88131 (TAT158)	tumor de próstata	tejido normal de próstata

DNA88131 (TAT158)	tumor de estómago	tejido normal de estómago
DNA88131 (TAT158)	tumor de vejiga	tejido normal de vejiga
DNA88131 (TAT158)	tumor cerebral	tejido cerebral normal
DNA95930 (TAT110)	tumor de colon	tejido normal de colon
DNA95930 (TAT110)	tumor de pulmón	tejido normal de pulmón
DNA95930 (TAT110)	tumor de próstata	tejido normal de próstata
DNA95930 (TAT110)	tumor de endometrio	tejido normal de endometrio
DNA95930 (TAT110)	tumor ovárico	tejido ovárico normal
DNA95930 (TAT110)	tumor de mama	tejido normal de mama
DNA95930-1 (TAT210)	tumor de colon	tejido normal de colon
DNA95930-1 (TAT210)	tumor de pulmón	tejido normal de pulmón
DNA95930-1 (TAT210)	tumor de próstata	tejido normal de próstata
DNA95930-1 (TAT210)	tumor de endometrio	tejido normal de endometrio
DNA95930-1 (TAT210)	tumor ovárico	tejido ovárico normal
DNA95930-1 (TAT210)	tumor de mama	tejido normal de mama
DNA96917 (TAT159)	tumor pancreático	tejido pancreático normal
DNA96917 (TAT159)	tumor de pulmón	tejido normal de pulmón
DNA96917 (TAT159)	tumor de hígado	tejido normal de hígado
DNA96917 (TAT159)	tumor de próstata	tejido normal de próstata
DNA96930 (TAT112)	tumor de mama	tejido normal de mama
DNA96930 (TAT112)	tumor de colon	tejido normal de colon
DNA96930 (TAT112)	tumor de pulmón	tejido normal de pulmón
DNA96930 (TAT112)	tumor ovárico	tejido ovárico normal
DNA96930 (TAT112)	tumor pancreático	tejido pancreático normal
DNA96930 (TAT112)	tumor de estómago	tejido normal de estómago
DNA96936 (TAT147)	tumor de mama	tejido normal de mama
DNA96936 (TAT147)	tumor de colon	tejido normal de colon
DNA96936 (TAT147)	tumor de próstata	tejido normal de próstata
DNA96936 (TAT147)	tumor uterino	tejido uterino normal
DNA98565 (TAT145)	tumor cerebral	tejido cerebral normal
DNA98565 (TAT145)	tumor de colon	tejido normal de colon
DNA246435 (TAT152)	tumor cerebral	tejido cerebral normal
DNA246435 (TAT152)	tumor de colon	tejido normal de colon
DNA98591 (TAT162)	tumor de colon	tejido normal de colon
DNA98591 (TAT162)	tumor de intestino delgado	tejido normal de intestino delgado

DNA98591 (TAT162)	tumor ovárico	tejido ovárico normal
DNA98591 (TAT162)	tumor de esófago	tejido normal de esófago
DNA108809 (TAT114)	tumor de colon	tejido normal de colon
DNA108809 (TAT114)	tumor de pulmón	tejido normal de pulmón
DNA108809 (TAT114)	tumor ovárico	tejido ovárico normal
DNA108809 (TAT114)	tumor cerebral	tejido cerebral normal
DNA143493 (TAT103)	tumor de mama	tejido normal de mama
DNA167234 (TAT130)	tumor de próstata	tejido normal de próstata
DNA235621 (TAT166)	tumor de próstata	tejido normal de próstata
DNA176766 (TAT132)	tumor de riñón	tejido normal de riñón
DNA176766 (TAT132)	tumor uterino	tejido uterino normal
DNA236463 (TAT150)	tumor de riñón	tejido normal de riñón
DNA236463 (TAT150)	tumor uterino	tejido uterino normal
DNA181162 (TAT129)	tumor de próstata	tejido normal de próstata
DNA188221 (TAT111)	tumor de colon	tejido normal de colon
DNA188221 (TAT111)	tumor de hígado	tejido normal de hígado
DNA188221 (TAT111)	tumor de pulmón	tejido normal de pulmón
DNA233876 (TAT146)	tumor de colon	tejido normal de colon

Molécula	**regulación por incremento en la expresión en:**	**cuando se compara con:**
DNA233876 (TAT146)	tumor de hígado	tejido normal de hígado
DNA233876 (TAT146)	tumor de pulmón	tejido normal de pulmón
DNA193891 (TAT148)	tumor de próstata	tejido normal de próstata
DNA193891 (TAT148)	tumor de mama	tejido normal de mama
DNA248170 (TAT187)	tumor de mama	tejido normal de mama
DNA248170 (TAT187)	tumor de próstata	tejido normal de próstata
DNA194628 (TAT118)	tumor de riñón	tejido normal de riñón
DNA246415 (TAT167)	tumor de riñón	tejido normal de riñón
DNA215609 (TAT113)	tumor de colon	tejido normal de colon
DNA220432 (TAT128)	tumor de próstata	tejido normal de próstata
DNA226094 (TAT164)	tumor de mama	tejido normal de mama
DNA226094 (TAT164)	tumor cerebral	tejido cerebral normal
DNA226094 (TAT164)	tumor ovárico	tejido ovárico normal
DNA226094 (TAT164)	tumor de pulmón	tejido normal de pulmón
DNA226165 (TAT122)	tumor de mama	tejido normal de mama

DNA226165 (TAT122)	tumor de endometrio	tejido normal de endometrio
DNA226165 (TAT122)	tumor de pulmón	tejido normal de pulmón
DNA226165 (TAT122)	tumor de colon	tejido normal de colon
DNA226237 (TAT117)	tumor de riñón	tejido normal de riñón
DNA246450 (TAT168)	tumor de riñón	tejido normal de riñón
DNA246450 (TAT168)	tumor cerebral	tejido cerebral normal
DNA226456 (TAT144)	tumor de mama	tejido normal de mama
DNA226456 (TAT144)	tumor cerebral	tejido cerebral normal
DNA226456 (TAT144)	tumor de endometrio	tejido normal de endometrio
DNA226456 (TAT144)	tumor de riñón	tejido normal de riñón
DNA226456 (TAT144)	tumor de pulmón	tejido normal de pulmón
DNA237637 (TAT188)	tumor de mama	tejido normal de mama
DNA237637 (TAT188)	tumor cerebral	tejido cerebral normal
DNA237637 (TAT188)	tumor de endometrio	tejido normal de endometrio
DNA237637 (TAT188)	tumor de riñón	tejido normal de riñón
DNA237637 (TAT188)	tumor de pulmón	tejido normal de pulmón
DNA226539 (TAT126)	tumor de colon	tejido normal de colon
DNA226539 (TAT126)	tumor de endometrio	tejido normal de endometrio
DNA226539 (TAT126)	tumor ovárico	tejido ovárico normal
DNA226539 (TAT126)	tumor pancreático	tejido pancreático normal
DNA236511 (TAT151)	tumor de colon	tejido normal de colon
DNA236511 (TAT151)	tumor de endometrio	tejido normal de endometrio
DNA236511 (TAT151)	tumor ovárico	tejido ovárico normal
DNA236511 (TAT151)	tumor pancreático	tejido pancreático normal
DNA226771 (TAT115)	tumor de colon	tejido normal de colon
DNA227087 (TAT163)	tumor de mama	tejido normal de mama
DNA227087 (TAT163)	tumor de colon	tejido normal de colon
DNA227087 (TAT163)	tumor endocrino	tejido endocrino normal
DNA227087 (TAT163)	tumor de riñón	tejido normal de riñón
DNA227087 (TAT163)	tumor de hígado	tejido normal de hígado
DNA227087 (TAT163)	tumor de pulmón	tejido normal de pulmón
DNA227087 (TAT163)	tumor pancreático	tejido pancreático normal
DNA227087 (TAT163)	tumor uterino	tejido uterino normal
DNA227087 (TAT163)	tumor de próstata	tejido normal de próstata
DNA227087 (TAT163)	tumor de vejiga	tejido normal de vejiga
DNA266307 (TAT227)	tumor de mama	tejido normal de mama

DNA266307 (TAT227)	tumor de colon	tejido normal de colon
DNA266307 (TAT227)	tumor endocrino	tejido endocrino normal
DNA266307 (TAT227)	tumor de riñón	tejido normal de riñón
DNA266307 (TAT227)	tumor de hígado	tejido normal de hígado

Molécula	**regulación por incremento en la expresión en:**	**cuando se compara con:**
DNA266307 (TAT227)	tumor de pulmón	tejido normal de pulmón
DNA266307 (TAT227)	tumor pancreático	tejido pancreático normal
DNA266307 (TAT227)	tumor uterino	tejido uterino normal
DNA266307 (TAT227)	tumor de próstata	tejido normal de próstata
DNA266307 (TAT227)	tumor de vejiga	tejido normal de vejiga
DNA266311 (TAT228)	tumor de mama	tejido normal de mama
DNA266311 (TAT228)	tumor de colon	tejido normal de colon
DNA266311 (TAT228)	tumor endocrino	tejido endocrino normal
DNA266311 (TAT228)	tumor de riñón	tejido normal de riñón
DNA266311 (TAT228)	tumor de hígado	tejido normal de hígado
DNA266311 (TAT228)	tumor de pulmón	tejido normal de pulmón
DNA266311 (TAT228)	tumor pancreático	tejido pancreático normal
DNA266311 (TAT228)	tumor uterino	tejido uterino normal
DNA266311 (TAT228)	tumor de próstata	tejido normal de próstata
DNA266311 (TAT228)	tumor de vejiga	tejido normal de vejiga
DNA266312 (TAT229)	tumor de mama	tejido normal de mama
DNA266312 (TAT229)	tumor de colon	tejido normal de colon
DNA266312 (TAT229)	tumor endocrino	tejido endocrino normal
DNA266312 (TAT229)	tumor de riñón	tejido normal de riñón
DNA266312 (TAT229)	tumor de hígado	tejido normal de hígado
DNA266312 (TAT229)	tumor de pulmón	tejido normal de pulmón
DNA266312 (TAT229)	tumor pancreático	tejido pancreático normal
DNA266312 (TAT229)	tumor uterino	tejido uterino normal
DNA266312 (TAT229)	tumor de próstata	tejido normal de próstata
DNA266312 (TAT229)	tumor de vejiga	tejido normal de vejiga
DNA266313 (TAT230)	tumor de mama	tejido normal de mama
DNA266313 (TAT230)	tumor de colon	tejido normal de colon
DNA266313 (TAT230)	tumor endocrino	tejido endocrino normal
DNA266313 (TAT230)	tumor de riñón	tejido normal de riñón

DNA266313 (TAT230)	tumor de hígado	tejido normal de hígado
DNA266313 (TAT230)	tumor de pulmón	tejido normal de pulmón
DNA266313 (TAT230)	tumor pancreático	tejido pancreático normal
DNA266313 (TAT230)	tumor uterino	tejido uterino normal
DNA266313 (TAT230)	tumor de próstata	tejido normal de próstata
DNA266313 (TAT230)	tumor de vejiga	tejido normal de vejiga
DNA227224 (TAT121)	tumor de mama	tejido normal de mama
DNA227224 (TAT121)	tumor de endometrio	tejido normal de endometrio
DNA227224 (TAT121)	tumor de pulmón	tejido normal de pulmón
DNA227224 (TAT121)	tumor de piel	tejido normal de piel
DNA247486 (TAT183)	tumor de mama	tejido normal de mama
DNA247486 (TAT183)	tumor de endometrio	tejido normal de endometrio
DNA247486 (TAT183)	tumor de pulmón	tejido normal de pulmón
DNA247486 (TAT183)	tumor de piel	tejido normal de piel
DNA227578 (TAT165)	tumor cerebral	tejido cerebral normal
DNA227800 (TAT131)	tumor de próstata	tejido normal de próstata
DNA227800 (TAT131)	tumor de riñón	tejido normal de riñón
DNA227904 (TAT140)	tumor de mama	tejido normal de mama
DNA228199 (TAT127)	tumor uterino	tejido uterino normal
DNA228199 (TAT127)	tumor de trompa de Falopio	tejido normal de trompa de Falopio
DNA228199 (TAT127)	tumor ovárico	tejido ovárico normal
DNA228199 (TAT127)	tumor de pulmón	tejido normal de pulmón
DNA228201 (TAT116)	tumor de colon	tejido normal de colon
DNA247488 (TAT189)	tumor de colon	tejido normal de colon
DNA236538 (TAT190)	tumor de colon	tejido normal de colon
DNA247489 (TAT191)	tumor de colon	tejido normal de colon

Molécula	**regulación por incremento en la expresión en:**	**cuando se compara con:**
DNA231312 (TAT143)	tumor de colon	tejido normal de colon
DNA231542 (TAT100)	tumor cerebral	tejido cerebral normal
DNA231542 (TAT100)	glioma	tejido glial normal
DNA231542-1 (TAT284)	tumor cerebral	tejido cerebral normal
DNA231542-1 (TAT284)	glioma	tejido glial normal
DNA231542-2 (TAT285)	tumor cerebral	tejido cerebral normal

DNA231542-2 (TAT285)	glioma	tejido glial normal
DNA297393 (TAT285-1)	tumor cerebral	tejido cerebral normal
DNA297393 (TAT285-1)	glioma	tejido glial normal
DNA232754 (TAT125)	tumor de pulmón	tejido normal de pulmón
DNA236246 (TAT153)	tumor de mama	tejido normal de mama
DNA236343 (TAT104)	tumor de mama	tejido normal de mama
DNA236493 (TAT141)	tumor de mama	tejido normal de mama
DNA236493 (TAT141)	tumor de glioblastoma	tejido glial normal
DNA236534 (TAT102)	tumor de mama	tejido normal de mama
DNA236534 (TAT102)	tumor de pulmón	tejido normal de pulmón
DNA236534 (TAT102)	tumor pancreático	tejido pancreático normal
DNA236534 (TAT102)	tumor de próstata	tejido normal de próstata
DNA236534 (TAT102)	tumor de vejiga	tejido normal de vejiga
DNA247480 (TAT142)	tumor de pulmón	tejido normal de pulmón
DNA264454 (TAT106)	tumor de mama	tejido normal de mama
DNA264454 (TAT106)	tumor de próstata	tejido normal de próstata
DNA264454 (TAT106)	tumor ovárico	tejido ovárico normal

EJEMPLO 7: Utilización de TAT como una Sonda de Hibridación

El siguiente método describe el uso de una secuencia nucleótida codificando TAT como una sonda de hibridación para, por ejemplo, el diagnóstico de la presencia de un tumor en un mamífero.

También se puede emplear el ADN comprendiendo la secuencia de codificación de un TAT de longitud total o maduro como se divulga en la presente, como una sonda para el cribado en busca de ADNs homólogos (como aquellos codificando los variantes naturales del TAT) en bibliotecas de ADN de tejido humano o en bibliotecas genómicas de tejido humano.

La hibridación y el lavado de filtros conteniendo cualquiera de las bibliotecas de ADN se realiza bajo las siguientes condiciones de alta rigurosidad. La hibridación de sonda radiomarcada derivada de TAT a los filtros se realiza en una solución de formamida al 50%, 5x SSC, SDS al 0.1%, pirofosfato sódico al 0.1%, 50 mM de fosfato sódico, pH 6.8, 2x solución de Denhardt y sulfato de dextrano al 10% a 42 °C durante 20

DNA231542-2 (TAT285)	glioma	tejido glial normal
DNA297393 (TAT285-1)	tumor cerebral	tejido cerebral normal
DNA297393 (TAT285-1)	glioma	tejido glial normal
DNA232754 (TAT125)	tumor de pulmón	tejido normal de pulmón
DNA236246 (TAT153)	tumor de mama	tejido normal de mama
DNA236343 (TAT104)	tumor de mama	tejido normal de mama
DNA236493 (TAT141)	tumor de mama	tejido normal de mama
DNA236493 (TAT141)	tumor de glioblastoma	tejido glial normal
DNA236534 (TAT102)	tumor de mama	tejido normal de mama
DNA236534 (TAT102)	tumor de pulmón	tejido normal de pulmón
DNA236534 (TAT102)	tumor pancreático	tejido pancreático normal
DNA236534 (TAT102)	tumor de próstata	tejido normal de próstata
DNA236534 (TAT102)	tumor de vejiga	tejido normal de vejiga
DNA247480 (TAT142)	tumor de pulmón	tejido normal de pulmón
DNA264454 (TAT106)	tumor de mama	tejido normal de mama
DNA264454 (TAT106)	tumor de próstata	tejido normal de próstata
DNA264454 (TAT106)	tumor ovárico	tejido ovárico normal

EJEMPLO 7: Utilización de TAT como una Sonda de Hibridación

El siguiente método describe el uso de una secuencia nucleótida codificando TAT como una sonda de hibridación para, por ejemplo, el diagnóstico de la presencia de un tumor en un mamífero.

También se puede emplear el ADN comprendiendo la secuencia de codificación de un TAT de longitud total o maduro como se divulga en la presente, como una sonda para el cribado en busca de ADNs homólogos (como aquellos codificando los variantes naturales del TAT) en bibliotecas de ADN de tejido humano o en bibliotecas genómicas de tejido humano.

La hibridación y el lavado de filtros conteniendo cualquiera de las bibliotecas de ADN se realiza bajo las siguientes condiciones de alta rigurosidad. La hibridación de sonda radiomarcada derivada de TAT a los filtros se realiza en una solución de formamida al 50%, 5x SSC, SDS al 0.1%, pirofosfato sódico al 0.1%, 50 mM de fosfato sódico, pH 6.8, 2x solución de Denhardt y sulfato de dextrano al 10% a 42 °C durante 20

horas. El lavado de los filtros se lleva a cabo en una solución acuosa de 0.1x SSC y SDS al 0.1% a 42°C.

Luego se pueden identificar los ADNs que tienen una identidad de secuencia deseada con el ADN codificando el TAT de secuencia natural total utilizando técnicas estándar conocidas en el arte.

EJEMPLO 8: Expresión de TAT en *E. coli*

Este ejemplo ilustra la preparación de una forma no glicosilada de TAT mediante expresión recombinante en *E. coli*.

La secuencia de ADN codificando el TAT inicialmente se amplifica empleando unos iniciadores de la PCR seleccionados. Los iniciadores deben contener unos sitios de enzimas de restricción que corresponden a los sitios de restricción en el vector de expresión seleccionado. Se puede emplear una variedad de vectores de expresión. Un ejemplo de un vector apropiado es pBR322 (derivado de *E. coli*; consultar Bolivar y colaboradores, Gene, 2:95 (1977)) que contiene genes para la resistencia a ampicilina y tetraciclina. El vector se digiere mediante la enzima de restricción y se desfosforila. Entonces las secuencias amplificadas de la PCR se ligan al vector. El vector preferiblemente va a incluir secuencias que codifican un gen de resistencia a los antibióticos, un promotor trp, un líder polihis (incluyendo los seis primeros codones STII, la secuencia polihis y el sitio de segmentación de la enterocinasa), la región de codificación del TAT, un terminador de transcripción lambda y un gen argU.

Luego la mezcla de ligado se utiliza para transformar la cepa *E. coli* seleccionada utilizando los métodos descritos en Sambrook y colaboradores, supra. Se identifican los transformantes por su capacidad de crecer en placas LB y luego se seleccionan las colonias resistentes a los antibióticos. Se puede aislar ADN plásmido y se puede confirmar mediante análisis de restricción y secuenciación de ADN.

Los clones seleccionados se pueden cultivar durante la noche en un medio de cultivo líquido como caldo LB complementado con antibióticos. Posteriormente el cultivo que se dejó durante la noche se puede utilizar para inocular un cultivo de mayor escala. Posteriormente las células se cultivan hasta alcanzar una densidad óptica deseada durante la cual se activa el promotor de expresión.

Después de cultivar las células durante varias horas más, se pueden recolectar las células mediante centrifugado. Las células aglomeradas obtenidas mediante centrifugado se pueden solubilizar empleando diversos agentes conocidos en el arte y luego la proteína TAT solubilizada se puede purificar empleando una columna de metal quelado bajo condiciones que permiten una fuerte unión de la proteína.

El TAT se puede expresar en *E. coli* en una forma de marcado poli-His utilizando el siguiente procedimiento. El ADN codificando TAT inicialmente se amplifica utilizando unos iniciadores de la PCR seleccionados. Los iniciadores van a contener unos sitios de enzimas de restricción que corresponden a los sitios de enzimas de restricción en el vector de expresión seleccionado y otras secuencias útiles que contribuyen a una iniciación de traslación eficiente y confiable, una purificación rápida en una columna de quelado de metal y una extracción proteolítica con enterocinasa. Luego la PCR amplificada, las secuencias marcadas poli-His se ligan en un vector de expresión que se utiliza para transformar un hospedero *E. coli* con base en la cepa 52 (W3110 fuhA(tonA) Ion galE rpoHts(htpRts) clpP(lacIq). Primero se cultivan los transformantes en LB conteniendo 50 mg/ml de carbenicilina a 30°C con agitación hasta alcanzar una O.D.600 de 3-5. Posteriormente se diluyen los cultivos 50-100 veces en el medio CRAP (que se prepara mezclando 3.57 g $(NH_4)_2SO_4$, 0.71 g citrato de sodio·2H2O, 1.07 g KCl, 5.36 g Difco extracto de levadura, 5.36 g Sheffield hycase SF en 500 mL de agua, al igual que 110 mM de MPOS, pH 7.3, 0.55% (p/v) glucosa y 7 mM $MgSO_4$) y se cultiva durante aproximadamente 20-30 horas a 30°C con agitación. Las muestras se

retiran para verificar la expresión mediante análisis SDS-PAGE y el volumen del cultivo se centrifuga para aglomerar las células. Las células aglomeradas se congelan hasta la purificación y el replegamiento.

Se vuelve a suspender la pasta *E. coli* de 0.5 a 1 L fermentaciones (pellets de 6-10 g) en 10 volúmenes (p/v) en 7 M guanidina, 20 mM Tris, pH 8 solución amortiguadora. Se agrega sulfito de sodio sólido y tetrationato de sodio para alcanzar las concentraciones finales de 0.1M y 0.02 M, respectivamente y la solución se revuelve durante la noche a 4°C. Este paso produce una proteína desnaturalizada donde todos los residuos de cisteína están bloqueados por sulfitolización. La solución se centrifuga a 40,000 rpm en una Beckman Ultracentifuge durante 30 min. El sobrenadante se diluye con 3-5 volúmenes de solución amortiguadora de columna de metal quelado (6 M guanidina, 20 mM Tris, pH 7.4) y se filtra a través de filtros de 0.22 micrones para su clarificación. El extracto clarificado se carga en una columna de metal quelado de 5 ml Qiagen Ni-NTA equilibrado en la solución amortiguadora de la columna de metal quelado. La solución se lava con solución amortiguadora adicional conteniendo 50 mM de imidazole (Calbiochem, grado Utrol), pH 7.4. La proteína se leviga con solución amortiguadora conteniendo 250 mM de imidazole. Las fracciones que contienen la proteína deseada se combinan y se almacenan a 4°C. La concentración de la proteína se estima mediante su absorción a 280 nm utilizando el coeficiente de extinción basado en su secuencia de aminoácido.

Las proteínas se repliegan al diluir la muestra lentamente en una solución amortiguadora de replegamiento recién preparada que consiste en: 20 mM de Tris, pH 8.6, 0.3 M de NaCl, 2.5 M de urea, 5 mM de cisteína, 20 mM de glicina y 1 mM de EDTA. Los volúmenes replegamiento se escogen de manera que la concentración final de la proteína está entre 50 y 100 microgramos/ml. La solución de replegamiento se revuelve suavemente a 4°C durante 12-36 horas. La reacción de replegamiento se extingue mediante la adición de TFA para una concentración final de 0.4%

(pH de aproximadamente 3). Antes de una purificación posterior de la proteína, la solución se filtra a través de un filtro de 0.22 micrones y se añade acetonitrilo a 2- 10% de la concentración final. La proteína replegada se somete a cromatografía en una columna de fase inversa Poros R1/H empleando una solución amortiguadora móvil de TFA al 0.1 % con levigado con un gradiente de acetonitrilo de 10 a 80%. Se analizan alícuotas de fracciones con una absorción de A280 en geles de poliacrilamida SDS y las fracciones que contienen la proteína homogénea replegada se combinan. Generalmente las especies replegadas en forma apropiada de la mayoría de las proteínas se levigan a las concentraciones más bajas de acetonitrilo puesto que esas especies son las más compactas donde sus interiores hidrófobos están protegidos de la interacción con la resina de fase inversa. Normalmente las especies agregadas se levigan a unas concentraciones más altas de acetonitrilo. Además de descomponer las formas de proteínas que se replegaron en forma diferente a la forma deseada, el paso de fase inversa también extrae las endotoxinas de las muestras.

Las fracciones que contienen el polipéptido TAT plegado deseado se combinan y se extrae el acetonitrilo utilizando un flujo suave de nitrógeno dirigido a la solución. Las proteínas se formulan en 20 mM Hepes, pH 6.8 con 0.14 M de cloruro de sodio y mannitol al 4% mediante diálisis o por filtración con gel utilizando resinas G25 Superfine (Pharmacia) equilibradas en la formulación de la solución amortiguadora y se filtran en forma estéril.

Algunos de los polipéptidos TAT divulgados en la presente se han expresado en forma exitosa y se han purificado utilizando esta(s) técnica(s).

EJEMPLO 9: Expresión de TAT en Células de Mamíferos

Este ejemplo ilustra la preparación de una forma potencialmente glicosilada de TAT mediante la expresión recombinante en células de mamíferos.

Se utiliza el vector, pRK5 (consultar EP 307,247, publicada el 15 de marzo de 1989), como el vector de expresión. Opcionalmente, El ADN del TAT se liga al pRK5 con unas enzimas de restricción seleccionadas para permitir la inserción del ADN del TAT empleando métodos de ligado como se describe en Sambrook y colaboradores, supra. El vector resultante se llama pRK5- TAT.

En una incorporación, las células hospederas seleccionadas pueden ser 293 células. Las células humanas 293 (ATCC CCL 1573) se cultivan hasta confluencia en placas de cultivo de tejido en un medio como DMEM complementado con suero fetal bovino y opcionalmente, con componentes nutrientes y/o antibióticos. Se mezclan alrededor de 10µg pRK5- ADN del TAT con alrededor de 1 µg de ADN codificando el gen VA ARN [Thimmappaya y colaboradores, Cell, 31:543 (1982)] y se disuelven en 500 µl de 1 mM de Tris-HCl, 0.1 mM de EDTA, 0.227 M de $CaCl_2$. A esta mezcla se agrega, en forma de goteo, 50 µl de 50 mM de HEPES (pH 7.35), 280 mM de NaCl, 1.5 mM de $NaPO_4$, y se permite la formación de un precipitado durante 10 minutos a 25°C. El precipitado se suspende y se agrega a las 293 células y se deja precipitar durante unas cuatro horas a 37°C. El medio de cultivo se extrae mediante aspirado y se agregan 2 ml de glicerol en PBS al 20% durante 30 segundos. Luego las 293 células se lavan con un medio libre de suero, se añade medio fresco y las células se incuban durante unos 5 días.

Aproximadamente 24 horas después de las transinfecciones, se extrae el medio de cultivo y se reemplaza por un medio de cultivo (solo) o con un medio de cultivo que contiene 200 µCi/ml ^{35}S-cisteína y 200 µCi/ml ^{35}S-metionina. Después de una incubación de 12 horas, se recolecta el medio acondicionado, se concentra en un filtro giratorio y se carga en un gel de SDS al 15%. El gel procesado se puede secar y exponer a una película durante un periodo de tiempo seleccionado para revelar la presencia del polipéptido TAT. Los cultivos que contienen las células transinfectadas se

pueden someter a una incubación posterior (en un medio libre de suero) y el medio se prueba en unos bioanálisis seleccionados.

En una técnica alterna, se puede introducir el TAT en 293 células de forma transiente utilizando el método de sulfato de dextrano descrito por Somparyrac y colaboradores, Proc. Natl. Acad. Sci., 12:7575 (1981). Se cultivan 293 células hasta su máxima densidad en un matraz con agitación centrífuga y se añaden 700 µg de pRK5- ADN del TAT. Primero las células se concentran del matraz con agitación centrífuga mediante centrifugado y se lavan con PBS. El precipitado ADN-dextrano se incuba en el pellet de la célula durante cuatro horas. Las células se trataron con glicerol al 20% durante 90 segundos, se lavaron con medio de cultivo para tejidos y se volvieron a introducir en el matraz con agitación centrífuga conteniendo medio de cultivo de tejidos, 5 µg/ml de insulina bovina y 0.1 µg/ml de transferina bovina. Después de unos cuatro días, el medio acondicionado se sometió a centrifugado y se filtró para extraer las células y los desechos. Luego la muestra que contiene el TAT expresado se puede concentrar y purificar mediante cualquier método que se seleccione, como diálisis y/o cromatografía en columna.

En otra incorporación, el TAT se puede expresar en células CHO. El pRK5-TAT se puede transinfectar en las células CHO utilizando unos reactivos muy conocidos como $CaPO_4$ o DEAE-dextrano. Como se describe arriba, los cultivos de células se pueden incubar, y el medio se puede reemplazar con un medio de cultivo (solo) o con un medio que contiene un radiomarcado como ^{35}S-metionina. Después de determinar la presencia del polipéptido TAT, el medio de cultivo se puede reemplazar con un medio libre de suero. Preferiblemente, los cultivos se incuban durante unos 6 días, y luego se cosecha el medio acondicionado. Después el medio que contiene el TAT expresado se puede concentrar y purificar por cualquier método seleccionado.

El TAT marcador de epítopo también se puede expresar en células hospederas CHO. El TAT se puede subclonar del vector pRK5. El inserto

del subclon se puede someter a una PCR para que se fusione en el marco con un marcador de epítopo seleccionado como un marcador poli-his en un vector de expresión Baculovirus. El inserto TAT marcado poli-his se puede entonces subclonar en un vector estimulado SV40 que contiene un marcador de selección como un DHFR para seleccionar los clones estables. Finalmente, las células CHO se pueden transinfectar (como se describe arriba) con el vector estimulado SV40. Se puede realizar la marcación como se describe arriba, para verificar la expresión. El medio de cultivo que contiene el TAT marcado poli-His expresado se puede concentrar y purificar mediante cualquier método seleccionado, como mediante la cromatografía Ni^{2+}-afinidad de quelado.

El TAT también se puede expresar en células CHO y/o células COS mediante un procedimiento de expresión transiente o en células CHO por medio de otro procedimiento de expresión estable.

La expresión estable en las células CHO se realiza utilizando el siguiente procedimiento. Las proteínas se expresan como una construcción IgG (inmunoadhesina), donde las secuencias de codificación para las formas solubles (por ejemplo dominios extracelulares) de las respectivas proteínas se fusionan a una secuencia de región constante IgG1 que contiene la bisagra, CH2 y los dominios CH2 y/o es una forma marcada poli-His.

Después de la amplificación por PCR, los respectivos ADNs se subclonan en un vector de expresión CHO utilizando técnicas estándar como se describe en Ausubel y colaboradores, Current Protocols of Molecular Biology, Unidad 3.16, John Wiley and Sons (1997). Los vectores de expresión CHO se construyen de manera que tengan unos sitios de restricción compatibles 5' y 3' del ADN de interés para permitir una replicación múltiple conveniente de los ADNc. El vector de expresión empleado en las células CHO es como se describe en Lucas y colaboradores, Nucl. Acids Res. 24:9 (1774-1779 (1996), y utiliza el promotor/intensificador temprano SV40 para estimular la expresión del

ADNc de interés y la reductasa dihidrofolato (DHFR). La expresión DHFR permite la selección para una conservación estable del plásmido después de la transinfección.

Se introducen doce microgramos del ADN plásmido deseado en aproximadamente 10 millones de células CHO empleando unos reactivos de transinfección disponibles en el comercio Superfect® (Quiagen), Dosper® o Fugene® (Boehringer Mannheim). Las células se cultivan como se describe en Lucas y colaboradores supra. Se congelan aproximadamente 3 x 107 células en una ampolleta para futuros cultivos y producción como se describe abajo.

Las ampolletas que contienen el ADN plásmido se descongelan poniéndolas en un baño de agua con giramiento en espiral. Los contenidos se extraen mediante pipeta y se colocan en un tubo centrífugo que contiene 10 mLs de medio y se centrifugan a 1000 rpm durante 5 minutos. Se aspira el sobrenadante y las células se vuelven a suspender en 10 mL de medio selectivo (0.2 μm filtrado PS20 con 0.2 μm de suero fetal bovino diafiltrado al 5%). Luego las células se distribuyen en alícuotas en un agitador centrífugo 100 mL que contiene 90 mL de medio selectivo. Después de 1-2 días, las células se transfieren a un agitador centrífugo 250 mL que se encuentra lleno con 150 mL de medio de crecimiento selectivo y se incuban a 37°C. Después de otros 2-3 días, se siembran 3 x 10^5 células/mL en agitadores centrífugos 250 mL, 500 mL y 2000 mL. El medio de las células se cambia por un medio fresco mediante centrifugado y se vuelven a suspender en el medio de producción. Aunque se puede emplear cualquier medio CHO apropiado, en realidad se puede utilizar un medio de producción descrito en la Patente de Estados Unidos No. 5,122,469, otorgada el 16 de junio de 1992. Se siembran 1.2 x 10^6 células/mL en un agitador centrífugo de producción 3L. El día 0, se determina el número del pH de las células. El día 1, se toma una muestra del agitador centrífugo y se comienza el rociado con aire filtrado. El día 2, se toma una muestra del agitador centrífugo, la temperatura se cambia a

arriba para las proteínas poli-His marcadas. Se evalúa la homogeneidad mediante geles poliacrilamidas SDS y mediante la secuenciación de terminal N de aminoácidos mediante degradación de Edman.

Algunos de los polipéptidos TAT divulgados en la presente se han expresado en forma exitosa y se han purificado utilizando esta(s) técnica(s).

EJEMPLO 10: Expresión del TAT en Levadura

El siguiente método describe la expresión recombinante de TAT en levadura.

Primero se construyen unos vectores de expresión de levadura para la producción o la secreción intracelular de TAT del promotor ADH2/GAPDH. Se inserta el ADN codificando el TAT y el promotor en unos sitios de enzimas de restricción apropiados en el plásmido seleccionado para dirigir la expresión intracelular del TAT. Para la secreción, se puede clonar el ADN codificando el TAT en el plásmido seleccionado, junto con el ADN codificando el promotor ADH2/GAPDH, una señal peptídica TAT natural u otra señal peptídica, o por ejemplo, un factor alfa de levadura o una secuencia de señal/líder de invertasa o secretora y secuencias de enlazamiento (si se requieren) para la expresión del TAT.

Luego las células de levadura, como la cepa de levadura AB110, se pueden transformar con los plásmidos de expresión descritos arriba y se pueden cultivar en un medio de fermentación seleccionado. Los sobrenadantes de la levadura transformada se pueden analizar mediante precipitación con ácido tricloroacético al 10% y se pueden separar mediante SDS-PAGE, seguido por la coloración de los geles con mancha azul de Coomassie.

Posteriormente se puede aislar y purificar el TAT recombinante retirando las células de levadura del medio de fermentación mediante centrifugado y luego concentrando el medio utilizando unos filtros de cartucho seleccionados. El concentrado conteniendo el TAT se puede

purificar aún más empleando unas resinas seleccionadas por cromatografía en columna.

Algunos de los polipéptidos TAT divulgados en la presente se han expresado en forma exitosa y se han purificado utilizando esta(s) técnica(s).

EJEMPLO 11: Expresión de TAT en Células de Insectos Infectadas con Baculovirus

El siguiente método describe la expresión recombinante de TAT en células de insectos infectadas con Baculovirus.

La secuencia codificando el TAT se fusiona en sentido ascendente de un marcador de epítopo contenido en un vector de expresión de baculovirus. Dichos marcadores de epítopos incluyen poli-his marcadores y marcadores de inmunoglobulina (como regiones Fc de IgG). Se puede utilizar una variedad de plásmidos, incluyendo plásmidos derivados de plásmidos disponibles en el comercio como pVL1393 (Novagen). Brevemente, la secuencia codificando el TAT o la porción deseada de la secuencia de codificación del TAT como la secuencia codificando un dominio extracelular de una proteína transmembrana o la secuencia codificando la proteína madura, si la proteína es extracelular, es amplificada por una PCR con iniciadores complementarios a las regiones 5' y 3'. El iniciador 5' puede incorporar unos sitios de enzimas de restricción de flanco (seleccionados) Luego el producto se digiere con estas enzimas de restricción seleccionadas y se subclonan en el vector de expresión.

Se genera el baculovirus recombinante mediante la cotransinfección del plásmido de arriba y el ADN del virus BaculoGold™ (Pharmingen) en células de *Spodoptera frugiperda* ("Sf9") (ATCC CRL 1711) utilizando lipofectina (disponible en el comercio en GIBCO-BRL). Después de 4 - 5 días de incubación a 28C, los virus liberados se cosechan y se emplean para amplificaciones futuras. La infección viral y la expresión de la proteína se desarrollan como se describe por O'Reilley y colaboradores, Baculovirus

Expression Vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

Luego se pueden expresar los TAT marcados poli-his, por ejemplo, mediante una cromatografía de afinidad ^{2}-*Nichelate* como sigue. Se preparan extractos a partir de células recombinantes Sf9 infectadas con virus como lo describen Rupert y colaboradores, Nature, 362:175-179 (1993). Brevemente, se lavan las células Sf9, se vuelven a suspender en solución amortiguadora de sonicación (25 mL Hepes, pH 7.9; 12.5 mM de $MgCl_2$; 0.1 mM de EDTA; glicerol al 10%; NP-40 al 0.1%; 0.4 M de KCl), y se someten a sonicación dos veces durante 20 segundos sobre hielo. Los productos dispersados por sonicación se liberan mediante centrifugado, y el sobrenadante se diluye 50 veces en solución amortiguadora de carga (50 mM de fosfato, 300 mM de NaCl, glicerol al 10%, pH 7.8) y se filtran a través de un filtro de 0.45 μm. Se prepara una columna Ni^{2}+-NTA agarosa (disponible en el comercio en Qiagen) con un volumen de lecho de 5 mL, se lava con 25 mL de agua y se equilibra con 25 ml de solución amortiguadora de carga. El extracto de células filtradas se carga en la columna a 0.5 mL por minuto. La columna se lava hasta una línea base de A_{280} con solución amortiguadora de carga, y en ese punto se inicia la recolección de las fracciones. Luego la columna se lava con una solución amortiguadora secundaria de lavado (50 mM de fosfato; 300 mM de NaCl, glicerol al 10%, pH 6.0), que leviga la proteína que no está específicamente unida. Después de llegar nuevamente a la línea base A_{280}, la columna se desarrolla con un gradiente de 500 mM de imidazole en la solución amortiguadora secundaria. Se recolectan fracciones de 1 mL y se analizan mediante SDS-PAGE y mancha de plata o inmunotransferencia con Ni^{2}+-NTA-conjugado a fosfatasa alcalina (Qiagen). Se combinan las fracciones conteniendo el TAT His_{10}-marcado levigado y se someten a diálisis frente a una solución amortiguadora de carga.

Alternativamente, la purificación del TAT IgG marcado (o la Fc marcada) se puede realizar utilizando técnicas de cromatografía conocidas,

incluyendo por ejemplo, cromatografía en columna de Proteína A o Proteína G.

Algunos de los polipéptidos TAT divulgados en la presente se han expresado en forma exitosa y se han purificado utilizando esta(s) técnica(s).

EJEMPLO 12: La Preparación de Anticuerpos que Fijan el TAT

Este ejemplo ilustra la preparación de anticuerpos monoclonales que específicamente pueden fijar el TAT. Este ejemplo ilustra además la preparación de los anticuerpos monoclonales que específicamente pueden fijar el polipéptido TAT188(E16).

Las técnicas para producir los anticuerpos monoclonales son conocidas en el arte y se describen, por ejemplo, en Goding, supra. Los inmunógenos que se pueden utilizar incluyen TAT purificado, proteínas de fusión conteniendo el TAT y células que expresan el TAT recombinante en la superficie de la célula. Un artesano con destreza puede seleccionar el inmunógeno sin una experimentación indebida.

Ratones, como Balb/c, se inmunizan con el inmunógeno TAT emulsificado en potenciador completo de Freund y se inyecta de manera subcutánea o intraperitoneal en una cantidad de 1-100 microgramos. Alternativamente, el inmunógeno se emulsifica en MPL-TDM potenciador (Ribi Immunochemical Research, Hamilton, MT) y se inyecta en las almohadillas de las patas posteriores del animal. Luego los animales inmunizados reciben un refuerzo de 10 a 12 días más tarde con más inmunógeno emulsificado en el potenciador seleccionado. Posteriormente, durante varias semanas, también los animales pueden recibir refuerzos con inyecciones adicionales de inmunización. Periódicamente se pueden obtener muestras de suero de los ratones mediante un sangrado retroorbital para realizar pruebas de ELISA con el fin de detectar anticuerpos anti-TAT.

El polipéptido TAT188, E16, es una proteína doce-transmembrana donde los dos terminales C- y N- se encuentran a nivel intracelular.

TAT(E16) es una subunidad del transportador Na2+ del aminoácido independiente neutro grande, heterodimerizado con una cadena pesada común 4F2hc (CD98hc) para formar una unidad funcional (consultar el diagrama en la Fig. 155). Para preparar anticuerpos anti-TAT188 (anti-E16) apuntando a el(los) dominio(s) extracelular(es), se utilizaron células PC3, que en forma endógena expresan E16, como el inmunógeno. Brevemente, las células PC3 (22 x 10^6 células/ml) se inyectaron en ratones Balb-c. Los títulos de anticuerpos se analizaron contra 293 células de control y células 293-E16 (293 células transinfectadas con y produciendo E16). Se prepararon seis hibridomas que expresaron anticuerpos anti-E16. De estos, tres anticuerpos mostraron una buena especificidad para E16 mediante análisis FACS (consultar Fig. 156) donde la fijación en la superficie de la célula fue paralela a la expresión en superficie de E16. Específicamente la transinfección E16 ARNsi inhibió la expresión de E16 (consultar Ejemplo 17 en la presente) y también disminuyó la cantidad de anticuerpo anti-E16 fijado a la superficie de las células PC3. En la presente se hace referencia a estos anticuerpos como 3B5.1 (o 3B5), 12G12.1 (o 12G12) y 12B9.1 (o 12B9) y se utilizaron para experimentos futuros divulgados en la presente.

Se prepararon hibridomas expresando los anticuerpos de la invención como sigue. Después de detectar títulos apropiados de un anticuerpo, los animales "positivos" para anticuerpos se pueden inyectar con una inyección final intravenosa de TAT. Tres o cuatro días más tarde, se sacrifican los ratones y se cosechan las células del bazo. Luego las células del bazo se fusionan (utilizando polietilenglicol al 35%) a una estirpe celular murina mieloma seleccionada como P3X63 AgU.1, disponible en ATCC, No. CRL 1597. Las fusiones generan células de hibridoma que luego se pueden llevar a placas de tejido de cultivo de 96 pozos conteniendo medio HAT (hipoxantina, aminopterina y timidina) para inhibir la proliferación de células sin fusionar, híbridos de mieloma e híbridos de células de bazo.

Las células de hibridoma se cribaron en una prueba de ELISA para identificar la reactividad contra el TAT. La determinación de células de hibridoma "positivas" que secretan los anticuerpos monoclonales deseados contra el TAT está dentro del ámbito de las destrezas del arte.

Las células de hibridoma positivas se pueden inyectar de forma intraperitoneal en ratones singénicos Balb/c para producir ascitos conteniendo los anticuerpos monoclonales anti-TAT. Alternativamente, las células de hibridoma se pueden cultivar en unos matraces o frascos para el cultivo de tejidos. La purificación de los anticuerpos monoclonales producidos en los ascitos se puede lograr mediante precipitación con sulfato de amonio, seguido por una cromatografía de exclusión en gel. Alternativamente, se puede utilizar la cromatográfica de afinidad con base en la fijación de un anticuerpo a la proteína A o a la proteína G.

Los anticuerpos dirigidos hacia algunos de los polipéptidos TAT divulgados en la presente se han producido con éxito utilizando esta(s) técnica(s). Más específicamente, los anticuerpos monoclonales funcionales que son capaces de reconocer y fijarse a la proteína TAT (según se mide por un ELISA estándar, un análisis de clasificación FACS y/o un análisis de inmunohistoquímica) se han generado en forma exitosa contra las siguientes proteínas TAT, como se divulga en la presente: TAT110 (DNA95930), TAT210 (DNA95930-1), TAT113 (DNA215609), TAT126 (DNA226539), TAT151 (DNA236511), TAT111 (DNA188221), TAT146 (DNA233876), TAT112 (DNA96930), TAT145 (DNA98565), TAT152 (DNA246435), TAT141 (DNA236493), TAT114 (DNA108809), TAT104 (DNA236343), TAT100 (DNA231542), TAT284 (DNA231542-1), TAT285 (DNA231542-2), TAT285-1 (DNA297393), TAT144 (DNA226456), TAT188 (DNA237637), TAT123 (DNA210499), TAT211 (DNA219894), TAT102 (DNA236534), TAT127 (DNA228199) y TAT128 (DNA220432). Es interesante notar que los Solicitantes identificaron que los anticuerpos monoclonales preparados contra TAT111 (DNA188221) y TAT146 (DNA233876) son capaces de bloquear la activación del receptor EphB2R

codificado por las moléculas DNA188221 y DNA233876 por su polipéptido ligando asociado. Como tal, los anticuerpos y los métodos para utilizar esos anticuerpos para bloquear la activación del receptor EphB2R (por ejemplo, los polipéptidos TAT111 y TAT146) por su ligando asociado están comprendidos en la invención descrita en la presente. Por otra parte, los Solicitantes identificaron que los anticuerpos monoclonales dirigidos contra los polipéptidos TAT110 (DNA95930) y TAT210 (DNA95930-1) (por ejemplo, polipéptidos del receptor alfa IL-20) son capaces de inhibir la activación del receptor alfa IL20 mediante la proteína IL-19. Como tal, los anticuerpos y los métodos para utilizar esos anticuerpos para inhibir la activación del receptor alfa IL-20 (por ejemplo, los polipéptidos TAT110 y TAT210) mediante EL-19 se encuentran comprendidos en la invención descrita en la presente.

Además de tener éxito en la preparación de anticuerpos monoclonales dirigidos contra los polipéptidos TAT como se describe en la presente, muchos de esos anticuerpos monoclonales se han conjugado con éxito a una toxina celular para ser utilizados para dirigir la toxina celular a una célula (o tejido) que expresa un polipéptido TAT de interés (tanto *in vitro* como *in vivo*). Por ejemplo, se han generado unos anticuerpos monoclonales derivados de toxinas (por ejemplo, DM1) a los siguientes polipéptidos TAT como se describe en la presente: TAT110 (DNA95930), TAT210 (DNA95930-1), TAT112 (DNA96930), TAT113 (DNA215609), TAT111 (DNA188221) y TAT146 (DNA233876). Como se muestra abajo, otros conjugados útiles de toxinas incluyen las auristatinas MMAE y MMAF.

EJEMPLO 13: Purificación de Polipéptidos TAT Utilizando Anticuerpos Específicos

Se pueden purificar polipéptidos TAT naturales o recombinantes mediante una variedad de técnicas estándar en el arte de la purificación de las proteínas. Por ejemplo, se purifica un polipéptido pro-TAT, un polipéptido TAT maduro, o un polipéptido pre-TAT mediante

cromatografía de inmunoafinidad empleando anticuerpos específicos para el polipéptido TAT de interés. En general, se construye una columna de inmunoafinidad mediante el acoplamiento covalente del anticuerpo del polipéptido anti-TAT a una resina cromatográfica activada.

Se preparan inmunoglobulinas policlonales a partir de suero inmune bien sea mediante precipitación con sulfato de amonio o por purificación en Proteína A inmovilizada. (Pharmacia LKB Biotechnology, Piscataway, N.J.). Igualmente, se preparan anticuerpos monoclonales a partir de fluido de ascitos de ratones mediante precipitación con sulfato de amonio o cromatografía en Proteína A inmovilizada. La inmunoglobulina parcialmente purificada se adhiere de forma covalente a una resina cromatográfica como CnBr-activada SEPHAROSE™ (Pharmacia LKB Biotechnology). El anticuerpo se acopla a la resina, la resina se bloquea y la resina derivada se lava de acuerdo con las instrucciones del fabricante.

La citada columna de inmunoafinidad se utiliza en la purificación del polipéptido TAT preparando una fracción a partir de células conteniendo el polipéptido TAT en una forma soluble. Esta preparación se deriva mediante la solubilización de toda la célula o de una fracción subcelular obtenida a partir de un centrifugado diferencial mediante la adición de un detergente o por otros métodos bien conocidos en el arte. Alternativamente, se puede secretar un polipéptido TAT soluble conteniendo una secuencia de señal en una cantidad útil en el medio en el cual se cultivan las células.

Una preparación conteniendo un polipéptido TAT soluble se pasa por la columna de inmunoafinidad, y la columna se lava bajo condiciones que permiten la absorción preferencial del polipéptido TAT (*por ejemplo*, soluciones amortiguadoras de alta potencia iónica en presencia de un detergente). Luego la columna se leviga bajo condiciones que alteran la unión del anticuerpo/polipéptido TAT (*por ejemplo*, una solución amortiguadora de pH bajo, como de aproximadamente pH 2-3, o una alta concentración de un chaótropo como urea o un ión tiocianato), y se recolecta ese polipéptido TAT.

EJEMPLO 14: Análisis de Muerte de Células de Tumores *In Vitro*

Se pueden obtener las células de mamíferos que expresan el polipéptido TAT de interés mediante unas técnicas estándar de vector de expresión y de clonación. Alternativamente, muchas estirpes de células de tumor que expresan los polipéptidos TAT de interés están disponibles para el público, por ejemplo, mediante la ATCC y se pueden identificar de forma rutinaria utilizando el análisis estándar de ELISA o FACS. Los anticuerpos monoclonales de polipéptidos anti-TAT (y derivados de conjugados de toxinas de los mismos) se pueden utilizar en el análisis para determinar la capacidad del anticuerpo para matar las células expresando el polipéptido TAT *in vitro*.

Por ejemplo, se obtienen las células que expresan el polipéptido TAT de interés como se describe arriba y se llevan a placas de 96 pozos. En un análisis, se incluye el conjugado anticuerpo/toxina (o el anticuerpo desnudo) durante la incubación de la célula por un periodo de 4 días. En un segundo análisis independiente, las células se incuban durante 1 hora con el conjugado anticuerpo/toxina (o el anticuerpo desnudo) y luego se lava y se incuba en ausencia del conjugado anticuerpo/toxina por un periodo de 4 días. Posteriormente se mide la viabilidad de las células utilizando el CellTiter-Glo Luminescent Cell Viability Assay de Promega (Cat# G7571). Las células sin tratar sirven como un control negativo.

En un análisis específico, la capacidad de los anticuerpos monoclonales dirigidos contra el TAT112 (DNA96930) se analizaron para conocer la capacidad de matar las células expresando ese polipéptido. En un análisis, se preparó un vector de expresión denominado gD.NCA. Las secuencias codificando el polipéptido TAT112 insertadas en ese vector se estimulan por un promotor SV40 y el vector también contiene la señal temprana SV40 poli A. El vector gD.NCA se cotransinfectó en las células PC3 junto con un vector SV40 que expresa una neo resistencia en las células PC3 y se seleccionaron los transformantes positivos en 800 µg/ml

G418. Se aislaron los clones positivos en placas de 96 pozos y se analizaron mediante una citometría de flujo utilizando un anticuerpo monoclonal anti-TAT112 preparado como se describe arriba y denominado 3E6. Se seleccionó el clon 3 para el análisis y se encontró que expresaba un alto nivel del polipéptido TAT112 en su superficie. En un segundo análisis independiente, se obtuvo la cepa celular del cáncer pancreático, Hpaf II, de la ATCC y se utilizó en el análisis.

En otro análisis específico, se analizó la capacidad de los anticuerpos monoclonales dirigidos contra el TAT188 (DNA237637) para conocer la capacidad de matar las células que expresan ese polipéptido.

Primero se comprobó la capacidad de un anticuerpo anti-TAT188(E16) desnudo para afectar la actividad celular.

A. Los anticuerpos monoclonales anti-TAT188(E16) estimulan la internalización del TAT188(E16) y reducen la actividad de transporte de aminoácidos.

La unión de un anticuerpo anti-E16 a las células que expresan la proteína E16 produjo la internalización del anticuerpo que en forma paralela redujo el transporte de aminoácidos. Las células PC3 se incubaron con anticuerpos primarios durante el periodo de tiempo indicado con inhibidores de proteasoma (Sigma) a 37 °C. Luego se extraen los anticuerpos y las células se lavan con PBS varias veces. Las células se fijan con PFA al 4 %, luego se permeabilizan con PBS/0.1%TritonX-100. Después de bloquearlas con suero de cabra al 10% las células se colorearon con el anticuerpo secundario (conjugado cabra anti-ratón IgG-Cy3 (Jackson immunolabs)) y se analizaron mediante microscopio fluorescente. Los resultados muestran que un anticuerpo monoclonal anti-E16 se fija y se internaliza por las células PC3 que expresan el TAT188(E16) en su superficie (consultar Fig. 157). La comprobación de la internalización del anticuerpo anti-TAT188 destaca una característica útil de la interacción del anticuerpo-E16 porque la internalización de un anticuerpo fijado puede

resultar en la absorción de toxinas conjugadas a los anticuerpos anti-TAT188 (consultar secciones B y C de este ejemplo 14).

La función del TAT188(E16) como una proteína transportadora de aminoácidos se inhibió por la fijación del anticuerpo anti-TAT188(E16). El análisis funcional del transporte de leucina se efectuó como sigue. Se cultivaron células en una placa de 24 pozos y se trataron con el anticuerpo anti-E16 durante la noche. Luego las células se enjuagaron con solución de Hepes-Ringer tibia sin Na++ y se incubaron durante 10 min a 37 °C. Se cambió la solución amortiguadora por 200 µl/pozo de solución amortiguadora de análisis conteniendo aminoácido radiomarcado (L-[2, 3-^{3}H]-Arginina o L-[3, 4, 5-^{3}H(N)]-Leucina) 5 µCi/ml, 50 µl de Hepes-Ringer sin insodio). Después de incubar la mezcla radioactiva durante 30 sec a 37°C, se extrajo la solución amortiguadora mediante aspirado y las células se lavaron 4 veces con Hepes-Ringer helado, libre de sodio a 1ml/pozo. Las placas se secaron y a estas se agregaron 0.2ml/pozo de 0.2%SDS/0.2N NaOH. Se neutralizaron alícuotas de 0.1ml de los lisados de cada muestra con 0.1ml de 0.2N HCl, y luego se agregó un coctel de centelleo. La radioactividad ([^{3}H]) de los lisados se sometió a prueba utilizando un contador de centelleo. Los resultados de la Figura 158 también comprueban que la unión del anticuerpo anti-E16 a la proteína E16 en la superficie de la célula causa una reducción general del transporte del aminoácido aproximadamente después de 24 horas. La reducción se debe a la internalización concomitante del polipéptido E16 y a la pérdida del E16 disponible para la función de transporte.

B. La fijación de anticuerpos anti-TAT188(E16) de toxina-conjugada a células expresando el TAT188(E16) resulta en la apoptosis

La preparación de anti-T ATI 88-MC-vc-PAB-MMAE mediante la conjugación de anti-TAT188 y MC-vc-PAB-MMAE

Los anticuerpos ofrecen un método conveniente y altamente específico para administrar toxinas a las células donde una característica relativamente única de la célula, como la superficie de la proteína con especificidad de tejido, (o especificidad de enfermedad, o ambas) permite que la célula sea un blanco para matar mientras se evita matar las células sanas. Las citotoxinas auristatinas con acoplamiento a un anticuerpo mediante un enlazante MC-vc-PAB como se divulga en la presente arriba, son útiles para comprobar la apoptosis resultante de la internalización mediada por E16 del conjugado anticuerpo-toxina. El procedimiento se efectuó como sigue. Empleando un enlazante MC-vc-PAB, la toxina MMAE se enlazó de forma covalente a los residuos de cisteína en los anticuerpos anti-TAT. Se trata anti-TAT188, disuelto en 500mM de borato de sodio y 500 mM de cloruro de sodio con un pH 8.0 con una cantidad importante de 100mM de ditiotreitol (DTT). Después de una incubación a 37 °C durante unos 30 minutos, la solución amortiguadora se intercambia mediante levigado sobre resina Sephadex G25 y levigado con PBS con ImM DTPA. Se verifica el valor tiol/Ab al determinar la concentración reducida del anticuerpo a partir de la absorción a 280 nm de la solución y la concentración de tiol por la reacción con DTNB (Aldrich, Milwaukee, WI) y la determinación de la absorción a 412 nm. El anticuerpo reducido disuelto en PBS se enfría sobre hielo.

El reactivo del enlazante del fármaco, maleimidocaproil-valina-citrulina-PAB-monometil auristatina E (MMAE), es decir MC-vc-PAB-MMAE, disuelto en DMSO, se diluye en acetonitrilo y agua a una concentración conocida y se agrega al anticuerpo anti-TAT188(E16) frío y reducido en PBS. Después de más o menos una hora, se agrega una cantidad importante de maleimida para extinguir la reacción y taponar cualquier grupo tiol del anticuerpo sin reaccionar. La mezcla de la reacción se concentra mediante ultrafiltración por centrifugado y el anti-TAT188(E16)-MC-vc-PAB-MMAE se purificó y desalinizó mediante

levigado a través de una resina G25 m PBS, se filtró a través de filtros de 0.2 µm bajo condiciones estériles y se congeló para almacenamiento.

La preparación de anti-TAT188-MC-vc-PAB-MMAF mediante conjugación de anti-TAT188 y MC-vc-PAB-MMAF

Se preparó anti-TAT188-MC-val-cit-PAB-MMAF mediante conjugación de anti-TAT188(E16) y MC-val-cit-PAB-MMAF siguiendo el procedimiento divulgado arriba para la conjugación de MMAE. MMAF es un derivado de MMAE con una fenilalanina en el terminal C del fármaco.

Determinación de la Citotoxicidad in vitro de anti-TAT188 de toxina conjugada

Los anticuerpos de toxina conjugada 3B5, 12B9 y 12G12 de este experimento se prepararon como conjugados 3B5- MC-vc-PAB-MMAE, 12B9-MC-vc-PAB-MMAE y 3B5-MC-vc-PAB-MMAE, respectivamente. El conjugado anticuerpo-toxina, alfa-IL8-MC-vc-P AB-MMAE, se empleó como un control negativo. Se pusieron en contacto unas células PC3 (una estirpe celular de cáncer de próstata, disponible en ATCC y que expresaba E16 de forma endógena en su superficie) y células Colo205 (una estirpe celular de cáncer de colon humano, disponible en ATCC y que expresa E16 de forma endógena en su superficie) con concentraciones cada vez mayores de anticuerpos de toxina conjugada y se monitoreó para efectuar el seguimiento de la muerte celular mediante análisis de bioluminiscencia de viabilidad celular (CellTiterGlo™ Kit, Promega) como se describe arriba. Los resultados que se muestran en la Fig. 159A (células PC3) y la Fig. 159B (células Colo205) comprueban que tres anticuerpos anti-E16 de toxina conjugada promueven la muerte de las células que expresan el E16.

Otra toxina, MMAF, se enlazó al anticuerpo y se realizaron pruebas de citotoxicidad. Los anticuerpos enlazados con toxina 3B5 para este experimento se denominaron 3B5-MC-vc-PAB-MMAE y 3B5-MC-vc-PAB-MMAF. El conjugado anticuerpo-toxina, alfa-IL8-vc-MMAE, se

utilizó como un control negativo. El número promedio de moléculas de toxinas adheridas a cada anticuerpo fue de aproximadamente cuatro moléculas de toxina por anticuerpo como se muestra en la Fig. 160. Se sembraron estirpes celulares de cáncer de colon humano, Colo205, en placas de 96 pozos a 4000 células/pozo en 50 µl de medio de cultivo. El día siguiente, las células se incubaron con anticuerpos 3B5-MC-vc-PAB-MMAE, 3B5-MC-vc-PAB- MMAF o alfa-IL8-vc-MMAE a concentraciones de diferentes grados diluidas en 50 µl del medio de cultivo durante 72hrs. Se midió la viabilidad de las células utilizando un estuche de análisis de viabilidad celular luminiscente CellTiter-Glo™ (Promega). Los resultados que se muestran en la Fig. 160 comprueban que MC-vc-PAB-MMAE y MC-vc-PAB-MMAF-conjugados a anticuerpos anti-E16 tienen la capacidad de matar las células que expresan el E16, con la toxina MMAF (EC50 aproximadamente 0.06 µg/ml) matando más células que el MMAF (EC50 aproximadamente 0.007 µg/ml).

El desarrollo de los anticuerpos de la invención como un agente terapéutico en humanos requiere estudios tempranos de toxicidad en monos y/o primates. La secuencia de aminoácido E16 varía entre especies al igual que entre humano versus mono, rata y ratón donde la secuencia de aminoácidos E16 es 96.26%, 91.05% y 90.66%, respectivamente. Para determinar si los anticuerpos TAT188(E16) anti-humano también pudieron fijarse a y matar las células de los monos, se realizó el siguiente experimento *in vitro*. Se empleó una estirpe celular, COS7 (disponible en ATCC) de mono que expresa el E16 de mono en forma endógena en su superficie. Los anticuerpos anti-E16 3B5, 12B9 y 12G12 se pusieron en contacto con E16 de células de humano, de mono y de ratón respectivamente y se analizaron mediante análisis estándar de FACS. Los FACS mostraron que cada uno de los tres anticuerpos anti-humanos E16 3B5, 12G12 y 12B9 fijan células de mono COS7 expresando el E16 humano (consultar Fig. 161A) al igual que células humanas expresando el E16 humano (consultar Fig. 161B), pero no fijan las células de ratón que

expresan el E16 (consultar Fig. 161B). Por lo tanto, los anticuerpos anti-TAT188(E16) de la invención son útiles para la actividad clínica y para estudios de eficacia y toxicidad.

EJEMPLO 15: Análisis de Muerte de Células de Tumores *In Vivo*

Actividad anti-TAT112:

Para probar la eficacia de los anticuerpos monoclonales sin conjugar anti-TAT112, se inyectó el anticuerpo anti-TAT112 de forma intraperitoneal a ratones atímicos 24 horas antes de recibir 3 células de clon PC3.gD.NCA (obtenidas como se describe en el Ejemplo 14 arriba) de forma subcutánea en el flanco. Las inyecciones del anticuerpo continuaron dos veces por semana durante el resto del estudio. El volumen del tumor se midió dos veces por semana.

Para probar la eficacia de DM1-conjugado a anticuerpo anti-TAT112, se inocularon 3 células de clon PC3.gD.NCA (obtenidas como se describe en el Ejemplo 14 arriba) en el flanco de ratones atímicos. Cuando los tumores alcanzaron un volumen medio de aproximadamente 100mm3, los ratones se trataron con DM1-conjugado a anticuerpo anti-TAT112 de forma intravenosa, ya sea una o dos veces por semana.

Los resultados de los anteriores análisis comprobaron que tanto el anti-TAT112 sin conjugar como el DM1-conjugado a anticuerpo anti-TAT112 fueron sumamente eficaces para reducir el volumen del tumor en este modelo *in vivo*. Estos análisis comprueban que los anticuerpos monoclonales del polipéptido anti-TAT son eficaces para matar células de tumores que expresan un polipéptido TAT de interés.

Actividad anti-TAT118 (E16):

Para probar la eficacia de anticuerpos monoclonales anti-TAT188 sin conjugar, anticuerpos anti-TAT188 3B5, 12B9 y 12G12 para matar células *in vitro*, se llevó a cabo el siguiente procedimiento. Células PC3, que expresan el E16 de forma endógena, se inyectaron en ratones

atímicos desnudos (nu/nu) en el flanco dorsal. El mismo día de la inoculación celular, se inyectaron anticuerpos a 2 mg/kg de forma intraperitoneal dos veces por semana. Los ratones de control fueron inyectados con PBS sin anticuerpo. Se monitorearon diez ratones en cada grupo durante 4 semanas y el volumen del tumor se midió dos veces por semana. Al inyectar las células tumorales y el anticuerpo anti-E16 de forma simultánea, el procedimiento ensayó la capacidad de los anticuerpos administrados para inhibir el establecimiento de un tumor en los ratones xenoinjertados. Los resultados de la Fig. 162 muestran que en este experimento, el volumen medio del tumor no fue significativamente menor que el del anticuerpo de control.

Para probar la eficacia de MC-vc-PAB-MMAE y MC-vc-PAB-MMAF conjugado a anticuerpo anti-TAT188(E16), células Colo205 se inocularon en el flanco de ratones atímicos desnudos (nu/nu), y después de aproximadamente 1 semana, cuando el tumor había alcanzado un volumen medio de aproximadamente 100-200mm^3, los ratones se inyectaron con anticuerpo conjugado a 3 mg/kg de forma intravenosa una vez por semana. Los siguientes tratamientos del anticuerpo se realizaron el día en que los ratones que tenían el tumor se dividieron en grupos: (1) Control simulado - 0.2ml o menos de PBS, inyección intravenosa una vez por semana durante 4 semanas; (2) Anticuerpo de control - anti-IL8-MC-vc-PAB-MMAE en 0.2ml o menos PBS, inyección intravenosa una vez por semana durante 4 semanas; (3) Anticuerpo-3B5-MC-vc-PAB-MMAE en 0.2ml o menos PBS, inyección intravenosa una vez por semana durante 4 semanas; (4) Anticuerpo-3B5-MC-vc-PAB-MMAF en 0.2ml o menos PBS, inyección intravenosa una vez por semana durante 4 semanas. Se midió el volumen medio del tumor y se delineó frente al tiempo. Los resultados en la Fig. 163 muestran que el anticuerpo anti-E16 conjugado a MMAE- y MMAF- redujo significativamente el volumen del tumor en este modelo *in vivo*. Estos análisis comprueban que los anticuerpos monoclonales de polipéptido anti-TAT, como los anticuerpos anti-188(E16) son eficaces para matar las

células tumorales que expresan el polipéptido TAT de interés, como el polipéptido TAT188(E16).

EJEMPLO 16: Análisis de la Técnica Northern

Se efectuó el análisis de la técnica Northern indispensablemente como lo describe Sambrook y colaboradores, supra. El análisis de la técnica Northern empleando sondas derivadas de DNA231542, DNA231542-1, DNA231542-2 y DNA297393 evidencian una significativa regulación por incremento de la expresión en tejido humano de glioma cuando se compara con tejido humano normal de cerebro.

También se realizaron análisis de la técnica Northern empleando sondas derivadas de DNA237637 (TAT188, E16) y los resultados muestran que la expresión en mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado humanos evidencia una significativa regulación por incremento de la expresión en neoplasias humanas cuando se comparan con tejido humano normal del mismo tipo de tejido.

EJEMPLO 17: Modulando la Expresión TAT por ARNsi

Los ARNsi han comprobado ser útiles como una herramienta en estudios para modular la expresión del gen cuando los antagonistas tradicionales como moléculas pequeñas o anticuerpos han fallado. (Shi Y., *Trends in Genetics* 19(1):9-12 (2003)). ARNs bicatenarios, sintetizados in vitro, que tienen de 21 a 23 nucleótidos de largo pueden actuar como ARNs de interferencia (ARNi) y específicamente pueden inhibir la expresión del gen (Fire A., *Trends in Genetics* 391; 806-810 (1999)). Estos ARNi actúan mediando la degradación de sus ARNs objetivo. Aunque tienen menos de 30 nucleótidos de largo, no disparan un mecanismo de defensa antiviral celular. Dichos mecanismos incluyen la producción de interferona y una detención general de la síntesis de la proteína en la célula hospedera. Prácticamente los ARNsi se pueden sintetizar y luego clonar en vectores de ADN. Dichos vectores se pueden transinfectar y se puede hacer que

expresen el ARNsi a altos niveles. El alto nivel de la expresión del ARNsi se utiliza para "inactivar" o significativamente reducir la cantidad de proteína producida en una célula, y por lo tanto es útil en experimentos donde se cree que la sobreexpresión de una proteína se relaciona con un desorden como el cáncer. Los TAT son antígenos de tumor expresados en la superficie celular y la sobreexpresión del TAT se mostró mediante microarrays (pequeñas placas de cristal), Taqman™ y en análisis de hibridación *in situ*. Aunque en la presente se muestra que la fijación del anticuerpo a los TATs como el TAT188 inhibe la función del antígeno de tumor, y con respecto a anticuerpos conjugados a toxinas, causa la apoptosis) los ARNsi son antagonistas útiles para las proteínas de TAT limitando la producción celular del antígeno.

Métodos Experimentales y Resultados: *TAT188 (E16)*

Los siguientes experimentos comprueban que ARNsi específico para el TAT188 (al cual también se hace referencia en la presente como el E16) reduce la expresión del ARN, la producción de proteína, la expresión de la superficie celular y la función de la proteína. Estos procedimientos iguales o similares son útiles para comprobar la regulación por reducción del ARNsi de otros ARNs de TAT de la invención. Se puede seleccionar una secuencia objetivo mediante cualquier procedimiento apropiado (consultar, por ejemplo, Amarzguioui, M. And Prydz, H., Biochem Biophys Res Comm 316: 1050-1058 (2004).

A. La transinfección transiente de ARNsi de E16 reduce la expresión del E16-GFP en células PC3-E16- GFP.

Se utiliza la estirpe celular del cáncer de próstata, PC3, en este conjunto de experimentos puesto que sobreexpresa el TAT188 de forma endógena. Unas células PC3 se transinfectaron empleando oligos ARN 21 bicatenarios dirigidos contra el TAT188 (siTAT188), transinfectados por lámina o simulado. Estos oligos se enumeran a continuación.

oligos ARNsi:

TAT188(E16)

La secuencia de ADN de la región objetivo en la secuencia de codificación del TAT188(E16) se muestra abajo seguida de las secuencias de las hebras sentido y antisentido del ARNsi bicatenario.

AAGGAAGAGGCGCGGGAGAAG (SEQ ID NO: 155), es la secuencia del ácido nucleico en la región de codificación del TAT188 al cual se dirigen los oligos ARNsi utilizados en los ejemplos de ARNsi del TAT188 divulgados en la presente. Se pueden escoger otras secuencias como objetivos de acuerdo con la práctica estándar en el campo de silenciamiento de los genes del ARNsi.

TAT188(E16) oligos ARNsi:

Sentido: r(GGAAGAGGCGCGGGAGAAG)TT (SEQ ID NO: 156)

Antisentido: r(CUUCUCCCGCGCCUCUUCC)TT (SEQ ID NO: 157)

silámina ARNsi oligos:

Sentido: r(CUGGACUUCCAGAAGAACA)TT (SEQ ID NO: 158)

Antisentido: r(UGUUCUGGAAGUCCAG)TT (SEQ ID NO: 159)

(consultar Elbashir y colaboradores, Nature 411:494-498 (2001)

Dos días después de la transinfección, se analizó el nivel relativo del polipéptido TAT188(E16) versus la inactivación de la expresión del E16-EGFP mediante microscopio fluorescente monitoreando la fluorescencia EGFP. Una reducción en la expresión del El6 es paralela a una reducción en la producción de EGFP y en la bioluminiscencia detectada. Los resultados en la Fig. 164 comprueban que la transinfección transiente con ARNsi del TAT188si (E16) redujo la expresión de fusión del E16-EGFP. La proteína de fusión denominada "GFP" en las Figs. 164 y 164 hace

referencia a una fusión E16-EGFP producida por la clonación del gen E16 en el vector pEGFP-Cl de BDClontech).

B. La transinfección transiente del ARNsi del E16 reduce la producción del polipéptido TAT188(E16)

Se comprobó la reducción de la expresión del E16 mediante la reducción en la proteína E16 utilizando el análisis de inmunotransferencia. Los ARN oligos TAT188si (E16si) SEQ ID NO: 156 y SEQ ID NO: 157 y los lámina ARNsi oligos SEQ ID NO: 158 y SEQ ID NO: 159 empleados en este ejemplo fueron los mismos que se mostraron arriba en la parte A de este ejemplo. Se recolectaron las células PC3 transinfectadas en forma transiente con E16-EGFP 3 días después de la transinfección y se realizó una inmunotransferencia a una banda beta-tubulina para indicar la carga de la proteína. Se visualizó la inmunotransferencia de las fusiones TAT188(E16)- GFP utilizando un anticuerpo anti-GFP marcado de forma detectable (consultar Fig. 165, fila izquierda – transinfección simulada; fila derecha, E16 ARNsi transinfección). Los resultados comprueban que la presencia de TAT188(E16) ARNsi en una célula que expresa la proteína de fusión E16-GFP redujo la expresión de E16-GFP de forma significativa.

C. La transinfección transiente del ARNsi de E16 inhibe la actividad transportadora endógena de aminoácidos E16 en las células PC3.

Se realizó el mismo procedimiento para la transinfección transiente del ARNsi del TAT188(E16) y lámina ARNsi (control). La prueba y las células de control se probaron para identificar la actividad E16 como una proteína de transporte del aminoácido como sigue, utilizando el análisis de transporte de leucina descrito en la presente arriba. Los resultados en la Fig. 166 muestran que una reducción en el ARNsi del E16 en la producción de proteína en el ARN del E16 es consistente con la

reducción de la cantidad de E16 disponible para funcionar en el transporte del aminoácido en la célula.

D. La transinfección transiente del ARNsi de E16 reduce la expresión de superficie del E16

Este ejemplo se llevó a cabo como sigue. Cuarenta y ocho horas después de la transinfección de ARNsi, se analizaron las células mediante FACS. Brevemente, las células se incubaron con anticuerpos primarios durante una hora sobre hielo. Luego las células se marcaron con un anticuerpo secundario conjugado con un tinte fluorescente (como los ejemplos no limitantes Cy-3 o PE) y se recolectó la información con FACScaliber™ o FACScan™. Se analizó la información recolectada empleando, como ejemplos no limitantes, Cellquest™ (BD Clontech) o FlowJo ™ (Tree Star, Inc.). Como se divulga en el Ejemplo 12 en la presente, los tres anticuerpos monoclonales anti-E16 3b5, 12B9 y 12G12, mostraron que se fijan menos a las células PC3 después de la transinfección del ARNsi del TAT188(E16) consistente con una expresión reducida de la proteína E16 en presencia del ARNsi objetivo (consultar Fig. 156).

E. La transinfección transiente del ARNsi del TAT188(E16) inhibe la proliferación de las células PC3

La proliferación de las células PC3 transinfectadas con ARNsi apuntando al TAT188 se pueden analizar de acuerdo con el siguiente protocolo comercial (Promega Corp. Technical Bulletin TB288; Mendoza y colaboradores (2202) Cancer Res. 62:5485-5488). Por ejemplo, se llevaron células PC3 a unas placas a $3.5x10^5$ células/pozo en placas de 6 pozos. El día siguiente, se utilizó ARN bicatenario (ARNds, SEQ ID NO: 155 y SEQ ID NO: 156) o lámina ARNsi de control (utilizando ARNsi oligos SEQ ID NO: 157 y SEQ ID NO: 158) a una concentración final de 100 nM para transinfectar las células utilizando Lipofectamin2000 (Invitrogen). Después de 6 hrs de transinfección, las células se resembraron en placas de 96 pozos,

y se midió la viabilidad celular utilizando el Cell Counting Kit-8™ (Dojindo) en los periodos de tiempo indicados en la Fig. 167, donde RLU = unidades de luminiscencia relativa. Otras células que se pueden analizar mediante este método si expresan el TAT188(ya sea en forma endógena y/o después de la transinfección con ácido nucleico codificando el TAT188) incluyen sin limitación SKBR-3, BT474, MCF7 o MDA-MB-468 comprobando que la reducción de la expresión del E16 en las células de cáncer de mama resulta en la muerte de las células y que el cáncer de mama, al igual que el cáncer de próstata, colorrectal y otros tipos de cáncer, son blancos deseables para el tratamiento empleando la inactivación de la expresión del E16.

Estos datos indican que los ARNsis pueden "inactivar" o reducir la expresión de las proteínas del TAT como la proteína del TAT188 (E16). La reducción se muestra a través de la reducción de la expresión del ARN (en análisis de hibridación *in situ*), una producción reducida de proteína (como se muestra en los análisis de inmunotransferencia), una reducción de la expresión del TAT188 (E16) en la superficie celular (como se muestra en los análisis FACS y en los análisis de internalización), y una función reducida (como se muestra por la reducción del transporte del aminoácido). La reducción del ARNsi de las proteínas de TAT, como el TAT188 (E16) durante un periodo de tiempo resulta en una reducción de la proliferación celular en estirpes celulares de cáncer. Por lo tanto el ARNsi es útil para reducir las proteínas TAT como el TAT188 (E16). La reducción del nivel de las proteínas TAT, como el TAT188 (E16), causa una reducción en la proliferación de las estirpes celulares del cáncer. Puesto que el cáncer es un desorden de proliferación celular, estos resultados muestran que los antagonistas de los polipéptidos TAT como el TAT188 (E16) serían útiles para la reducción del crecimiento de células cancerosas. La reducción del crecimiento de células cancerosas sería útil para aliviar el cáncer en los mamíferos.

Ejemplo 18: Depósito de Materiales

La siguiente estirpe celular de hibridoma se depositó en la American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209 USA (ATCC):

Designación del Hibridoma/Anticuerpo	ATCC No.	Fecha de depósito
3B5.1	PTA-6193	Septiembre 8, 2004
12B9.1	PTA-6194	Septiembre 8, 2004
12G12.1	PTA-6195	Septiembre 8, 2004

Este depósito se realizó de acuerdo con las disposiciones del Tratado de Budapest sobre el Reconocimiento Internacional del Depósito de Microorganismos para Efectos de Procedimientos de Patentes y la regulación del mismo (Tratado de Budapest). Esto garantiza la preservación de un cultivo viable durante 30 años a partir de la fecha del depósito. La estirpe celular estará disponible en ATCC bajo los términos del Tratado de Budapest, y sujetos a un acuerdo entre Genentech, Inc. y la ATCC, que garantiza (a) que el acceso al cultivo estará disponible mientras la solicitud de patente se encuentra pendiente a quien designe el Comisionado como el que tiene derecho a esta bajo 37 CFR §1.14 y 35 USC §122, y (b) que todas las restricciones sobre disponibilidad para el público del cultivo así depositado serán retiradas de forma irrevocable cuando se otorgue la patente.

El beneficiario de la presente solicitud acepta que si el cultivo en depósito llegara a morir o se perdiera o destruyera cuando se cultive bajo condiciones apropiadas, de inmediato será reemplazado luego de una notificación, con un espécimen viable del mismo cultivo. La disponibilidad de la estirpe celular depositada no será interpretada como una licencia para poner en práctica la invención en contravención de los derechos otorgados bajo la autoridad de cualquier gobierno de conformidad con sus leyes sobre patentes.

La anterior especificación escrita se considera suficiente para que una persona con destreza en el arte ponga la invención en práctica. En ámbito de la presente invención no se deberá limitar por el constructo depositado, puesto que la incorporación depositada tiene la finalidad de una sola ilustración de ciertos aspectos de la invención, y cualquier construcción funcionalmente equivalente está dentro del ámbito de esta invención. El depósito de material de la presente no constituye una admisión de que la descripción escrita aquí contenida es inapropiada para poder poner en práctica cualquier aspecto de la invención, incluyendo el mejor modo de la misma, y no se deberá interpretar como una limitación al ámbito de las reivindicaciones de las ilustraciones específicas que representa. De hecho, varias modificaciones de la invención, además de las que se muestran en la presente, serán aparentes para aquellos con destreza en el arte, a partir de la anterior descripción y caen en el ámbito de las reivindicaciones adjuntas.

LO QUE SE REIVINDICA ES:

1. Un anticuerpo aislado que se fija a un polipéptido teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con:

a) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

b) el polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

c) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su péptido de señal asociado;

d) un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su péptido de señal asociado;

e) un polipéptido codificado por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

2. Un anticuerpo aislado que se fija a un polipéptido teniendo:

a) la secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154);

b) la secuencia de aminoácidos que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

c) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su secuencia de péptido de señal asociado;

d) una secuencia de aminoácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ

ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

e) una secuencia de aminoácidos codificada por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

f) una secuencia de aminoácidos codificada por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

3. El anticuerpo de la Reivindicación 1 el cual es un anticuerpo monoclonal.

4. El anticuerpo de la Reivindicación 1 el cual es un fragmento de anticuerpo.

5. El anticuerpo de la Reivindicación 1 el cual es un anticuerpo quimérico o humanizado.

6. El anticuerpo de la Reivindicación 1 el cual se conjuga con un agente inhibidor de crecimiento.

7. El anticuerpo de la Reivindicación 1 el cual se conjuga con un agente citotóxico.

8. El anticuerpo de la Reivindicación 7, donde el agente citotóxico se selecciona entre el grupo conformado por toxinas, antibióticos, isótopos radioactivos y enzimas nucleolíticas.

9. El anticuerpo de la Reivindicación 7, donde el agente citotóxico es una toxina.

10. El anticuerpo de la Reivindicación 9, donde la toxina se selecciona entre el grupo conformado por auristatina, maitansinoide y calicheamicina.

11. El anticuerpo de la Reivindicación 9, donde la toxina es un maitansinoide.

12. El anticuerpo de la Reivindicación 1 el cual se produce en bacterias.

13. El anticuerpo de la Reivindicación 1 el cual se produce en células CHO.

14. El anticuerpo de la Reivindicación 1 que induce la muerte de una célula a la cual se fija.

15. El anticuerpo de la Reivindicación 1 el cual se marca en forma detectable.

16. El anticuerpo de la Reivindicación 9, donde la toxina se selecciona entre el grupo conformado por MMAE (mono-metil auristatina E), MMAF, (auristatina E valeril bencihidrazona) y AFP (Auristatina F fenileno diamina).

17. El anticuerpo de la Reivindicación 9, donde la toxina se adhiere en forma covalente al anticuerpo mediante un enlazante.

18. El anticuerpo de la Reivindicación 17, donde el enlazante se selecciona entre el grupo conformado por maleimidocaproilo (MC), valina-citrulina (val-cit, vc), citrulina (2-amino-5-ureido ácido pentanoico), PAB (p-aminobencilcarbamoilo), Me (N-metil-valina citrulina), MC(PEG)6-OH (maleimidocaproilo-polietilenglicol), SPP (N-Succinimidil 4-(2-

piridiltio)pentanoato), SMCC (N-Succinimidil 4-(N-maleimidometil) ciclohexano-1 carboxilato) y MC-vc-PAB.

19. El anticuerpo de la Reivindicación 16, donde la toxina es MMAE.

20. El anticuerpo de la Reivindicación 16, donde la toxina es MMAF.

21. El anticuerpo de la Reivindicación 19 0 20, donde el enlazante es MC-vc-PAB.

22. El anticuerpo de una cualquiera de las Reivindicaciones 1-21 donde el anticuerpo fija el polipéptido TAT188.

23. El anticuerpo de la Reivindicación 22, donde el anticuerpo inhibe la proliferación o promueve la muerte celular de una célula que expresa el TAT188.

24. El anticuerpo de la Reivindicación 23, donde la célula es una célula cancerosa.

25. El anticuerpo de la Reivindicación 24, donde la célula cancerosa se selecciona entre el grupo de mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado.

26. Un anticuerpo aislado que compite con la fijación de los epítopos del polipéptido TAT188 unido por un anticuerpo producido por el hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No. PTA-6193), 12B9.1 (ATCC Acceso No. PTA-6194) y 12G12.1 (ATCC Acceso No. PTA-6195).

27. Un anticuerpo aislado teniendo la actividad biológica de un anticuerpo producido por el hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No. PTA-6193), 12B9.1 (ATCC Acceso No. PTA-6194) y 12G12.1 (ATCC Acceso No. PTA- 6195), donde la actividad biológica es la inhibición de la proliferación celular o la promoción de la muerte celular en una célula que expresa el TAT188.

28. Un anticuerpo aislado que comprende en una región de determinante complementario (CDR) correspondiente, una secuencia de aminoácidos teniendo por lo menos 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de una secuencia de aminoácidos de por lo menos 1, 2, 3, 4, 5 o 6 de los CDR(s) del anticuerpo producido por un hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No. PTA-6193), 12B9.1 (ATCC Acceso No. PTA-6194) y 12G12.1 (ATCC Acceso No. PTA-6195).

29. El anticuerpo de una cualquiera de las Reivindicaciones 1-21, donde el anticuerpo comprende en una región de determinante complementario (CDR) correspondiente, una secuencia de aminoácidos teniendo por lo menos 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de una secuencia de aminoácidos de por lo menos 1, 2, 3, 4, 5 o 6 de los CDR(s) del anticuerpo producido por un hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No. PTA- 6193), 12B9.1 (ATCC Acceso No. PTA-6194) y 12G12.1 (ATCC Acceso No. PTA-6195).

30. El anticuerpo de una cualquiera de las Reivindicaciones 1-21, donde el anticuerpo exhibe la actividad biológica de un anticuerpo producido por el hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No. PTA-6193), 12B9.1 (ATCC Acceso No. PTA-6194) y 12G12.1 (ATCC Acceso No. PTA-6195), donde la actividad biológica es la

inhibición de la proliferación celular o la promoción de la muerte celular en una célula que expresa el TAT188.

31. El anticuerpo de la Reivindicación 30, donde la célula es una célula cancerosa.

32. El anticuerpo de la Reivindicación 31, donde la célula cancerosa se selecciona entre el grupo conformado por mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado.

33. Un método para inhibir el crecimiento de una célula que expresa el TAT188, el método que comprende poner una célula en contacto con un anticuerpo de una cualquiera de las Reivindicaciones 1-21.

34. El método de la Reivindicación 33, donde la célula es una célula cancerosa.

35. El método de la Reivindicación 34, donde la célula cancerosa se selecciona entre el grupo conformado por mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado.

36. El método de la Reivindicación 35, donde la célula cancerosa es una célula de mamífero.

37. El método de la Reivindicación 36, donde la célula de mamífero es una célula humana.

38. Un método para inhibir el crecimiento de una célula que expresa el TAT188, el método que comprende poner una célula en contacto con un anticuerpo de una cualquiera de las Reivindicaciones 185-205, donde el anticuerpo comprende en una región de determinante complementario

(CDR) correspondiente, una secuencia de aminoácidos teniendo por lo menos 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de una secuencia de aminoácidos de por lo menos 1, 2, 3, 4, 5 o 6 de las CDR(s) del anticuerpo producido por un hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No. PTA-6193), 12B9.1 (ATCC Acceso No. PTA-6194) y 12G12.1 (ATCC Acceso No. PTA-6195).

39. Un método para detectar el nivel del polipéptido TAT188 expresado en una célula de prueba en relación con una célula de control, el método que comprende:

a) poner en contacto la célula de prueba y la célula de control con un anticuerpo anti-TAT188 aislado de la Reivindicación 26;

b) detectar la fijación del anticuerpo; y

c) determinar la fijación relativa del anticuerpo a la célula de prueba y de control.

40. El método de la Reivindicación 39, donde la célula de prueba y la célula de control se someten a lisis.

41. El método de la Reivindicación 39, donde la célula de prueba está en un tejido.

42. El método de la Reivindicación 41, donde el tejido es un tejido de tumor.

43. El método de la Reivindicación 42, donde el tejido de tumor se selecciona entre el grupo conformado por mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado.

44. Un método para detectar el nivel del polipéptido TAT188 o un polipéptido teniendo una identidad de secuencia por lo menos de 80% con una secuencia de aminoácidos que se muestra en la Figura 115 (SEQ ID No: 115) en una célula de prueba en relación con una célula de control, el método que comprende:

a) poner en contacto la célula de prueba y la célula de control con un anticuerpo aislado de una cualquiera de las Reivindicaciones 1-5, 12-15 y 26-28

b) detectar la fijación del anticuerpo; y

c) determinar la fijación relativa del anticuerpo a la célula de prueba y de control.

45. El método de la Reivindicación 44, donde el nivel del polipéptido TAT188 en la célula de prueba es mayor que el de la célula de control.

46. El método de la Reivindicación 45, donde el método diagnostica el cáncer en un tejido que contiene o que tiene contenida la célula de prueba.

47. Un ARN de interferencia de unión TAT (ARNsi) que se fija a un ácido nucleico teniendo una identidad de secuencia por lo menos de 80% con:

a) una secuencia nucleótida que se muestra en la Figura 37 (SEQ ID No:37; y

b) el complemento de (a), donde el ARNsi reduce la expresión del TAT188.

48. Un vector de expresión que comprende el ARNsi de la Reivindicación 47.

49. El vector de expresión de la Reivindicación 48, donde el citado ARNsi está enlazado operacionalmente a las secuencias de control reconocidas por una célula hospedera transinfectada con el vector.

50. Una célula hospedera que comprende el vector de expresión de la Reivindicación 49.

51. Una composición de sustancias que comprende:

a) el anticuerpo de la Reivindicación 1, o

b) el ARNsi de la Reivindicación 47,

en combinación con un excipiente.

52. La composición de las sustancias de la Reivindicación 51, donde el citado excipiente es un excipiente farmacéuticamente aceptable.

53. Un artículo de manufactura:

a) un envase; y

b) la composición de las sustancias de la Reivindicación 51 contenida en el citado envase.

54. El artículo de manufactura de la Reivindicación 53, que además comprende una etiqueta adherida al citado envase, o un inserto en el paquete incluido con el citado envase, haciendo referencia al uso de la citada composición de sustancias para el tratamiento terapéutico o la detección diagnóstica de un cáncer.

55. Un método para inhibir el crecimiento de una célula cancerosa que expresa un polipéptido teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con una secuencia de aminoácidos que se muestra en la Figura 115 (SEQ ID No: 115), el citado método que comprende poner en contacto la citada célula cancerosa con un ARNsi que se fija a un ácido nucleico codificando el aminoácido en la citada célula cancerosa, en esta forma inhibiendo el crecimiento de la citada célula cancerosa.

56. El método de la Reivindicación 55, donde el ácido nucleico tiene la secuencia que se muestra en la Figura 37 (SEQ ID No:37).

57. El método de la Reivindicación 44, donde la detección del nivel de expresión del polipéptido comprende emplear un anticuerpo en un análisis de inmunohistoquímica.

58. Un método para tratar o prevenir un desorden de proliferación celular asociado con una expresión o actividad aumentada de un polipéptido teniendo una identidad de secuencia de aminoácidos por lo menos de 80% con una secuencia de aminoácidos que se muestra en la Figura 115 (SEQ ID No: 115), el citado método que comprende administrar a un sujeto que necesite el citado tratamiento una cantidad efectiva de un antagonista de un polipéptido TAT188.

59. El método de la Reivindicación 58, donde el citado antagonista es un anticuerpo aislado de un polipéptido anti-TAT188 de una cualquiera de las Reivindicaciones 1-21 y 23-28.

60. El método de la Reivindicación 59, donde el desorden de proliferación celular es cáncer.

61. El método de la Reivindicación 60, donde el cáncer se selecciona entre el grupo conformado por mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado.

FIGURA 1

CCAAGAATTCGGCACGAGGGTGACCCGGACAGCTGTCCTCTCTGACACCACCCCGGCCTGC
CTCTTTGTTGCCATGAGAGCTGCCTACCTCTTCCTGCTATTCCTGCCTGCAGGCTTGCTGG
CTCAGGGCCAGTATGACCTGGACCCGCTGCCGCCGTTCCCTGACCACGTCCAGTACACCCA
CTATAGCGACCAGATCGACAACCCAGACTACTATGATTATCAAGAGGTGACTCCTCGGCCC
TCCGAGGAACAGTTCCAGTTCCAGTCCCAGCAGCAAGTCCAACAGGAAGTCATCCCAGCCC
CAACCCCAGAACCAGGAAATGCAGAGCTGGAGCCCACAGAGCCTGGGCCTCTTGACTGCCG
TGAGGAACAGTACCCGTGCACCCGCCTCTACTCCATACACAGGCCTTGCAAACAGTGTCTC
AACGAGGTCTGCTTCTACAGCCTCCGCCGTGTGTACGTCATTAACAAGGAGATCTGTGTTC
GTACAGTGTGTGCCCACGAGGAGCTCCTCCGAGCTGACCTCTGTCGGGACAAGTTCTCCAA
ATGTGGCGTGATGGCCAGCAGCGGCCTGTGCCAATCCGTGGCGGCCTCCTGTGCCAGGAGC
TGTGGGAGCTGCTAGGGTGGTGCTGGCATCCTGAGTCCTGGCCCTCCTGGGATCTGGGGCC
CTCGGGCCCTGCCTGACCTGGTGCTTTTTTCCCCATCCCCATGTTCCTTTTATTCTGTAAA
AAGTTAGTGGACTGCAGCCCTGGGGGTTGCAGGCTGCGGTGCCTCAGGCCCCTCCTTCAGC
CTGTGGCCACCTCTGGGGCACGATGGGGGCTCCCCACTGCCCAGTCTGCCCCTCGGGTTGG
GGGAGTATCCCAGGCCTCTCTGTGGGACCTGGGCCCCTGACGGGCCTTCTCAGCCCGTTTT
GAGGACAGACAGTCCCCCGAGGTAGGCTACATCCCCCCACCCCAGCTGGTCTGCTTGGATT
TCCTACAGCCCCCGTGGGCATGGACCACCTTTATTTTATACAAAATTAAAAACAAGTTTTT
AC

FIGURA 2A

CACTAACGCTCTTCCTAGTCCCCGGGCCAACTCGGACAGTTTGCTCATTTATTGCAACGGT
CAAGGCTGGCTTGTGCCAGAACGGCGCGCGCGCGACGCACGCACACACACGGGGGGAAACT
TTTTTAAAAATGAAAGGCTAGAAGAGCTCAGCGGCGGCGCGGGCCGTGCGCGAGGGCTCCG
GAGCTGACTCGCCGAGGCAGGAAATCCCTCCGGTCGCGACGCCCGGCCCCGCTCGGCGCCC
GCGTGGGATGGTGCAGCGCTCGCCGCCGGGCCCGAGAGCTGCTGCACTGAAGGCCGGCGAC
GATGGCAGCGCGCCCGCTGCCCGTGTCCCCCGCCCGCGCCCTCCTGCTCGCCCTGGCCGGT
GCTCTGCTCGCGCCCTGCGAGGCCCGAGGGGTGAGCTTATGGAACGAAGGAAGAGCTGATG
AAGTTGTCAGTGCCTCTGTTCGGAGTGGGGACCTCTGGATCCCAGTGAAGAGCTTCGACTC
CAAGAATCATCCAGAAGTGCTGAATATTCGACTACAACGGGAAAGCAAAGAACTGATCATA
AATCTGGAAAGAAATGAAGGTCTCATTGCCAGCAGTTTCACGGAAACCCACTATCTGCAAG
ACGGTACTGATGTCTCCCTCGCTCGAAATTACACGGTAATTCTGGGTCACTGTTACTACCA
TGGACATGTACGGGGATATTCTGATTCAGCAGTCAGTCTCAGCACGTGTTCTGGTCTCAGG
GGACTTATTGTGTTTGAAAATGAAAGCTATGTCTTAGAACCAATGAAAAGTGCAACCAACA
GATACAAACTCTTCCCAGCGAAGAAGCTGAAAAGCGTCCGGGGATCATGTGGATCACATCA
CAACACACCAAACCTCGCTGCAAAGAATGTGTTTCCACCACCCTCTCAGACATGGGCAAGA
AGGCATAAAAGAGAGACCCTCAAGGCAACTAAGTATGTGGAGCTGGTGATCGTGGCAGACA
ACCGAGAGTTTCAGAGGCAAGGAAAAGATCTGGAAAAAGTTAAGCAGCGATTAATAGAGAT
TGCTAATCACGTTGACAAGTTTTACAGACCACTGAACATTCGGATCGTGTTGGTAGGCGTG
GAAGTGTGGAATGACATGGACAAATGCTCTGTAAGTCAGGACCCATTCACCAGCCTCCATG
AATTTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAAATCCCATGACAATGCGCAGCT
TGTCAGTGGGGTTTATTTCCAAGGGACCACCATCGGCATGGCCCCAATCATGAGCATGTGC
ACGGCAGACCAGTCTGGGGGAATTGTCATGGACCATTCAGACAATCCCCTTGGTGCAGCCG
TGACCCTGGCACATGAGCTGGGCCACAATTTCGGGATGAATCATGACACACTGGACAGGGG
CTGTAGCTGTCAAATGGCGGTTGAGAAAGGAGGCTGCATCATGAACGCTTCCACCGGGTAC
CCATTTCCCATGGTGTTCAGCAGTTGCAGCAGGAAGGACTTGGAGACCAGCCTGGAGAAAG
GAATGGGGGTGTGCCTGTTTAACCTGCCGGAAGTCAGGGAGTCTTTCGGGGGCCAGAAGTG
TGGGAACAGATTTGTGGAAGAAGGAGAGGAGTGTGACTGTGGGGAGCCAGAGGAATGTATG
AATCGCTGCTGCAATGCCACCACCTGTACCCTGAAGCCGGACGCTGTGTGCGCACATGGGC
TGTGCTGTGAAGACTGCCAGCTGAAGCCTGCAGGAACAGCGTGCAGGGACTCCAGCAACTC
CTGTGACCTCCCAGAGTTCTGCACAGGGGCCAGCCCTCACTGCCCAGCCAACGTGTACCTG
CACGATGGGCACTCATGTCAGGATGTGGACGGCTACTGCTACAATGGCATCTGCCAGACTC
ACGAGCAGCAGTGTGTCACACTCTGGGGACCAGGTGCTAAACCTGCCCCTGGGATCTGCTT
TGAGAGAGTCAATTCTGCAGGTGATCCTTATGGCAACTGTGGCAAAGTCTCGAAGAGTTCC
TTTGCCAAATGCGAGATGAGAGATGCTAAATGTGGAAAAATCCAGTGTCAAGGAGGTGCCA

GCCGGCCAGTCATTGGTACCAATGCCGTTTCCATAGAAACAAACATCCCCCTGCAGCAAGG
AGGCCGGATTCTGTGCCGGGGGACCCACGTGTACTTGGGCGATGACATGCCGGACCCAGGG
CTTGTGCTTGCAGGCACAAAGTGTGCAGATGGAAAAATCTGCCTGAATCGTCAATGTCAAA
ATATTAGTGTCTTTGGGGTTCACGAGTGTGCAATGCAGTGCCACGGCAGAGGGGTGTGCAA
CAACAGGAAGAACTGCCACTGCGAGGCCCACTGGGCACCTCCCTTCTGTGACAAGTTTGGC
TTTGGAGGAAGCACAGACAGCGGCCCCATCCGGCAAGCAGATAACCAAGGTTTAACCATAG
GAATTCTGGTGACCATCCTGTGTCTTCTTGCTGCCGGATTTGTGGTTTATCTCAAAAGGAA
GACCTTGATACGACTGCTGTTTACAAATAAGAAGACCACCATTGAAAAACTAAGGTGTGTG
CGCCCTTCCCGGCCACCCCGTGGCTTCCAACCCTGTCAGGCTCACCTCGGCCACCTTGGAA
AAGGCCTGATGAGGAAGCCGCCAGATTCCTACCCACCGAAGGACAATCCCAGGAGATTGCT
GCAGTGTCAGAATGTTGACATCAGCAGACCCCTCAACGGCCTGAATGTCCCTCAGCCCCAG
TCAACTCAGCGAGTGCTTCCTCCCCTCCACCGGGCCCCACGTGCACCTAGCGTCCCTGCCA
GACCCCTGCCAGCCAAGCCTGCACTTAGGCAGGCCCAGGGGACCTGTAAGCCAAACCCCCC
TCAGAAGCCTCTGCCTGCAGATCCTCTGGCCAGAACAACTCGGCTCACTCATGCCTTGGCC
AGGACCCCAGGACAATGGGAGACTGGGCTCCG

FIGURA 2B

CCTGGCACCCCTCAGACCTGCTCCACAATATCCACACCAAGTGCCCAGATCCACCCACACC
GCCTATATTAAGTGAGAAGCCGACACCTTTTTTCAACAGTGAAGACAGAAGTTTGCACTAT
CTTTCAGCTCCAGTTGGAGTTTTTTGTACCAACTTTTAGGATTTTTTTTTAATGTTTAAAAC
ATCATTACTATAAGAACTTTGAGCTACTGCCGTCAGTGCTGTGCTGTGCTATGGTGCTCTG
TCTACTTGCACAGGTACTTGTAAATTATTAATTTATGCAGAATGTTGATTACAGTGCAGTG
CGCTGTAGTAGGCATTTTTACCATCACTGAGTTTTCCATGGCAGGAAGGCTTGTTGTGCTT
TTAGTATTTTAGTGAACTTGAAATATCCTGCTTGATGGGATTCTGGACAGGATGTGTTTGC
TTTCTGATCAAGGCCTTATTGGAAAGCAGTCCCCCAACTACCCCCAGCTGTGCTTATGGTA
CCAGATGCAGCTCAAGAGATCCCAAGTAGAATCTCAGTTGATTTTCTGGATTCCCCATCTC
AGGCCAGAGCCAAGGGGCTTCAGGTCCAGGCTGTGTTTGGCTTTCAGGGAGGCCCTGTGCC
CCTTGACAACTGGCAGGCAGGCTCCCAGGGACACCTGGGAGAAATCTGGCTTCTGGCCAGG
AAGCTTTGGTGAGAACCTGGGTTGCAGACAGGAATCTTAAGGTGTAGCCACACCAGGATAG
AGACTGGAACACTAGACAAGCCAGAACTTGACCCTGAGCTGACCAGCCGTGAGCATGTTTG
GAAGGGGTCTGTAGTGTCACTCAAGGCGGTGCTTGATAGAAATGCCAAGCACTTCTTTTTC
TCGCTGTCCTTTCTAGAGCACTGCCACCAGTAGGTTATTTAGCTTGGGAAAGGTGGTGTTT
CTGTAAGAAACCTACTGCCCAGGCACTGCAAACCGCCACCTCCCTATACTGCTTGGAGCTG
AGCAAATCACCACAAACTGTAATACAATGATCCTGTATTCAGACAGATGAGGACTTTCCAT
GGGACCACAACTATTTTCAGATGTGAACCATTAACCAGATCTAGTCAATCAAGTCTGTTTA
CTGCAAGGTTCAACTTATTAACAATTAGGCAGACTCTTTATGCTTGCAAAAACTACAACCA
ATGGAATGTGATGTTCATGGGTATAGTTCATGTCTGCTATCATTATTCGTAGATATTGGAC
AAAGAACCTTCTCTATGGGGCATCCTCTTTTTCCAACTTGGCTGCAGGAATCTTTAAAAGA
TGCTTTTAACAGAGTCTGAACCTATTTCTTAAACACTTGCAACCTACCTGTTGAGCATCAC
AGAATGTGATAAGGAAATCAACTTGCTTATCAACTTCCTAAATATTATGAGATGTGGCTTG
GGCAGCATCCCCTTGAACTCTTCACTCTTCAAATGCCTGACTAGGGAGCCATGTTTCACAA
GGTCTTTAAAGTGACTAATGGCATGAGAAATACAAAAATACTCAGATAAGGTAAAATGCCA
TGATGCCTCTGTCTTCTGGACTGGTTTTCACATTAGAAGACAATTGACAACAGTTACATAA
TTCACTCTGAGTGTTTTATGAGAAAGCCTTCTTTTGGGGTCAACAGTTTTCCTATGCTTTG
AAACAGAAAAATATGTACCAAGAATCTTGGTTTGCCTTCCAGAAAACAAAACTGCATTTCA
CTTTCCCGGTGTTCCCCACTGTATCTAGGCAACATAGTATTCATGACTATGGATAAACTAA
ACACGTGACACAAACACACACAAAAGGGAACCCAGCTCTAATACATTCCAACTCGTATAGC
ATGCATCTGTTTATTCTATAGTTATTAAGTTCTTTAAAATGTAAAGCCATGCTGGAAAATA
ATACTGCTGAGATACATACAGAATTACTGTAACTGATTACACTTGGTAATTGTACTAAAGC
CAAACATATATATACTATTAAAAAGGTTTACAGAATTTTATGGTGCATTACGTGGGCATTG
TCTTTTTAGATGCCCAAATCCTTAGATCTGGCATGTTAGCCCTTCCTCCAATTATAAGAGG
ATATGAACCAAAAAAAAAAAAAAAAAA

FIGURA 3

GGAGTTTCAAAGAAGCGCTAGTAAGGTCTCTGAGATCCTTGCACTAGCTACATCCTCAGGG
TAGGAGGAAGATGGCTTCCAGAAGCATGCGGCTGCTCCTATTGCTGAGCTGCCTGGCCAAA
ACAGGAGTCCTGGGTGATATCATCATGAGACCCAGCTGTGCTCCTGGATGGTTTTACCACA
AGTCCAATTGCTATGGTTACTTCAGGAAGCTGAGGAACTGGTCTGATGCCGAGCTCGAGTG
TCAGTCTTACGGAAACGGAGCCCACCTGGCATCTATCCTGAGTTTAAAGGAAGCCAGCACC
ATAGCAGAGTACATAAGTGGCTATCAGAGAAGCCAGCCGATATGGATTGGCCTGCACGACC
CACAGAAGAGGCAGCAGTGGCAGTGGATTGATGGGGCCATGTATCTGTACAGATCCTGGTC
TGGCAAGTCCATGGGTGGGAACAAGCACTGTGCTGAGATTAGCTCCAATAACAACTTTTTA
ACTTGGAGCAGCAACGAATGCAACAAGCGCCAACACTTCCTGTGCAAGTACCGACCATAGA
GCAAGAATCAAGATTCTGCTAACTCCTGCACAGCCCCGTCCTCTTCCTTTCTGCTAGCCTG
GCTAAATCTGCTCATTATTTCAGAGGGGAAACCTAGCAAACTAAGAGTGATAAGGGCCCTA
CTACACTGGCTTTTTTAGGCTTAGAGACAGAAACTTTAGCATTGGCCCAGTAGTGGCTTCT
AGCTCTAAATGTTTGCCCCGCCATCCCTTTCCACAGTATCCTTCTTCCCTCCTCCCCTGTC
TCTGGCTGTCTCGAGCAGTCTAGAAGAGTGCATCTCCAGCCTATGAAACAGCTGGGTCTTT
GGCCATAAGAAGTAAAGATTTGAAGACAGAAGGAAGAAACTCAGGAGTAAGCTTCTAGACC
CCTTCAGCTTCTACACCCTTCTGCCCTCTCTCCATTGCCTGCACCCCACCCCAGCCACTCA
ACTCCTGCTTGTTTTTCCTTTGGCCATAGGAAGGTTTACCAGTAGAATCCTTGCTAGGTTG
ATGTGGGCCATACATTCCTTTAATAAACCATTGTGTACATAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURA 4

GGGGAGCAGAGAGGAGGCAATGGCCACCATGGAGAACAAGGTGATCTGCGCCCTGGTCCTG
GTGTCCATGCTGGCCCTCGGCACCCTGGCCGAGGCCCAGACAGAGACGTGTACAGTGGCCC
CCCGTGAAAGACAGAATTGTGGTTTTCCTGGTGTCACGCCCTCCCAGTGTGCAAATAAGGG
CTGCTGTTTCGACGACACCGTTCGTGGGGTCCCCTGGTGCTTCTATCCTAATACCATCGAC
GTCCCTCCAGAAGAGGAGTGTGAATTTTAGACACTTCTGCAGGGATCTGCCTGCATCCTGA
CGCGGTGCCATCCCCAGCACGGTGATTAGTCCCAGAGCTCGGCTGCCACCTCCACCGGACA
CCTCAGACACGCTTCTGCAGCTGTGCCTCGGCTCACAACACAGATTGACTGCTCTGACTTT
GACTACTCAAAATTGGCCTAAAAATTAAAAGAGCTCGATATTAAAAA

FIGURA 5

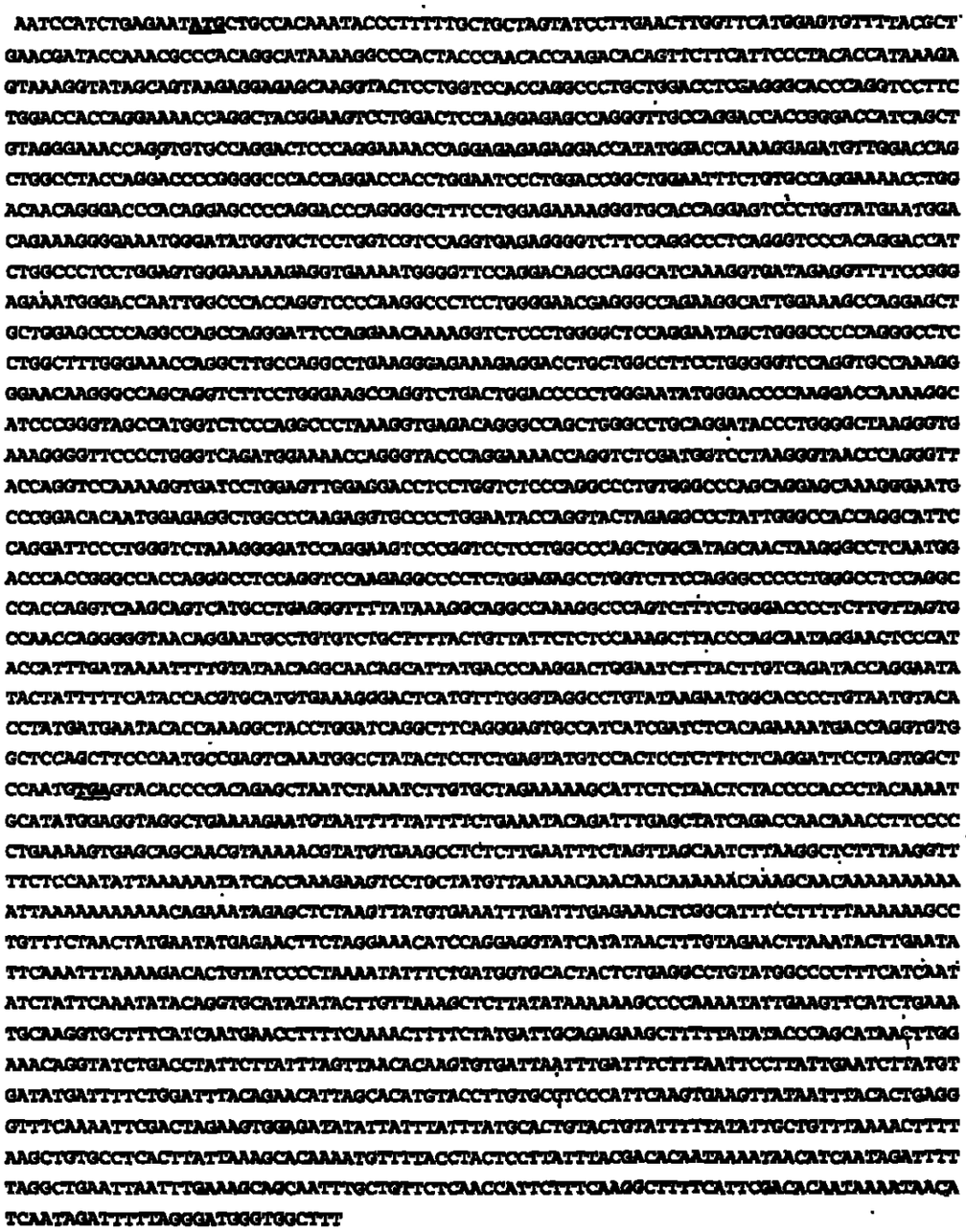


FIGURA 6

CATGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCC
TGTGTCTCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGA
ATGTCCTACAATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCA
GTATTTCATATATGGGCAAAAGAAATGGCTGAATAAATCAGAATGCAGAAATATCAATAGA
ACCTACTGTGATCTTTCTGCTGAAACTTCTGACTACGAACACCAGTATTATGCCAAAGTTA
AGGCCATTTGGGGAACAAAGTGTTCCAAATGGGCTGAAAGTGGACGGTTCTATCCTTTTTT
AGAAACACAAATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTT
GTCCTGACAGCTCCAGAGAAGTGGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAAC
AAATATACTCCAATCTGAAGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTG
GTCCCAGTGTGTGACCAACCACACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTAC
TGCGTACACGTGGAGTCCTTCGTCCCAGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGC
AGTGTGCCAGGACTTTGAAAGATCAATCATCAGAGTTCAAGGCTAAAATCATCTTCTGGTA
TGTTTTGCCCATATCTATTACCGTGTTTCTTTTTTCTGTGATGGGCTATTCCATCTACCGA
TATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTTGATTTTGATTTATGGAAATGAAT
TTGACAAAAGATTCTTTGTGCCTGCTGAAAAAATCGTGATTAACTTTATCACCCTCAATAT
CTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGAAAAAGCAGTGATGTA
TCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGAGGAAGAGGAGG
TGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAGAAAACAC
GGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAGCAGAACAATACCCCCGGATAAAACA
GTCATTGAATATGAATATGATGTCAGAACCACTGACATTTGTGCGGGGCCTGAAGAGCAGG
AGCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTT
GGCAGTCTTGGGCCCGCAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGAC
CCCCTGGCGCAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCC
TGGTCGACTGGGATCCCCAAACTGGCAGGCTGTGTATTCCTTCGCTGTCCAGCTTCGACCA
GGATTCAGAGGGCTGCGAGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCT
AGACTCTATGAGGAGCCGGCTCCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGC
AATTCATGGAGGAATGGGGGTTATATGTGCAGATGGAAAACTGATGCCAACACTTCCTTTT
GCCTTTTGTTTCCTGTGCAAACAAGTGAGTCACCCCTTTGATCCCAGCCATAAAGTACCTG
GGATGAAAGAAGTTTTTTCCAGTTTGTCAGTGTCTGTGAGAA

FIGURA 7

AGCGGGCGTGGGACTGAGCAGTCTGCTGCCCCCCGACATGTGACCCAGCCCCGCCGCCCAT
GCGGGCTCCCGGCCGCCCGGCCCTGCGGCCGCTGCCGCTGCCGCCGCTGCTGCTGTTGCTC
CTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGTGTCTCTGGTGGTTTGCCTAAACCTGCAA
ACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACAATGGACTCCACCAGAGGGTCT
TCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTCATATATGGGCAAAAGAAATGGCTG
AATAAATCAGAATGCAGAAATATCAATAGAACCTACTGTGATCTTTCTGCTGAAACTTCTG
ACTACGAACACCAGTATTATGCCAAAGTTAAGGCCATTTGGGGAACAAAGTGTTCCAAATG
GGCTGAAAGTGGACGGTTCTATCCTTTTTTAGAAACACAAATTGGCCCACCAGAGGTGGCA
CTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCCAGAGAAGTGGAAGAGAA
ATCCAGAAGACCTTCCTGTTTCCATGCAACAAATATACTCCAATCTGAAGTATAACGTGTC
TGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCACACGCTGGTG
CTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCTTCGTCCCAGGGC
CCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCATC
AGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTTCTT
TTTTCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAG
CAAATTTGATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAA
AATCGTGATTAACTTTATCACCCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGAT
ATGAGTTTACTGGGAAAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGA
ACCTGAGGCCCCCTCAGGAGGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGAT
GGAAATTTTTTGTGACTCTGAAGAAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCC
CTCAGCAGAACAATACCCCCGGATAAAACAGTCATTGAATATGAATATGATGTCAGAACCA
CTGACATTTGTGCGGGGCCTGAAGAGCAGGAGCTCAGTTTGCAGGAGGAGGTGTCCACACA
AGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAGTCTTGGGCCCGCAAACGTTACAGTAC
TCATACACCCCTCAGCTCCAAGACTTAGACCCCCTGGCGCAGGAGCACACAGACTCGGAGG
AGGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGATCCCCAAACTGGCAGGCT
GTGTATTCCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCGAGCCTTCTGAGGGG
GATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTGTATGAGGAGCCGGCTCCAGACAGGC
CACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGTTATATGTGCA
GATGGAAAACTGATGCCAACACTTCCTTTTGCCTTTTGTTTCCTGTGCAAACAAGTGAGTC
ACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTTCCAGTTTGTCAGT
GTCTGTGAGAA

FIGURA 8

GCTTGGTCTAGTGCCCATTTACTCTGGACTCCGGATGGCTGCCGCACGCCTCTGCCTCTCC
CTGCTGCTCCTGTCCACCTGCGTGGCTCTGTTACTACAGCCACTGCTGGGTGCCCAGGGAG
CCCCACTGGAGCCAGTGTACCCAGGGGACAATGCCACACCAGAGCAGATGGCCCAGTATGC
AGCTGATCTCCGTAGATACATCAACATGCTGACCAGGCCTAGGTATGGGAAAAGACACAAA
GAGGACACGCTGGCCTTCTCGGAGTGGGGGTCCCCGCATGCTGCTGTCCCCAGGGAGCTCA
GCCCGCTGGACTTATAATGCCACCTTCTGTCTCCTACGACTCCATGAGCAGCGCCAGCCCA
GCTCTCCCCTCTGCACCCTTGGCTCTGGCCAAAGCTTGCTCCCTGCTCCCACACAGGCTCA
ATAAAGCAAGTCAAAGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURA 9

GAAGGACAGCAGGGCCAACAGTCACAGCAGCCCTGACCAGAGCATTCCTGGAGCTCAAGCT
CCTCTACAAAGAGGTGGACAGAGAAGACAGCAGAGACCATGGGACCCCCCTCAGCCCCTCC
CTGCAGATTGCATGTCCCCTGGAAGGAGGTCCTGCTCACAGCCTCACTTCTAACCTTCTGG
AACCCACCCACCACTGCCAAGCTCACTATTGAATCCACGCCGTTCAATGTCGCAGAGGGGA
AGGAGGTTCTTCTACTCGCCCACAACCTGCCCCAGAATCGTATTGGTTACAGCTGGTACAA
AGGCGAAAGAGTGGATGGCAACAGTCTAATTGTAGGATATGTAATAGGAACTCAACAAGCT
ACCCCAGGGCCCGCATACAGTGGTCGAGAGACAATATACCCCAATGCATCCCTGCTGATCC
AGAACGTCACCCAGAATGACACAGGATTCTATACCCTACAAGTCATAAAGTCAGATCTTGT
GAATGAAGAAGCAACCGGACAGTTCCATGTATACCCGGAGCTGCCCAAGCCCTCCATCTCC
AGCAACAACTCCAACCCCGTGGAGGACAAGGATGCTGTGGCCTTCACCTGTGAACCTGAGG
TTCAGAACACAACCTACCTGTGGTGGGTAAATGGTCAGAGCCTCCCGGTCAGTCCCAGGCT
GCAGCTGTCCAATGGCAACATGACCCTCACTCTACTCAGCGTCAAAAGGAACGATGCAGGA
TCCTATGAATGTGAAATACAGAACCCAGCGAGTGCCAACCGCAGTGACCCAGTCACCCTGA
ATGTCCTCTATGGCCCAGATGGCCCCACCATTTCCCCCTCAAAGGCCAATTACCGTCCAGG
GGAAAATCTGAACCTCTCCTGCCACGCAGCCTCTAACCCACCTGCACAGTACTCTTGGTTT
ATCAATGGGACGTTCCAGCAATCCACACAAGAGCTCTTTATCCCCAACATCACTGTGAATA
ATAGCGGATCCTATATGTGCCAAGCCCATAACTCAGCCACTGGCCTCAATAGGACCACAGT
CACGATGATCACAGTCTCTGGAAGTGCTCCTGTCCTCTCAGCTGTGGCCACCGTCGGCATC
ACGATTGGAGTGCTGGCCAGGGTGGCTCTGATATAGCAGCCCTGGTGTATTTTCGATATTT
CAGGAAGACTGGCAGATTGGACCAGACCCTGAATTCTTCTAGCTCCTCCAATCCCATTTTA
TCCCATGGAACCACTAAAAACAAGGTCTGCTCTGCTCCTGAAGCCCTATATGCTGGAGATG
GACAACTCAATGAAAATTTAAAGGGAAAACCCTCAGGCCTGAGGTGTGTGCCACTCAGAGA
CTTCACCTAACTAGAGACAGGCAAACTGCAAACCATGGTGAGAAATTGACGACTTCACACT
ATGGACAGCTTTTCCCAAGATGTCAAAACAAGACTCCTCATCATGATAAGGCTCTTACCCC
CTTTTAATTTGTCCTTGCTTATGCCTGCCTCTTTCGCTTGGCAGGATGATGCTGTCATTAG
TATTTCACAAGAAGTAGCTTCAGAGGGTAACTTAACAGAGTATCAGATCTATCTTGTCAAT
CCCAACGTTTTACATAAAATAAGAGATCCTTTAGTGCACCCAGTGACTGACATTAGCAGCA
TCTTTAACACAGCCGTGTGTTCAAATGTACAGTGGTCCTTTTCAGAGTTGGACTTCTAGAC
TCACCTGTTCTCACTCCCTGTTTTAATTCAACCCAGCCATGCAATGCCAAATAATAGAATT
GCTCCCTACCAGCTGAACAGGGAGGAGTCTGTGCAGTTTCTGACACTTGTTGTTGAACATG
GCTAAATACAATGGGTATCGCTGAGACTAAGTTGTAGAAATTAACAAATGTGCTGCTTGGT
TAAAATGGCTACACTCATCTGACTCATTCTTTATTCTATTTTAGTTGGNTTGTATCTTGCC
TAAGGTGCGTAGTCCAACTCTTGGTATTACCCTCCTAATAGTCATACTAGTAGTCATACTC
CCTGGTGTAGTGTATTCTCTAAAAGCTTTAAATGTCTGCATGCAGCCAGCCATCAAATAGT
GAATGGGTCTCTCTTTTGGCTGGAATTACAAAACTCAGAGAAATGTGTCATCAGGAGAACATC

ATAACCCATGAAGGATAAAAGCCCCAAATGGTGGTAACTGATAATAGCACTAATGCTTTAA
GATTTGGTCACACTCTCACCTAGGTGAGCGCATTGAGCCAGTGGTGCATAAATGCTACATA
CTCCAACTGAAATGTTAAGGAAGAAGATAGATCCAATTAAAAAAAATTAAAACCAATTTAA
AAAAAAAAAGAACACAGGAGATTCCAGTCTACTTGAGTTAGCATAATACAGAAGTCCCCTC
TACTTTAACTTTTACAAAAAAGTAACCTGAACTAATCTGATGTTAACCAATGTATTTATTT
CTGTGGTTCTGTTTCCTTGTTCCAATTTGACAAAACCCACTGTTCTTGTATTGTATTGCCC
AGGGGGAGCTATCACTGTACTTGTAGAGTGGTGCTGCTTTAATTCATAAATCACAAATAAA
AGCCAATTAGCTCTAAAAAAAAAAAA

FIGURA 10

GACCTGGGGGGGCGAGGGCCCCCCAAACTCAGTTCGGATCCTACCCGAGTGAGGCGGCGCC
ATGGAGCTCCGGGTGCTGCTCTGCTGGGCTTCGTTGGCCGCAGCTTTGGAAGAGACCCTGC
TGAACACAAAATTGGAAACTGCTGATCTGAAGTGGGTGACATTCCCTCAGGTGGACGGGCA
GTGGGAGGAACTGAGCGGCCTGGATGAGGAACAGCACAGCGTGCGCACCTACGAAGTGTGT
GACGTGCAGCGTGCCCCGGGCCAGGCCCACTGGCTTCGCACAGGTTGGGTCCCACGGCGGG
GCGCCGTCCACGTGTACGCCACGCTGCGCTTCACCATGCTCGAGTGCCTGTCCCTGCCTCG
GGCTGGGCGCTCCTGCAAGGAGACCTTCACCGTCTTCTACTATGAGAGCGATGCGGACACG
GCCACGGCCCTCACGCCAGCCTGGATGGAGAACCCCTACATCAAGGTGGACACGGTGGCCG
CGGAGCATCTCACCCGGAAGCGCCCTGGGGCCGAGGCCACCGGGAAGGTGAATGTCAAGAC
GCTGCGTCTGGGACCGCTCAGCAAGGCTGGCTTCTACCTGGCCTTCCAGGACCAGGGTGCC
TGCATGGCCCTGCTATCCCTGCACCTCTTCTACAAAAAGTGCGCCCAGCTGACTGTGAACC
TGACTCGATTCCCGGAGACTGTGCCTCGGGAGCTGGTTGTGCCCGTGGCCGGTAGCTGCGT
GGTGGATGCCGTCCCCGCCCCTGGCCCCAGCCCCAGCCTCTACTGCCGTGAGGATGGCCAG
TGGGCCGAACAGCCGGTCACGGGCTGCAGCTGTGCTCCGGGGTTCGAGGCAGCTGAGGGGA
ACACCAAGTGCCGAGCCTGTGCCCAGGGCACCTTCAAGCCCCTGTCAGGAGAAGGGTCCTG
CCAGCCATGCCCAGCCAATAGCCACTCTAACACCATTGGATCAGCCGTCTGCCAGTGCCGC
GTCGGGTACTTCCGGGCACGCACAGACCCCCGGGGTGCACCCTGCACCACCCCTCCTTCGG
CTCCGCGGAGCGTGGTTTCCCGCCTGAACGGCTCCTCCCTGCACCTGGAATGGAGTGCCCC
CCTGGAGTCTGGTGGCCGAGAGGACCTCACCTACGCCCTCCGCTGCCGGGAGTGCCGACCC
GGAGGCTCCTGTGCGCCCTGCGGGGGAGACCTGACTTTTGACCCCGGCCCCCGGGACCTGG
TGGAGCCCTGGGTGGTGGTTCGAGGGCTACGTCCTGACTTCACCTATACCTTTGAGGTCAC
TGCATTGAACGGGGTATCCTCCTTAGCCACGGGGCCCGTCCCATTTGAGCCTGTCAATGTC
ACCACTGACCGAGAGGTACCTCCTGCAGTGTCTGACATCCGGGTGACGCGGTCCTCACCCA
GCAGCTTGAGCCTGGCCTGGGCTGTTCCCCGGGCACCCAGTGGGGCTGTGCTGGACTACGA
GGTCAAATACCATGAGAAGGGCGCCGAGGGTCCCAGCAGCGTGCGGTTCCTGAAGACGTCA
GAAAACCGGGCAGAGCTGCGGGGGCTGAAGCGGGGAGCCAGCTACCTGGTGCAGGTACGGG
CGCGCTCTGAGGCCGGCTACGGGCCCTTCGGCCAGGAACATCACAGCCAGACCCAACTGGA
TGAGAGCGAGGGCTGGCGGGAGCAGCTGGCCCTGATTGCGGGCACGGCAGTCGTGGGTGTG
GTCCTGGTCCTGGTGGTCATTGTGGTCGCAGTTCTCTGCCTCAGGAAGCAGAGCAATGGGA
GAGAAGCAGAATATTCGGACAAACACGGACAGTATCTCATCGGACATGGTACTAAGGTCTA
CATCGACCCCTTCACTTATGAAGACCCTAATGAGGCTGTGAGGGAATTTGCAAAAGAGATC
GATGTCTCCTACGTCAAGATTGAAGAGGTGATTGGTGCAGGTGAGTTTGGCGAGGTGTGTC
GGGGGCGGCTCAAGGCCCCAGGGAAGAAGGAGAGCTGTGTGGCAATCAAGACCCTGAAGGG
TGGCTACACGGAGCGGCAGCGGCGTGAGTTTCTGAGCGAGGCCTCCATCATGGGCCAGTTC
GAGCACCCCAATATCATCCGCCTGGAGGGCGTGGTCACCAACAGCATGCCCGTCATGATTC

TCACAGAGTTCATGGAGAACGGCGCCCTGGACTCCTTCCTGCGGCTAAACGACGGACAGTT
CACAGTCATCCAGCTCGTGGGCATGCTGCGGGGCATCGCCTCGGGCATGCGGTACCTTGCC
GAGATGAGCTACGTCCACCGAGACCTGGCTGCTCGCAACATCCTAGTCAACAGCAACCTCG
TCTGCAAAGTGTCTGACTTTGGCCTTTCCCGATTCCTGGAGGAGAACTCTTCCGATCCCAC
CTACACGAGCTCCCTGGGAGGAAAGATTCCCATCCGATGGACTGCCCCGGAGGCCATTGCC
TTCCGGAAGTTCACTTCCGCCAGTGATGCCTGGAGTTACGGGATTGTGATGTGGGAGGTGA
TGTCATTTGGGGAGAGGCCGTACTGGGACATGAGCAATCAGGACGTGATCAATGCCATTGA
ACAGGACTACCGGCTGCCCCCGCCCCCAGACTGTCCCACCTCCCTCCACCAGCTCATGCTG
GACTGTTGGCAGAAAGACCGGAATGCCCGGCCCCGCTTCCCCCAGGTGGTCAGCGCCCTGG
ACAAGATGATCCGGAACCCCGCCAGCCTCAAAATCGTGGCCCGGGAGAATGGCGGGGCCTC
ACACCCTCTCCTGGACCAGCGGCAGCCTCACTACTCAGCTTTTGGCTCTGTGGGCGAGTGG
CTTCGGGCCATCAAAATGGGAAGATACGAAGAAAGTTTCGCAGCCGCTGGCTTTGGCTCCT
TCGAGCTGGTCAGCCAGATCTCTGCTGAGGACCTGCTCCGAATCGGAGTCACTCTGGCGGG
ACACCAGAAGAAAATCTTGGCCAGTGTCCAGCACATGAAGTCCCAGGCCAAGCCGGGAACC
CCGGGTGGGACAGGAGGACCGGCCCCGCAGTACTGACCTGCAGGAACTCCCCACCCCAGGG
ACACCGCCTCCCCATTTTCCGGGGCAGAGTGGGGACTCACAGAGGCCCCCAGCCCTGTGCC
CCGCTGGATTGCACTTTGAGCCCGTGGGGTGAGGAGTTGGCAATTTGGAGAGACAGGATTT
GGGGGTTCTGCCATAATAGGAGGGGAAAATCACCCCCCAGCCACCTCGGGGAACTCCAGAC
CAAGGGTGAGGGCGCCTTTCCCTCAGGACTGGGTGTGACCAGAGGAAAAGGAAGTGCCCAA
CATCTCCCAGCCTCCCCAGGTGCCCCCCTCACCTTGATGGGTGCGTTCCCGCAGACCAAAG
AGAGTGTGACTCCCTTGCCAGCTCCAGAGTGGGGGGGCTGTCCCAGGGGGCAAGAAGGGGT
GTCAGGGCCCAGTGACAAAATCATTGGGGTTTGTAGTCCCAACTTGCTGCTGTCACCACCA
AACTCAATCATTTTTTTCCCTTGTAAATGCCCCTCCCCCAGCTGCTGCCTTCATATTGAAG
GTTTTTGAGTTTTGTTTTTGGTCTTAATTTTTCTCCCCGTTCCCTTTTTGTTTCTTCGTTT
TGTTTTTCTACCGTCCTTGTCATAACTTTGTGTTGGAGGGAACCTGTTTCACTATGGCCTC
CTTTGCCCAAGTTGAAACAGGGGCCCATCATCATGTCTGTTTCCAGAACAGTGCCTTGGTC
ATCCCACATCCCCGGACCCCGCCTGGGACCCCCAAGCTGTGTCCTATGAAGGGGTGTGGGG
TGAGGTAGTGAAAAGGGCGGTAGTTGGTGGTGGAACCCAGAAACGGACGCCGGTGCTTGGA
GGGGTTCTTAAATTATATTTAAAAAAGTAACTTTTTGTATAAATAAAAGAAAATGGGACGT
GTCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGA

FIGURA 11

GCTCTATGCCGCCTACCTTGCTCTCGCCGCTGCTGCCGGAGCCGAAGCAGAGAAGGCAGCG
GGTCCCGTGACCGTCCCGAGAGCCCCGCGCTCCCGACCAGGGGGCGGGGGCGGCCCCGGGG
AGGGCGGGGCAGGGGCGGGGGGAAGAAAGGGGGTTTTGTGCTGCGCCGGGAGGGCCGGCGC
CCTCTTCCGAATGTCCTGCGGCCCCAGCCTCTCCTCACGCTCGCGCAGTCTCCGCCGCAGT
CTCAGCTGCAGCTGCAGGACTGAGCCGTGCACCCGGAGGAGACCCCCGGAGGAGGCGACAA
ACTTCGCAGTGCCGCGACCCAACCCCAGCCCTGGGTAGCCTGCAGCATGGCCCAGCTGTTC
CTGCCCCTGCTGGCAGCCCTGGTCCTGGCCCAGGCTCCTGCAGCTTTAGCAGATGTTCTGG
AAGGAGACAGCTCAGAGGACCGCGCTTTTCGCGTGCGCATCGCGGGCGACGCGCCACTGCA
GGGCGTGCTCGGCGGCGCCCTCACCATCCCTTGCCACGTCCACTACCTGCGGCCACCGCCG
AGCCGCCGGGCTGTGCTGGGCTCTCCGCGGGTCAAGTGGACTTTCCTGTCCCGGGGCCGGG
AGGCAGAGGTGCTGGTGGCGCGGGGAGTGCGCGTCAAGGTGAACGAGGCCTACCGGTTCCG
CGTGGCACTGCCTGCGTACCCAGCGTCGCTCACCGACGTCTCCCTGGCGCTGAGCGAGCTG
CGCCCCAACGACTCAGGTATCTATCGCTGTGAGGTCCAGCACGGCATCGATGACAGCAGCG
ACGCTGTGGAGGTCAAGGTCAAAGGGGTCGTCTTTCTCTACCGAGAGGGCTCTGCCCGCTA
TGCTTTCTCCTTTTCTGGGGCCCAGGAGGCCTGTGCCCGCATTGGAGCCCACATCGCCACC
CCGGAGCAGCTCTATGCCGCCTACCTTGGGGGCTATGAGCAATGTGATGCTGGCTGGCTGT
CGGATCAGACCGTGAGGTATCCCATCCAGACCCCACGAGAGGCCTGTTACGGAGACATGGA
TGGCTTCCCCGGGGTCCGGAACTATGGTGTGGTGGACCCGGATGACCTCTATGATGTGTAC
TGTTATGCTGAAGACCTAAATGGAGAACTGTTCCTGGGTGACCCTCCAGAGAAGCTGACAT
TGGAGGAAGCACGGGCGTACTGCCAGGAGCGGGGTGCAGAGATTGCCACCACGGGCCAACT
GTATGCAGCCTGGGATGGTGGCCTGGACCACTGCAGCCCAGGGTGGCTAGCTGATGGCAGT
GTGCGCTACCCCATCGTCACACCCAGCCAGCGCTGTGGTGGGGGCTTGCCTGGTGTCAAGA
CTCTCTTCCTCTTCCCCAACCAGACTGGCTTCCCCAATAAGCACAGCCGCTTCAACGTCTA
CTGCTTCCGAGACTCGGCCCAGCCTTCTGCCATCCCTGAGGCCTCCAACCCAGCCTCCAAC
CCAGCCTCTGATGGACTAGAGGCTATCGTCACAGTGACAGAGACCCTGGAGGAACTGCAGC
TGCCTCAGGAAGCCACAGAGAGTGAATCCCGTGGGGCCATCTACTCCATCCCCATCATGGA
GGACGGAGGAGGTGGAAGCTCCACTCCAGAAGACCCAGCAGAGGCCCCTAGGACGCTCCTA
GAATTTGAAACACAATCCATGGTACCGCCCACGGGGTTCTCAGAAGAGGAAGGTAAGGCAT
TGGAGGAAGAAGAGAAATATGAAGATGAAGAAGAGAAAGAGGAGGAAGAAGAAGAGGAGGA
GGTGGAGGATGAGGCTCTGTGGGCATGGCCCAGCGAGCTCAGCAGCCCGGGCCCTGAGGCC
TCTCTCCCCACTGAGCCAGCAGCCCAGGAGAAGTCACTCTCCCAGGCGCCAGCAAGGGCAG
TCCTGCAGCCTGGTGCATCACCACTTCCTGATGGAGAGTCAGAAGCTTCCAGGCCTCCAAG
GGTCCATGGACCACCTACTGAGACTCTGCCCACTCCCAGGGAGAGGAACCTAGCATCCCCA
TCACCTTCCACTCTGGTTGAGGCAAGAGAGGTGGGGGAGGCAACTGGTGGTCCTGAGCTAT
CTGGGGTCCCTCGAGGAGAGAGCGAGGAGACAGGAAGCTCCGAGGGTGCCCCTTCCCTGCT

TCCAGCCACACGGGCCCCTGAGGGTACCAGGGAGCTGGAGGCCCCCTCTGAAGATAATTCT
GGAAGAACTGCCCCAGCAGGGACCTCAGTGCAGGCCCAGCCAGTGCTGCCCACTGACAGCG
CCAGCCGAGGTGGAGTGGCCGTGGTCCCCGCATCAGGTGACTGTGTCCCCAGCCCCTGCCA
CAATGGTGGGACATGCTTGGAGGAGGAGGAAGGGGTCCGCTGCCTATGTCTGCCTGGCTAT
GGGGGGGACCTGTGCGATGTTGGCCTCCGCTTCTGCAACCCCGGCTGGGACGCCTTCCAGG
GCGCCTGCTACAAGCACTTTTCCACACGAAGGAGCTGGGAGGAGGCAGAGACCCAGTGCCG
GATGTACGGCGCGCATCTGGCCAGCATCAGCACACCCGAGGAACAGGACTTCATCAACAAC
CGGTACCGGGAGTACCAGTGGATCGGACTCAACGACAGGACCATCGAAGGCGACTTCTTGT
GGTCGGATGGCGTCCCCCTGCTCTATGAGAACTGGAACCCTGGGCAGCCTGACAGCTACTT
CCTGTCTGGAGAGAACTGCGTGGTCATGGTGTGGCATGATCAGGGACAATGGAGTGACGTG
CCCTGCAACTACCACCTGTCCTACACCTGCAAGATGGGGCTGGTGTCCTGTGGGCCGCCAC
CGGAGCTGCCCCTGGCTCAAGTGTTCGGCCGCCCACGGCTGCGCTATGAGGTGGACACTGT
GCTTCGCTACCGGTGCCGGGAAGGACTGGCCCAGCGCAATCTGCCGCTGATCCGATGCCAA
GAGAACGGTCGTTGGGAGGCCCCCCAGATCTCCTGTGTGCCCAGAAGACCTGCCCGAGCTC
TGCACCCAGAGGAGGACCCAGAAGGACGTCAGGGGAGGCTACTGGGACGCTGGAAGGCGCT
GTTGATCCCCCCTTCCAGCCCCATGCCAGGTCCCTAGGGGGCAAGGCCTTGAACACTGCCG
GCCACAGCACTGCCCTGTCACCCAAATTTTCCCTCACACCTTGCGCTCCCGCCACCACAGG
AAGTGACAACATGACGAGGGGTGGTGCTGGAGTCCAGGTGACAGTTCCTGAAGGGGCTTCT
GGGAAATACCTAGGAGGCTCCAGCCCAGCCCAGGCCCTCTCCCCCTACCCTGGGCACCAGA
TCTTCCATCAGGGCCGGAGTAAATCCCTAAGTGCCTCAACTGCCCTCTCCCTGGCAGCCAT
CTTGTCCCCTCTATTCCTCTAGGGAGCACTGTGCCCACTCTTTCTGGGTTTTCCAAGGGAA
TGGGCTTGCAGGATGGAGTGTCTGTAAAATCAACAGGAAATAAAACTGTGTATGAGCCCA

FIGURA 12

TGTGGCACTGCCTGCGTACCCAACCCCAGCCCTGGGTAGCCTGCAGCATGGCCCAGCTGTT
CCTGCCCCTGCTGGCAGCCCTGGTCCTGGCCCAGGCTCCTGCAGCTTTAGCAGATGTTCTG
GAAGGAGACAGCTCAGAGGACCGCGCTTTTCGCGTGCGCATCGCGGGCGACGCGCCACTGC
AGGGCGTGCTCGGCGGCGCCCTCACCATCCCTTGCCACGTCCACTACCTGCGGCCACCGCC
GAGCCGCCGGGCTGTGCTGGGCTCTCCGCGGGTCAAGTGGACTTTCCTGTCCCGGGGCCGG
GAGGCAGAGGTGCTGGTGGCGCGGGGAGTGCGCGTCAAGGTGAACGAGGCCTACCGGTTCC
GCGTGGCACTGCCTGCGTACCCAGCGTCGCTCACCGACGTCTCCCTGGCGCTGAGCGAGCT
GCGCCCCAACGACTCAGGTATCTATCGCTGTGAGGTCCAGCACGGCATCGATGACAGCAGC
GACGCTGTGGAGGTCAAGGTCAAAGGGGTCGTCTTTCTCTACCGAGAGGGCTCTGCCCGCT
ATGCTTTCTCCTTTTCTGGGGCCCAGGAGGCCTGTGCCCGCATTGGAGCCCACATCGCCAC
CCCGGAGCAGCTCTATGCCGCCTACCTTGGGGGCTATGAGCAATGTGATGCTGGCTGGCTG
TCGGATCAGACCGTGAGGTATCCCATCCAGACCCCACGAGAGGCCTGTTACGGAGACATGG
ATGGCTTCCCCGGGGTCCGGAACTATGGTGTGGTGGACCCGGATGACCTCTATGATGTGTA
CTGTTATGCTGAAGACCTAAATGGAGAATTGTTCCTGGGTGACCCTCCAGAGAAGCTGACA
TTGGAGGAAGCACGGGCGTACTGCCAGGAGCGGGGTGCAGAGATTGCCACCACGGGCCAAC
TGTATGCAGCCTGGGATGGTGGCCTGGACCACTGCAGCCCAGGGTGGCTAGCTGATGGCAG
TGTGCGCTACCCCATCGTCACACCCAGCCAGCGCTGTGGTGGGGGCTTGCCTGGTGTCAAG
ACTCTCTTCCTCTTCCCCAACCAGACTGGCTTCCCCAATAAGCACAGCCGCTTCAACGTCT
ACTGCTTCCGAGACTCGGCCCAGCCTTCTGCCATCCCTGAGGCCTCCAACCCAGCCTCCAA
CCCAGCCTCTGATGGACTAGAGGCTATCGTCACAGTGACAGAGACCCTGGAGGAACTGCAG
CTGCCTCAGGAAGCCACAGAGAGTGAATCCCGTGGGGCCATCTACTCCATCCCCATCATGG
AGGACGGAGGAGGTGGAAGCTCCACTCCAGAAGACCCAGCAGAGGCCCCTAGGACGCTCCT
AGAATTTGAAACACAATCCATGGTACCGCCCACGGGGTTCTCAGAAGAGGAAGGTAAGGCA
TTGGAGGAAGAAGAGAAATATGAAGATGAAGAAGAGAAAGAGGAGGAAGAAGAAGAGGAGG
AGGTGGAGGATGAGGCTCTGTGGGCATGGCCCAGCGAGCTCAGCAGCCCGGGCCCTGAGGC
CTCTCTCCCCACTGAGCCAGCAGCCCAGGAGGAGTCACTCTCCCAGGCGCCAGCAAGGGCA
GTCCTGCAGCCTGGTGCATCACCACTTCCTGATGGAGAGTCAGAAGCTTCCAGGCCTCCAA
GGGTCCATGGACCACCTACTGAGACTCTGCCCACTCCCAGGGAGAGGAACCTAGCATCCCC
ATCACCTTCCACTCTGGTTGAGGCAAGAGAGGTGGGGGAGGCAACTGGTGGTCCTGAGCTA
TCTGGGGTCCCTCGAGGAGAGAGCGAGGAGACAGGAAGCTCCGAGGGTGCCCCTTCCCTGC
TTCCAGCCACACGGGCCCCTGAGGGTACCAGGGAGCTGGAGGCCCCCTCTGAAGATAATTC
TGGAAGAACTGCCCCAGCAGGGACCTCAGTGCAGGCCCAGCCAGTGCTGCCCACTGACAGC
GCCAGCCGAGGTGGAGTGGCCGTGGTCCCCGCATCAGGTAATTCTGCCCAAGGCTCAACTG
CCCTCTCTATCCTACTCCTTTTCTTCCCCCTGCAGCTCTGGGTCACCTGACCTGTAGTCCT

CTATGGGTCTCCAATCTCGGATATCCACCTTGTGGGTATCTCAGCTCTCCGCGTCTTTACC
CTGTGATCCCAGCCCCGCCACTGACCATCTGTGACCCTTCCCTGCCATTGGGCCCTCCACC
TGTGGCTCACATCTCGCCAGCCCCACAGAGCATCCTCAGGCCTCTCCAAGGGTCCTCATCA
CCTATTGCAGCCTTCAGGGCTCGGCCTATTTTCCACTACTCCCTTCATCCGCCTGTGTGCC
GTCCCCTTTAGCTGCCTCCTATTGATCTCAGGGAAGCCTGGGAGTCCCTTCTCACCCCTCA
ACCTCCGGAGTCCAGGAGAACCCGTACCCCCACAGAGCCTTAAGCAACTACTTCTGTGAAG
TATTTTTTGACTGTTTCATGGAAAACAAGCCTTGGAAATAAATCTCTATTAAACCGC

FIGURA 13

AGCCAGCTCAGGCTACACTATCCCAGGATCAGCATGGCCGTCCGCCAGTGGGTAATCGCCC
TGGCCTTGGCTGCCCTCCTTGTTGTGGACAGGGAAGTGCCAGTGGCAGCAGGAAAGCTCCC
TTTCTCAAGAATGCCCATCTGTGAACACATGGTAGAGTCTCCAACCTGTTCCCAGATGTCC
AACCTGGTCTGCGGCACTGATGGGCTCACATATACGAATGAATGCCAGCTCTGCTTGGCCC
GGATAAAAACCAAACAGGACATCCAGATCATGAAAGATGGCAAATGCTGATCCCACAGGAG
CACCTCAAGCCATGAAGTGTCAGCTGGAGAACAGTGGTGGGCATGGAGAGGATATGACATG
AAATAAAAGATCCAGCCCAAAAAAAAAAGAAAAAA

FIGURA 14

CGTTTCCAATGCACGTACAGCCCGTACACACCGTGTGCTGGGACACCCCACAGTCAGCCGC
ATGGCTCCCCTGTGCCCCAGCCCCTGGCTCCCTCTGTTGATCCCGGCCCCTGCTCCAGGCC
TCACTGTGCAACTGCTGCTGTCACTGCTGCTTCTGGTGCCTGTCCATCCCCAGAGGTTGCC
CCGGATGCAGGAGGATTCCCCCCTGGGAGGAGGCTCTTCTGGGGAAGATGACCCACTGGGC
GAGGAGGATCTGCCCAGTGAAGAGGATTCACCCAGAGAGGAGGATCCACCCGGAGAGGAGG
ATCTACCTGGAGAGGAGGATCTACCTGGAGAGGAGGATCTACCTGAAGTTAAGCCTAAATC
AGAAGAAGAGGGCTCCCTGAAGTTAGAGGATCTACCTACTGTTGAGGCTCCTGGAGATCCT
CAAGAACCCCAGAATAATGCCCACAGGGACAAAGAAGGGGATGACCAGAGTCATTGGCGCT
ATGGAGGCGACCCGCCCTGGCCCCGGGTGTCCCCAGCCTGCGCGGGCCGCTTCCAGTCCCC
GGTGGATATCCGCCCCCAGCTCGCCGCCTTCTGCCCGGCCCTGCGCCCCCTGGAACTCCTG
GGCTTCCAGCTCCCGCCGCTCCCAGAACTGCGCCTGCGCAACAATGGCCACAGTGTGCAAC
TGACCCTGCCTCCTGGGCTAGAGATGGCTCTGGGTCCCGGGCGGGAGTACCGGGCTCTGCA
GCTGCATCTGCACTGGGGGGCTGCAGGTCGTCCGGGCTCGGAGCACACTGTGGAAGGCCAC
CGTTTCCCTGCCGAGATCCACGTGGTTCACCTCAGCACCGCCTTTGCCAGAGTTGACGAGG
CCTTGGGGCGCCCGGGAGGCCTGGCCGTGTTGGCCGCCTTTCTGGAGGAGGGCCCGGAAGA
AAACAGTGCCTATGAGCAGTTGCTGTCTCGCTTGGAAGAAATCGCTGAGGAAGGCTCAGAG
ACTCAGGTCCCAGGACTGGACATATCTGCACTCCTGCCCTCTGACTTCAGCCGCTACTTCC
GATATGAGGGGTCTCTGACTACACCGCCCTGTGCCCAGGGTGTCATCTGGACTGTGTTTAA
CCAGACAGTGATGCTGAGTGCTAAGCAGCTCCACACCCTCTCTGACACCCTGTGGGGACCT
GGTGACTCTCGGCTACAGCTGAACTTCCGAGCGACGCAGCCTTTGAATGGGCGAGTGATTG
AGGCCTCCTTCCCTGCTGGAGTGGACAGCAGTCCTCGGGCTGCTGAGCCAGTCCAGCTGAA
TTCCTGCCTGGCTGCTGGTGACATCCTAGCCCTGGTTTTTGGCCTCCTTTTTGCTGTCACC
AGCGTCGCGTTCCTTGTGCAGATGAGAAGGCAGCACAGAAGGGGAACCAAAGGGGGTGTGA
GCTACCGCCCAGCAGAGGTAGCCGAGACTGGAGCCTAGAGGCTGGATCTTGGAGAATGTGA
GAAGCCAGCCAGAGGCATCTGAGGGGGAGCCGGTAACTGTCCTGTCCTGCTCATTATGCCA
CTTCCTTTTAACTGCCAAGAAATTTTTTAAAATAAATATTTATAATA

FIGURA 15

TCCACATCCAGGTCTTTGTGCTCCTCGCTTGCCTGTTCCTTTTCCACGCATTTTCCAGGAT
AACTGTGACTCCAGGCCCGCAATGGATGCCCTGCAACTAGCAAATTCGGCTTTTGCCGTTG
ATCTGTTCAAACAACTATGTGAAAAGGAGCCACTGGGCAATGTCCTCTTCTCTCCAATCTG
TCTCTCCACCTCTCTGTCACTTGCTCAAGTGGGTGCTAAAGGTGACACTGCAAATGAAATT
GGACAGGTTCTTCATTTTGAAAATGTCAAAGATGTACCCTTTGGATTTCAAACAGTAACAT
CGGATGTAAACAAACTTAGTTCCTTTTACTCACTGAAACTAATCAAGCGGCTCTACGTAGA
CAAATCTCTGAATCTTTCTACAGAGTTCATCAGCTCTACGAAGAGACCGTATGCAAAGGAA
TTGGAAACTGTTGACTTCAAAGATAAATTGGAAGAAACGAAAGGTCAGATCAACAACTCAA
TTAAGGATCTCACAGATGGCCACTTTGAGAACATTTTAGCTGACAACAGTGTGAACGACCA
GACCAAAATCCTTGTGGTTAATGCTGCCTACTTTGTTGGCAAGTGGATGAAGAAATTTTCT
GAATCAGAAACAAAAGAATGTCCTTTCAGAGTCAACAAGACAGACACCAAACCAGTGCAGA
TGATGAACATGGAGGCCACGTTCTGTATGGGAAACATTGACAGTATCAATTGTAAGATCAT
AGAGCTTCCTTTTCAAAATAAGCATCTCAGCATGTTCATCCTACTACCCAAGGATGTGGAG
GATGAGTCCACAGGCTTGGAGAAGATTGAAAAACAACTCAACTCAGAGTCACTGTCACAGT
GGACTAATCCCAGCACCATGGCCAATGCCAAGGTCAAACTCTCCATTCCAAAATTTAAGGT
GGAAAAGATGATTGATCCCAAGGCTTGTCTGGAAAATCTAGGGCTGAAACATATCTTCAGT
GAAGACACATCTGATTTCTCTGGAATGTCAGAGACCAAGGGAGTGGCCCTATCAAATGTTA
TCCACAAAGTGTGCTTAGAAATAACTGAAGATGGTGGGGATTCCATAGAGGTGCCAGGAGC
ACGGATCCTGCAGCACAAGGATGAATTGAATGCTGACCATCCCTTTATTTACATCATCAGG
CACAACAAAACTCGAAACATCATTTTCTTTGGCAAATTCTGTTCTCCTTAAGTGGCATAGC
CCATGTTAAGTCCTCCCTGACTTTTCTGTGGATGCCGATTTCTGTAAACTCTGCATCCAGA
GATTCATTTTCTAGATACAATAAATTGCTAATGTTGCTGGATCAGGAAGCCGCCAGTACTT
GTCATATGTAGCCTTCACACAGATAGACCTTTTTTTTTTTTCCAATTCTATCTTTTGTTTC
CTTTTTTCCCATAAGACAATGACATACGCTTTTAATGAAAAGGAATCACGTTAGAGGAAAA
ATATTTATTCATTATTTGTCAAATTGTCCGGGTTAGTTGGCAGAAATACAGTCTTCCACAA
AGAAAATTCCTATAAGGAAGATTTGGAAGCTCTTCTTCCCAGCACTATGCTTTCCTTCTTT
GGGATAGAGAATGTTCCAGACATTCTCGCTTCCCTGAAAGACTGAAGAAAGTGTAGTGCAT
GGGACCCACGAAACTGCCCTGGCTCCAGTGAAACTTGGGCACATGCTCAGGCTACTATAGG
TCCAGAAGTCCTTATGTTAAGCCCTGGCAGGCAGGTGTTTATTAAAATTCTGAATTTTGGG
GATTTTCAAAAGATAATATTTTACATACACTGTATGTTATAGAACTTCATGGATCAGATCT
GGGGCAGCACCCTATAAATCAACACCTTAATATGCTGCAACAAAATGTAGAATATTCAGAC
AAAATGGATACATAAAGACTAAGTAGCCCATAAGGGGTCAAAATTTGCTGCCAAATGCGTA
TGCCACCAACTTACAAAAACACTTCGTTCGCAGAGCTTTTCAGATTGTGGAATGTTGGATA
AGGAATTATAGACCTCTAGTAGCTGAAATGCAAGACCCCAAGAGGAAGTTCAGATCTTAAT

TGGCAGGGGCTTCTAGCTGACTTGCACAGGGATTCTCACAATAGCCGATATCAGAATTTGT
GTTGAAGGAACTTGTCTCTTCATCTAATATGATAGCGGGAAAAGGAGAGGAAACTACTGCC
TTTAGAAAATATAAGTAAAGTGATTAAAGTGCTCACGTTACCTTGACACATAGTTTTTCAG
TCTATGGGTTTAGTTACTTTAGATGGCAAGCATGTAACTTATATTAATAGTAATTTGTAAA
GTTGGTTGGATAAGCTATCCGTGTTGCAGGTTCATGGATTACTTCTCTATAAAAAATATGT
ATTTACCAAAAATTTTGTGACATTCCTTCTCCCATCTCTTCCTTGACCTGCATTGTAAATA
GGTTCTTCTTGTTCTGAGATTCAATATTGAATTTTTCCTATGCTATTGACAATAAAATATT
ATTGAACTACA

FIGURA 16

GCCGAGCCGGGGCTCCGGAAGCCGGCGGGGGCGCCGCGGCCGTGCGGGGCGTCAATGGATC
GCCACTCCAGCTACATCTTCATCTGGCTGCAGCTGGAGCTCTGCGCCATGGCCGTGCTGCT
CACCAAAGGTGAAATTCGATGCTACTGTGATGCTGCCCACTGTGTAGCCACTGGTTATATG
TGTAAATCTGAGCTCAGCGCCTGCTTCTCTAGACTTCTTGATCCTCAGAACTCAAATTCCC
CACTCACCCATGGCTGCCTGGACTCTCTTGCAAGCACGACAGACATCTGCCAAGCCAAACA
GGCCCGAAACCACTCTGGCACCACCATACCCACATTGGAATGCTGTCATGAAGACATGTGC
AATTACAGAGGGCTGCACGATGTTCTCTCTCCTCCCAGGGGTGAGGCCTCAGGACAAGGAA
ACAGGTATCAGCATGATGGTAGCAGAAACCTTATCACCAAGGTGCAGGAGCTGACTTCTTC
CAAAGAGTTGTGGTTCCGGGCAGCGGTCATTGCCGTGCCCATTGCTGGAGGGCTGATTTTA
GTGTTGCTTATTATGTTGGCCCTGAGGATGCTTCGAAGTGAAAATAAGAGGCTGCAGGATC
AGCGGCAACAGATGCTCTCCCGTTTGCACTACAGCTTTCACGGACACCATTCCAAAAAGGG
GCAGGTTGCAAAGTTAGACTTGGAATGCATGGTGCCGGTCAGTGGGCACGAGAACTGCTGT
CTGACCTGTGATAAAATGAGACAAGCAGACCTCAGCAACGATAAGATCCTCTCGCTTGTTC
ACTGGGGCATGTACAGTGGGCACGGGAAGCTGGAATTCGTATGACGGAGTCTTATCTGAAC
TACACTTACTGAACAGCTTGAAGGCCTTTTGAGTTCTGCTGGACAGGAGCACTTTATCTGA
AGACAAACTCATTTAATCATCTTTGACAGACAAAATGACCTCTGCAAACAGAATCTTGGAT
ATTTCTTCTGAAGGATTATTTGCACAGACTTAAATACAGTTAAATGTGTTATTTGCTTTTA
AAATTATAAAAAGCAAAGAGAAGACTTTGTACACACTGTCACCAGGGTTATTTGCATCCAA
GGGAGCTGGAATTGAGTACCTAAATAAACAAAAATGTGCCCTAAAAA

FIGURA 17A

GTCACTTAGGAAAAGGTGTCCTTTCGGGCAGCCGGGCTCAGCATGGAGGAACAGAAGGAATG
ACACTCTGGACAGCACCCGGACCCTGTACTCCAGCGCGTCTCGGAGCACAGACTTGTCTTA
CAGTGAAAGCGACTTGGTGAATTTTATTCAAGCAAATTTTAAGAAACGAGAATGTGTCTTC
TTTACCAAAGATTCCAAGGCCACGGAGAATGTGTGCAAGTGTGGCTATGCCCAGAGCCAGC
ACATGGAAGGCACCCAGATCAACCAAAGTGAGAAATGGAACTACAAGAAACACACCAAGGA
ATTTCCTACCGACGCCTTTGGGGATATTCAGTTTGAGACACTGGGGAAGAAAGGGAAGTAT
ATACGTCTGTCCTGCGACACGGACGCGGAAATCCTTTACGAGCTGCTGACCCAGCACTGGC
ACCTGAAAACACCCAACCTGGTCATTTCTGTGACCGGGGGCGCCAAGAACTTCGCCCTGAA
GCCGCGCATGCGCAAGATCTTCAGCCGGCTCATCTACATCGCGCAGTCCAAAGGTGCTTGG
ATTCTCACGGGAGGCACCCATTATGGCCTGACGAAGTACATCGGGGAGGTGGTGAGAGATA
ACACCATCAGCAGGAGTTCAGAGGAGAATATTGTGGCCATTGGCATAGCAGCTTGGGGCAT
GGTCTCCAACCGGGACACCCTCATCAGGAATTGCGATGCTGAGGGCTATTTTTTAGCCCAG
TACCTTATGGATGACTTCACAAGGGATCCACTGTATATCCTGGACAACAACCACACACATT
TGCTGCTCGTGGACAATGGCTGTCATGGACATCCCACTGTCGAAGCAAAGCTCCGGAATCA
GCTAGAGAAGCATATCTCTGAGCGCACTATTCAAGATTCCAACTATGGTGGCAAGATCCCC
ATTGTGTGTTTTGCCCAAGGAGGTGGAAAAGAGACTTTGAAAGCCATCAATACCTCCATCA
AAAATAAAATTCCTTGTGTGGTGGTGGAAGGCTCGGGCCGGATCGCTGATGTGATCGCTAG
CCTGGTGGAGGTGGAGGATGCCCCGACATCTTCTGCCGTCAAGGAGAAGCTGGTGCGCTTT
TTACCCCGCACGGTGTCCCGGCTGTCTGAGGAGGAGACTGAGAGTTGGATCAAATGGCTCA
AAGAAATTCTCGAATGTTCTCACCTATTAACAGTTATTAAAATGGAAGAAGCTGGGGATGA
AATTGTGAGCAATGCCATCTCCTACGCTCTATACAAAGCCTTCAGCACCAGTGAGCAAGAC
AAGGATAACTGGAATGGGCAGCTGAAGCTTCTGCTGGAGTGGAACCAGCTGGACTTAGCCA
ATGATGAGATTTTCACCAATGACCGCCGATGGGAGTCTGCTGACCTTCAAGAAGTCATGTT
TACGGCTCTCATAAAGGACAGACCCAAGTTTGTCCGCCTCTTTCTGGAGAATGGCTTGAAC
CTACGGAAGTTTCTCACCCATGATGTCCTCACTGAACTCTTCTCCAACCACTTCAGCACGC
TTGTGTACCGGAATCTGCAGATCGCCAAGAATTCCTATAATGATGCCCTCCTCACGTTTGT
CTGGAAACTGGTTGCGAACTTCCGAAGAGGCTTCCGGAAGGAAGACAGAAATGGCCGGGAC
GAGATGGACATAGAACTCCACGACGTGTCTCCTATTACTCGGCACCCCCTGCAAGCTCTCT
TCATCTGGGCCATTCTTCAGAATAAGAAGGAACTCTCCAAAGTCATTTGGGAGCAGACCAG
GGGCTGCACTCTGGCAGCCCTGGGAGCCAGCAAGCTTCTGAAGACTCTGGCCAAAGTGAAG
AACGACATCAATGCTGCTGGGGAGTCCGAGGAGCTGGCTAATGAGTACGAGACCCGGGCTG
TTGAGCTGTTCACTGAGTGTTACAGCAGCGATGAAGACTTGGCAGAACAGCTGCTGGTCTA
TTCCTGTGAAGCTTGGGGTGGAAGCAACTGTCTGGAGCTGGCGGTGGAGGCCACAGACCAG
CATTTCACCGCCCAGCCTGGGGTCCAGAATTTTCTTTCTAAGCAATGGTATGGAGAGATTT
CCCCAGCAGCAGCAAGAACTGGAAGATTATCCTGTGTCTGTTTATTATACCCTTGGTGGGCTG

TGGCTTTGTATCATTTAGGAAGAAACCTGTCGACAAGCACAAGAAGCTGCTTTGGTACTAT
GTGGCGTTCTTCACCTCCCCCTTCGTGGTCTTCTCCTGGAATGTGGTCTTCTACATCGCCT
TCCTCCTGCTGTTTGCCTACGTGCTGCTCATGGATTTCCATTCGGTGCCACACCCCCCCGA
GCTGGTCCTGTACTCGCTGGTCTTTGTCCTCTTCTGTGATGAAGTGAGACAGTGGTACGTA
AATGGGGTGAATTATTTTACTGACCTGTGGAATGTGATGGACACGCTGGGGCTTTTTTACT
TCATAGCAGGAATTGTATTTCGGCTCCACTCTTCTAATAAAAGCTCTTTGTATTCTGGACG
AGTCATTTTCTGTCTGGACTACATTATTTTCACTCTAAGATTGATCCACATTTTTACTGTA
AGCAGAAACTTAGGACCCAAGATTATAATGCTGCAGAGGATGCTGATCGATGTGTTCTTCT
TCCTGTTCCTCTTTGCGGTGTGGATGGTGGCCTTTGGCGTGGCCAGGCAAGGGATCCTTAG
GCAGAATGAGCAGCGCTGGAGGTGGATATTCCGTTCGGTCATCTACGAGCCCTACCTGGCC
ATGTTCGGCCAGGTGCCCAGTGACGTGGATGGTACCACGTATGACTTTGCCCACTGCACCT
TCACTGGGAATGAGTCCAAGCCACTGTGTGTGGAGCTGGATGAGCACAACCTGCCCCGGTT
CCCCGAGTGGATCACCATCCCCCTGGTGTGCATCTACATGTTATCCACCAACATCCTGCTG
GTCAACCTGCTGGTCGCCATGTTTGGCTACACGGTGGGCACCGTCCAGGAGAACAATGACC
AGGTCTGGAAGTTCCAGAGGTACTTCCTGGTGCAGGAGTACTGCA

FIGURA 17B

GCCGCCTCAATATCCCCTTCCCCTTCATCGTCTTCGCTTACTTCTACATGGTGGTGAAGAA
GTGCTTCAAGTGTTGCTGCAAGGAGAAAAACATGGAGTCTTCTGTCTGCTGTTTCAAAAAT
GAAGACAATGAGACTCTGGCATGGGAGGGTGTCATGAAGGAAAACTACCTTGTCAAGATCA
ACACAAAAGCCAACGACACCTCAGAGGAAATGAGGCATCGATTTAGACAACTGGATACAAA
GCTTAATGATCTCAAGGGTCTTCTGAAAGAGATTGCTAATAAAATCAAATAAAACTGTATG
AAACTCTAATGGAGAAAAATCTAATTATAGCAAGATCATATTAAGGAATGCTGATGAACAA
TTTTGCTATCGACTACTAAATGAGAGATTTTCAGACCCCTGGGTACATGGTGGATGATTTT
AAATCACCCTAGTGTGCTGAGACCTTGAGAATAAAGTGTGTGATTGGTTTCATACTTGAAG
ACGGATATAAAGGAAGAATATTTCCTTTATGTGTTTCTCCAGAATGGTGCCTGTTTCTCTC
TGTGTCTCAATGCCTGGGACTGGAGGTTGATAGTTTAAGTGTGTTCTTACCGCCTCCTTTT
TCCTTTAATCTTATTTTTGATGAACACATATATAGGAGAACATCTATCCTATGAATAAGAA
CCTGGTCATGCTTTACTCCTGTATTGTTATTTTGTTCATTTCCAATTGATTCTCTACTTTT
CCCTTTTTTTGTATTATGTGACTAATTAGTTGGCATATTGTTAAAAGTCTCTCAAATTAGGC
CAGATTCTAAAACATGCTGCAGCAAGAGGACCCCGCTCTCTTCAGGAAAAGTGTTTTCATT
TCTCAGGATGCTTCTTACCTGTCAGAGGAGGTGACAAGGCAGTCTCTTGCTCTCTTGGACT
CACCAGGCTCCTATTGAAGGAACCACCCCCATTCCTAAATATGTGAAAAGTCGCCCAAAAT
GCAACCTTGAAAGGCACTACTGACTTTGTTCTTATTGGATACTCCTCTTATTTATTATTTT
TCCATTAAAAATAATAGCTGGCTATTATAGAAAATTTAGACCATACAGAGATGTAGAAAGA
ACATAAATTGTCCCCATTACCTTAAGGTAATCACTGCTAACAATTTCTGGATGGTTTTTCA
AGTCTATTTTTTTTCTATGTATGTCTCAATTCTCTTTCAAAATTTTACAGAATGTTATCAT
ACTACATATATACTTTTTATGTAAGCTTTTTCACTTAGTATTTTATCAAATATGTTTTTAT
TATATTCATAGCCTTCTTAAACATTATATCAATAATTGCATAATAGGCAACCTCTAGCGAT
TACCATAATTTTGCTCATTGAAGGCTATCTCCAGTTGATCATTGGGATGAGCATCTTTGTG
CATGAATCCTATTGCTGTATTTGGGAAAATTTTCCAAGGTTAGATTCCAATAAATATCTAT
TTATTATTAAATATTAAAATATCGATTTATTATTAAAACCATTTATAAGGCTTTTTCATAA
ATGTATAGCAAATAGGAATTATTAACTTGAGCATAAGATATGAGATACATGAACCTGAACT
ATTAAAATAAAATATTATATTTAACCCTAGTTTAAGAAGAAGTCAATATGCTTATTTAAAT
ATTATGGATGGTGGGCAGATCACTTGAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGG
CAAAACCACATCTCTACTAAAAATAAAAAAATTAGCTGGGTGTGGTGGTGCACTCCTGTAA
TCCCAGCTACTCAGAAGGCTGAGGTACAAGAATTGCTGGAACCTGGGAGGCGGAGGTTGCA
GTGAACCAAGATTGCACCACTGCACTCCAGCCGGGGTGACAGAGTGAGACTCCGACTGAAA
ATAAATAAATAAATAAATAAATAAATAAATAAATAAATATTATGGATGGTGAAGGGAATGG
TATAGAATTGGAGAGATTATCTTACTGAACACCTGTAGTCCCAGCTTTCTCTGGAAGTGGT
GGTATTTGAGCAGGATGTGCACAAGGCAATTGAAATGCCCATAATTAGTTTCTCAGCTTTG
[illegible]AGGAGGAAATTTTAGAAGGAAGCTACTAAAAGATCT

AATTTGAAAAACTACAAAAGCATTAACTAAAAAAGTTTATTTTCCTTTTGTCTGGGCAGTA
GTGAAAATAACTACTCACAACATTCACTATGTTTGCAAGGAATTAACACAAATAAAAGATG
CCTTTTTTACTTAAACGCCAAGACAGAAAACTTGCCCAATACTGAGAAGCAACTTGCATTAG
AGAGGGAACTGTTAAATGTTTTCAACCCAGTTCATCTGGTGGATGTTTTTGCAGGTTACTC
TGAGAATTTTGCTTATGAAAAATCATTATTTTTAGTGTAGTTCACAATAATGTATTGAACA
TACTTCTAATCAAAGGTGCTATGTCCTTGTGTATGGTACTAAATGTGTCCTGTGTACTTTT
GCACAACTGAGAATCCTGCGGCTTGGTTTAATGAGTGTGTTCATGAAATAAATAATGGAGG
AATTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAA

FIGURA 18

GAATTCGCCCTTGGCTCAGCATGAGGAACAGAAGGAATGACACTCTGGACAGCACCCGGAC
CCTGTACTCCAGCGCGTCTCGGAGCACAGACTTGTCTTACAGTGAAAGCGACTTGGTGAAT
TTTATTCAAGCAAATTTTAAGAAACGAGAATGTGTCTTCTTTACCAAAGATTCCAAGGCCA
CGGAGAATGTGTGCAAGTGTGGCTATGCCCAGAGCCAGCACATGGAAGGCACCCAGATCAA
CCAAAGTGAGAAATGGAACTACAAGAAACACACCAAGGAATTTCCTACCGACGCCTTTGGG
GATATTCAGTTTGAGACACTGGGGAAGAAAGGGAAGTATATACGTCTGTCCTGCGACACGG
ACGCGGAAATCCTTTACGAGCTGCTGACCCAGCACTGGCACCTGAAAACACCCAACCTGGT
CATTTCTGTGACCGGGGGCGCCAAGAACTTCGCCCTGAAGCCGCGCATGCGCAAGATCTTC
AGCCGGCTCATCTACATCGCGCAGTCCAAAGGTGCTTGGATTCTCACGGGAGGCACCCATT
ATGGCCTGATGAAGTACATCGGGGAGGTGGTGAGAGATAACACCATCAGCAGGAGTTCAGA
GGAGAATATTGTGGCCATTGGCATAGCAGCTTGGGGCATGGTCTCCAACCGGGACACCCTC
ATCAGGAATTGCGATGCTGAGGGCTATTTTTTAGCCCAGTACCTTATGGATGACTTCACAA
GAGATCCACTGTATATCCTGGACAACAACCACACACATTTGCTGCTCGTGGACAATGGCTG
TCATGGACATCCCACTGTCGAAGCAAAGCTCCGGAATCAGCTAGAGAAGTATATCTCTGAG
CGCACTATTCAAGATTCCAACTATGGTGGCAAGATCCCCATTGTGTGTTTTGCCCAAGGAG
GTGGAAAAGAGACTTTGAAAGCCATCAATACCTCCATCAAAAATAAAATTCCTTGTGTGGT
GGTGGAAGGCTCGGGCCAGATCGCTGATGTGATCGCTAGCCTGGTGGAGGTGGAGGATGCC
CTGACATCTTCTGCCGTCAAGGAGAAGCTGGTGCGCTTTTTACCCCGCACGGTGTCCCGGC
TGCCTGAGGAGGAGACTGAGAGTTGGATCAAATGGCTCAAAGAAATTCTCGAATGTTCTCA
CCTATTAACAGTTATTAAAATGGAAGAAGCTGGGGATGAAATTGTGAGCAATGCCATCTCC
TACGCTCTATACAAAGCCTTCAGCACCAGTGAGCAAGACAAGGATAACTGGAATGGGCAGC
TGAAGCTTCTGCTGGAGTGGAACCAGCTGGACTTAGCCAATGATGAGATTTTCACCAATGA
CCGCCGATGGGAGTCTGCTGACCTTCAAGAAGTCATGTTTACGGCTCTCATAAAGGACAGA
CCCAAGTTTGTCCGCCTCTTTCTGGAGAATGGCTTGAACCTACGGAAGTTTCTCACCCATG
ATGTCCTCACTGAACTCTTCTCCAACCACTTCAGCACGCTTGTGTACCGGAATCTGCAGAT
CGCCAAGAATTCCTATAATGATGCCCTCCTCACGTTTGTCTGGAAACTGGTTGCGAACTTC
CGAAGAGGCTTCCGGAAGGAAGACAGAAATGGCCGGGACGAGATGGACATAGAACTCCACG
ACGTGTCTCCTATTACTCGGCACCCCCTGCAAGCTCTCTTCATCTGGGCCATTCTTCAGAA
TAAGAAGGAACTCTCCAAAGTCATTTGGGAGCAGACCAGGGGCTGCACTCTGGCAGCCCTG
GGAGCCAGCAAGCTTCTGAAGACTCTGGCCAAAGTGAAGAACGACATCAATGCTGCTGGGG
AGTCCGAGGAGCTGGCTAATGAGTACGAGACCCGGGCTGTTGAGCTGTTCACTGAGTGTTA
CAGCAGCGATGAAGACTTGGCAGAACAGCTGCTGGTCTATTCCTGTGAAGCTTGGGGTGGA
AGCAACTGTCTGGAGCTGGCGGTGGAGGCCACAGACCAGCATTTCATCGCCCAGCCTGGGG
TCCAGAATTTTCTTTCTAAGCAATGGTATGGAGAGATTTCCCGAGACACCAAGAACTGGAA
[illegible]

AAACCTGTCGACAAGCACAAGAAGCTGCTTTGGTACTATGTGGCGTTCTTCACCTCCCCCT
TCGTGGTCTTCTCCTGGAATGTGGTCTTCTACATCGCCTTCCTCCTGCTGTTTGCCTACGT
GCTGCTCATGGATTTCCATTCGGTGCCACACCCCCCCGAGCTGGTCCTGTACTCGCTGGTC
TTTGTCCTCTTCTGTGATGAAGTGAGACAGTGGTACGTAAATGGGGTGAATTATTTTACTG
ACCTGTGGAATGTGATGGACACGCTGGGGCTTTTTTACTTCATAGCAGGAATTGTATTTCG
GCTCCACTCTTCTAATAAAAGCTCTTTGTATTCTGGACGAGTCATTTTCTGTCTGGACTAC
ATTATTTTCACTCTAAGATTGATCCACATTTTTACTGTAAGCAGAAACTTAGGACCCAAGA
TTATAATGCTGCAGAGGATGCTGATCGATGTGTTCTTCTTCCTGTTCCTCTTTGCGGTGTG
GATGGTGGCCTTTGGCGTGGCCAGGCAAGGGATCCTTAGGCAGAATGAGCAGCGCTGGAGG
TGGATATTCCGTTCGGTCATCTACGAGCCCTACCTGGCCATGTTCGGCCAGGTGCCCAGTG
ACGTGGATGGTACCACGTATGACTTTGCCCACTGCACCTTCACTGGGAATGAGTCCAAGCC
ACTGTGTGTGGAGCTGGATGAGCACAACCTGCCCCGGTTCCCCGAGTGGATCACCATCCCC
CTGGTGTGCATCTACATGTTATCCACCAACATCCTGCTGGTCAACCTGCTGGTCGCCATGT
TTGGCTACACGGTGGGCACCGTCCAGGAGAACAATGACCAGGTCTGGAAGTTCCAGAGGTA
CTTCCTGGTGCAGGAGTACTGCAGCCGCCTCAATATCCCCTTCCCCTTCATTGTCTTCGCT
TACTTCTACATGGTGGTGAAGAAGTGCTTCAAGTGTTGCTGCAAGGAGAAAAACATGGAGT
CTTCTGTCTGCTGTTTCAAAAATGAAGACAATGAGACTCTGGCATGGGAGGGTGTCATGAA
GGAAAACTACCTTGTCAAGATCAACACAAAAGCCAACGACACCTCAGAGGAAATGAGGCAT
CGATTTAGACAACTGGATACAAAGCTTAATGATCTCAAGGGTCTTCTGAAAGAGATTGCTA
ATAAAATCAAATAAAACTGTATGAACTCTAATGGAGAAAAATCTAATTATAGCAAGATCAT
ATTAAGGAATGCTGATGAACAATTTTGCTATCGACTACTAAATGAGAGATTTTCAGACCCC
TGGGTACATGGTGGATGATTTTAAATCACCCTAGTGTGCTGAGACCTTGAGAATAAAGTGT
GTGATTGGTTTCATACTTGAAGACGGATATAAAGGAAGAATATTTCCTTTATGTGTTTCTC
CAGAATGGTGCCTGTTTCTCTCTGTGTCTCAATGCCTGGGACTGGAGGTTGATAGTTTAAG
TGTGTTCTTACCGCCTCCTTTTTCCTTTAATCTTATTTTTGATGAACACATATATAGGAGA
ACATCTATCCTATGAATAAGAACCTGGTCATGCTTTAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAA

FIGURA 19

GCGGCGGGGGCGAGGTGAGGTGTTGGCAGTGGAAAGGGGTTCGGGCTCGGGGGGCGGGGGG
ACGCGGAGCGGATGGCCCGCGCCGGCCGCAGGGGCGGATAAAAAGCCGTCGCGCTGCGGGA
GTGGGCGGGAGGGAGAGGGGGTGTCCGAGGGCCACAAGAGTATGACGGGGCTGTACGAGCT
GGTGTGGCGGGTGCTGCACGCGCTGCTCTGTCTGCACCGCACGCTCACCTCCTGGCTCCGC
GTTCGGTTCGGCACCTGGAACTGGATCTGGCGGCGCTGCTGCCGCGCCGCCTCTGCCGCGG
TCCTAGCGCCGCTCGGCTTCACGCTCCGCAAGCCCCCGGCAGTCGGCAGGAACCGCCGTCA
CCACCGGCACCCGCGCGGGGGGGTCGTGCCTGGCAGCCGCACACCACCGGATGCGCTGGCGC
GCGGACGGTCGTTCCTTGGAGAAGCTGCCTGTGCATATGGGCCTGGTGATCACCGAGGTGG
AGCAGGAACCCAGCTTCTCGGACATCGCGAGCCTCGTGGTGTGGTGTATGGCCGTGGGCAT
CTCCTACATTAGCGTCTACGACCACCAAGGTATTTTCAAAAGAAATAATTCCAGATTGATG
GATGAAATTTTAAAACAACAGCAAGAACTTCTGGGCCTAGATTGTTCAAAATACTCACCAG
AATTTGCAAATAGTAATGACAAAGATGATCAAGTTTTAAATTGCCATTTGGCAGTGAAGGT
GCTGTCTCCGGAAGATGGAAAAGCAGATATTGTAAGAGCTGCTCAGGACTTTTGCCAGTTA
GTAGCCCAGAAGCAAAGGAGACCCACAGATTTGGATGTAGATACGTTAGCCAGTTTACTTA
GTTCAAATGGTTGTCCTGATCCTGATTTAGTATTGAAGTTCGGTCCTGTGGACAGCACATT
AGGCTTTCTTCCCTGGCACATCAGATTGACTGAGATTGTCTCTTTGCCTTCCCACCTAAAC
ATCAGTTATGAGGACTTTTTCTCTGCCCTTCGTCAATATGCAGCCTGTGAACAGCGTCTGG
GAAAGTAGTGGTCATTGGTTGCATAATTTGATTTGAGGCTTGTGGAGGAAAGGAACCAAGT
GACTCTGATGTTTACAAAGCACCTATGAAACCCTGTACACACCTATGAAACCCTGTACACA
CCTAGTTCATAATCCTCATAATTTATCAACAAACACAAAGAAGTGTCTTACTTGAGAGTGA
GTGTGTGCGTGTGTGCGTGCACACATGTGCACGTTTGTATGTGTGGAAATAAACATAAATG
GGGACGTGTTGGAGAAGGAAATACATAGACCTACAACTTTGAGCATATAGCAGTGATGTTT
TAGGAACTGAAATGTCACACTTAATAAAGTCTTCAGCCCAGCTACTTCCCTGTTTTCGTGG
GGAGAAGGGGGCCTGATTAGAACTGTTCTGGTTGTGTTTGGCGGGAGGGGAATAAATTTTG
TTCAGTCCTTCTTAGTGACCAAACTTTAATTTTTAGGAATAACATATTGACTTACTGAACT
GAAGCATTCTGAGTTGAAAGGAGCCCCAGAGGAAAGGAGTTCTGTGTTGCTCACATGTTAA
AAGCTTGCTCACCTTCAGAGCAGAGGGAATACCTATCTTCAGATATCCGCCCATTTTCATC
TCTTCATTATAGTCAAACAGTGTGACTTGAGAGTGTTGCTCTGGTGTCTGTATTCTGGCTT
ATGAAGATTATTTGAAAAAGAACTCTTACTACATTGAAATGCAGACTTTTAAAAATTTAAA
TATTGGATTAGGCAGTCAAAAAACCAAACAAGCATAAAAGGTCAGTAAGTTGTAATCTTAA
AAGTAAAGGTGGAAAACTCATTATAAATGGAAGAAAAGTTTTGATTTCCTTTTTTGTTTGA
TGGGCAGTATGCCATATTATATCCAAAGTTGGTTTAAAAAATACTTCCATCAACTATTTTT
ATTTAAAATAAACATTTGAGGGAAGTTACCAAGGCAGCTTTTTTCCTCAAAAGTAACCTGT
TCCTCTTTGGAACAGCACATTTTAGGGGCATGGTTAATACCTGAGATTTTTACTCAGTAAA
[illegible]AAGTAGGATTGAAGGCCTCTGTGGGGAATAT

AATGTTACCAAAGCCTATAAAAATAAATTTTACATGTTCTTTTTTATGACAGAGAGCAGCA
CTGGTTCTGTTATTTTTTAAAATGAATAACTGATTTCTTGACAGGTGTTTAGTATTTCTTC
CCTCACTGCTGATTCTTGGATAGAAACCATTCTTTATATTTGATAGACTGTTTTCAGAAAA
CTCTTATCAACAAGTGTACAATAGTTATCTAAAACTGTACATTTAGAATGGAGCAGTTTAA
TACTAGATCTCAGAAATTTGAAAAATAGCAAAGAAGACTGGATTTGGAAAGCATGGTCTAC
AATTGCTTGTCAATTCTGAAGCTATGAAGAATAAATGTTTCAGCTTTGGATTATGAAACCC
CGTTTATGATTTTTAAAACACACTTGAAATAAAAATGATTGAACTAAATTTTGGTCCGGTG
ACATCATTTTGCACTGCATAGCCCATTATACATTGTAAGTCTTTTTGTTTTAATTTATTAC
TGAGAGTTTTCTGTGAAGCTACAGCATATCTAATCAGAGAATTTCTGATTTGTTTCCTTAT
ACTGTGATTATGTTATATTAAGGCATTTGTAGTGCAGCTTAAGACTGAATTTATGCTTTTT
GTAAACATGATCGGTGTAAATGTCTTATAAACATTCTGGAGTATGTATAACTTTAATGAAT
GAAATTTAATGGACCTGATTAAAATAAAGGGATTTAATCTTTGTTAAAGTGAAGTTAGTGG
AATAAATTACCTACTGGAATATAGTCCAAGTCACTAAAGGTTTAATATTTGCATTTTTGTG
CTTTTATTTTCTCCTTCCATTCATAATTATATACTTGAAAGTACATCTGTAGCCTATGACT
TGAGTCTCTTGAACTTCTAGGAAGAGGCAAACTACAAACTACTAGGATTCTGATTTCAGAT
ATAGGCATTCCAGAATCTTCTCTTTACGAGTTCACCTGCTAGTATAATCTCCACAACTTGA
ATGGCCTTGGTTGTTCTGTAATTGCTGCCAAAATCATCACAAGCTGTACGTCATCAAGGCT
CCCTTTGCACTCCCAAGAAGAACTGTTCATTTTAAACAAAAGTGTATGTCTTTATTTGTAT
TGGAAAATATTGTCTTTAAATTGTTTCTTGTTGACACTTCCCACAATGGAAAAATTACCAA
ATTAAACCTGTTTTATGGATGGCAGCTTGGAGCATAGCAAGAATTTGGAGGATTTGAATTC
CATTCCCAGTTCTCATTGCGTTTTGTTTCTTAAAACTATAATAATTGGTTACTGTTATAAA
GTTTAAAAGGTGGTTTTAATATGAATAGCAAATTCTGGTATATTGTGACTAACGCTTAAGA
ATGCCTGTCTTTGAGAGGAAGGTGTTATAATATTAATCAACAGTGCCAAATACACTGTGCA
TATCTACAATTTAATCTTTGAATGTTTGTTACTGGATTAGCTCCCTCCTCCTTCTGTGTGA
TGGTACCATGCATAGAGTCAATCAAATCCTTGTGATGTTTTGTATGGACTTTGACAACATG
TAACTAATGTGTAAAGCAAGTTTTTATGATTAAGGAATCAAATTTATTGAATTTTATTATT
GAAAGTTGAAACTTAACATGTATAAACAAAAAACAATAAAATAATAAACTATTTTCATTGA
CTATAAAAAAAAAGCCCACGCGTCCGCC

FIGURA 20

GAGGTGAGGTGTTGGCAGTGGAAAGGGGTTCGGGCTCGGGGGGCGGGGGGACGCGGAGCGA
TGGCCCGCGCCGGCCGCAGGGGCGGATAAAAAGCCGTCGCGCTGCGGGAGTGGGCGGGAGG
GAGAGGGGGTGTCCGAGGGCCACAAGAGTATGACGGGGCTGTACGAGCTGGTGTGGCGGGT
GCTGCACGCGCTGCTCTGTCTGCACCGCACGCTCACCTCCTGGCTCCGCGTTCGGTTCGGC
ACCTGGAACTGGATCTGGCGGCGCTGCTGCCGCGCCGCCTCTGCCGCGGTCCTAGCGCCGC
TCGGCTTCACGCTCCGCAAGCCCCCGGCAGTCGGCAGGAACCGCCGTCACCACCGGCACCC
GCGCGGGGGGTCGTGCCTGGCAGCCGCACACCACCGGATGCGCTGGCGCGCGGACGGTCGT
TCCTTGGAGAAGCTGCCTGTGCATATGGGCCTGGTGATCACCGAGGTGGAGCAGGAACCCA
GCTTCTCGGACATCGCGAGCCTCGTGGTGTGGTGTATGGCCGTGGGCATCTCCTACATTAG
CGTCTACGACCACCAAGGTATTTTCAAAAGAAATAATTCCAGATTGATGGATGAAATTTTA
AAACAACAGCAAGAACTTCTGGGCCTAGATTGTTCAAAATACTCACCAGAATTTGCAAATA
GTAATGACAAAGATGATCAAGTTTTAAATTGCCATTTGGCAGTGAAGGTGCTGTCTCCGGA
AGATGGAAAAGCAGATATTGTAAGAGCTGCTCAGGACTTTTGCCAGTTAGTAGCCCAGAAG
CAAAAGAGACCCACAGATTTGGATGTAGATACGTTAGCCAGTTTACTTAGTTCAAATGGTT
GTCCTGATCCTGATTTAGTATTGAAGTTCGGTCCTGTGGACAGCACATTAGGCTTTCTTCC
CTGGCACATCAGATTGACTGAGATTGTCTCTTTGCCTTCCCACCTAAACATCAGTTATGAG
GACTTTTTCTCTGCCCTTCGTCAATATGCAGCCTGTGAACAGCGTCTGGGAAAGTAGTGGT
CATTGGTTGCA

FIGURA 21

GGGGGGCGGAACCAGCCTGCACGCGCTGGCTCCGGGTGACAGCCGCGCGCCTCGGCCAGGA
TCTGAGTGATGAGACGTGTCCCCACTGAGGTGCCCCACAGCAGCAGGTGTTGAGCATGGGC
TGAGAAGCTGGACCGGCACCAAAGGGCTGGCAGAAATGGGCGCCTGGCTGATTCCTAGGCA
GTTGGCGGCAGCAAGGAGGAGAGGCCGCAGCTTCTGGAGCAGAGCCGAGACGAAGCAGTTC
TGGAGTGCCTGAACGGCCCCCTGAGCCCTACCCGCCTGGCCCACTATGGTCCAGAGGCTGT
GGGTGAGCCGCCTGCTGCGGCACCGGAAAGCCCAGCTCTTGCTGGTCAACCTGCTAACCTT
TGGCCTGGAGGTGTGTTTGGCCGCAGGCATCACCTATGTGCCGCCTCTGCTGCTGGAAGTG
GGGGTAGAGGAGAAGTTCATGACCATGGTGCTGGGCATTGGTCCAGTGCTGGGCCTGGTCT
GTGTCCCGCTCCTAGGCTCAGCCAGTGACCACTGGCGTGGACGCTATGGCCGCCGCCGGCC
CTTCATCTGGGCACTGTCCTTGGGCATCCTGCTGAGCCTCTTTCTCATCCCAAGGGCCGGC
TGGCTAGCAGGGCTGCTGTGCCCGGATCCCAGGCCCCTGGAGCTGGCACTGCTCATCCTGG
GCGTGGGGCTGCTGGACTTCTGTGGCCAGGTGTGCTTCACTCCACTGGAGGCCCTGCTCTC
TGACCTCTTCCGGGACCCGGACCACTGTCGCCAGGCCTACTCTGTCTATGCCTTCATGATC
AGTCTTGGGGGCTGCCTGGGCTACCTCCTGCCTGCCATTGACTGGGACACCAGTGCCCTGG
CCCCCTACCTGGGCACCCAGGAGGAGTGCCTCTTTGGCCTGCTCACCCTCATCTTCCTCAC
CTGCGTAGCAGCCACACTGCTGGTGGCTGAGGAGGCAGCGCTGGGCCCCACCGAGCCAGCA
GAAGGGCTGTCGGCCCCCTCCTTGTCGCCCCACTGCTGTCCATGCCGGGCCCGCTTGGCTT
TCCGGAACCTGGGCGCCCTGCTTCCCCGGCTGCACCAGCTGTGCTGCCGCATGCCCCGCAC
CCTGCGCCGGCTCTTCGTGGCTGAGCTGTGCAGCTGGATGGCACTCATGACCTTCACGCTG
TTTTACACGGATTTCGTGGGCGAGGGGCTGTACCAGGGCGTGCCCAGAGCTGAGCCGGGCA
CCGAGGCCCGGAGACACTATGATGAAGGCGTTCGGATGGGCAGCCTGGGGCTGTTCCTGCA
GTGCGCCATCTCCCTGGTCTTCTCTCTGGTCATGGACCGGCTGGTGCAGCGATTCGGCACT
CGAGCAGTCTATTTGGCCAGTGTGGCAGCTTTCCCTGTGGCTGCCGGTGCCACATGCCTGT
CCCACAGTGTGGCCGTGGTGACAGCTTCAGCCGCCCTCACCGGGTTCACCTTCTCAGCCCT
GCAGATCCTGCCCTACACACTGGCCTCCCTCTACCACCGGGAGAAGCAGGTGTTCCTGCCC
AAATACCGAGGGGACACTGGAGGTGCTAGCAGTGAGGACAGCCTGATGACCAGCTTCCTGC
CAGGCCCTAAGCCTGGAGCTCCCTTCCCTAATGGACACGTGGGTGCTGGAGGCAGTGGCCT
GCTCCCACCTCCACCCGCGCTCTGCGGGGCCTCTGCCTGTGATGTCTCCGTACGTGTGGTG
GTGGGTGAGCCCACCGAGGCCAGGGTGGTTCCGGGCCGGGGCATCTGCCTGGACCTCGCCA
TCCTGGATAGTGCCTTCCTGCTGTCCCAGGTGGCCCCATCCCTGTTTATGGGCTCCATTGT
CCAGCTCAGCCAGTCTGTCACTGCCTATATGGTGTCTGCCGCAGGCCTGGGTCTGGTCGCC
ATTTACTTTGCTACACAGGTAGTATTTGACAAGAGCGACTTGGCCAAATACTCAGCGTAGA
AAACTTCCAGCACATTGGGGTGGAGGGCCTGCCTCACTGGGTCCCAGCTCCCCGCTCCTGT
TAGCCCCATGGGGCTGCCGGGCTGGCCGCCAGTTTCTGTTGCTGCCAAAGTAATGTGGCTC
[illegible]CAGCTGGGGGCTGGGGCGTCCCT

CTCCTCTCTCCCCAGTCTCTAGGGCTGCCTGACTGGAGGCCTTCCAAGGGGGTTTCAGTCT
GGACTTATACAGGGAGGCCAGAAGGGCTCCATGCACTGGAATGTGGGGACTCTGCAGGTGG
ATTACCCAGGCTCAGGGTTAACAGCTAGCCTCCTAGTTGAGACACACCTAGAGAAGGGTTT
TTGGGAGCTGAATAAACTCAGTCACCTGGTTTCCCATCTCTAAGCCCCTTAACCTGCAGCT
TCGTTTAATGTAGCTCTTGCATGGGAGTTTCTAGGATGAAACACTCCTCCATGGGATTTGA
ACATATGAAAGTTATTTGTAGGGGAAGAGTCCTGAGGGGCAACACACAAGAACCAGGTCCC
CTCAGCCCACAGCACTGTCTTTTTGCTGATCCACCCCCCTCTTACCTTTTATCAGGATGTG
GCCTGTTGGTCCTTCTGTTGCCATCACAGAGACACAGGCATTTAAATATTTAACTTATTTA
TTTAACAAAGTAGAAGGGAATCCATTGCTAGCTTTTCTGTGTTGGTGTCTAATATTTGGGT
AGGGTGGGGGATCCCCAACAATCAGGTCCCCTGAGATAGCTGGTCATTGGGCTGATCATTG
CCAGAATCTTCTTCTCCTGGGGTCTGGCCCCCCAAAATGCCTAACCCAGGACCTTGGAAAT
TCTACTCATCCCAAATGATAATTCCAAATGCTGTTACCCAAGGTTAGGGTGTTGAAGGAAG
GTAGAGGGTGGGGCTTCAGGTCTCAACGGCTTCCCTAACCACCCCTCTTCTCTTGGCCCAG
CCTGGTTCCCCCCACTTCCACTCCCCTCTACTCTCTCTAGGACTGGGCTGATGAAGGCACT
GCCCAAAATTTCCCCTACCCCCAACTTTCCCCTACCCCCAACTTTCCCCACCAGCTCCACA
ACCCTGTTTGGAGCTACTGCAGGACCAGAAGCACAAAGTGCGGTTTCCCAAGCCTTTGTCC
ATCTCAGCCCCCAGAGTATATCTGTGCTTGGGGAATCTCACACAGAAACTCAGGAGCACCC
CCTGCCTGAGCTAAGGGAGGTCTTATCTCTCAGGGGGGTTTAAGTGCCGTTTGCAATAATG
TCGTCTTATTTATTTAGCGGGGTGAATATTTTATACTGTAAGTGAGCAATCAGAGTATAAT
GTTTATGGTGACAAAATTAAAGCCTTTCTTATATGTTT

FIGURA 22

AACAAAAGCTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGCCCCTGGTTCGGCCCACCTCT
GAAGGTTCCAGAATCGATAGTGAATTCGTGGGGAAGCGCAGCCATGGCTCTGCGGAGGCTG
GGGGCCGCGCTGCTGCTGCTGCCGCTGCTCGCCGCCGTGGAAGAAACGCTAATGGACTCCA
CTACAGCGACTGCTGAGCTGGGCTGGATGGTGCATCCTCCATCAGGGTGGGAAGAGGTGAG
TGGCTACGATGAGAACATGAACACGATCCGCACGTACCAGGTGTGCAACGTGTTTGAGTCA
AGCCAGAACAACTGGCTACGGACCAAGTTTATCCGGCGCCGTGGCGCCCACCGCATCCACG
TGGAGATGAAGTTTTCGGTGCGTGACTGCAGCAGCATCCCCAGCGTGCCTGGCTCCTGCAA
GGAGACCTTCAACCTCTATTACTATGAGGCTGACTTTGACTCGGCCACCAAGACCTTCCCC
AACTGGATGGAGAATCCATGGGTGAAGGTGGATACCATTGCAGCCGACGAGAGCTTCTCCC
AGGTGGACCTGGGTGGCCGCGTCATGAAAATCAACACCGAGGTGCGGAGCTTCGGACCTGT
GTCCCGCAGCGGCTTCTACCTGGCCTTCCAGGACTATGGCGGCTGCATGTCCCTCATCGCC
GTGCGTGTCTTCTACCGCAAGTGCCCCCGCATCATCCAGAATGGCGCCATCTTCCAGGAAA
CCCTGTCGGGGGCTGAGAGCACATCGCTGGTGGCTGCCCGGGGCAGCTGCATCGCCAATGC
GGAAGAGGTGGATGTACCCATCAAGCTCTACTGTAACGGGGACGGCGAGTGGCTGGTGCCC
ATCGGGCGCTGCATGTGCAAAGCAGGCTTCGAGGCCGTTGAGAATGGCACCGTCTGCCGAG
GTTGTCCATCTGGGACTTTCAAGGCCAACCAAGGGGATGAGGCCTGTACCCACTGTCCCAT
CAACAGCCGGACCACTTCTGAAGGGGCCACCAACTGTGTCTGCCGCAATGGCTACTACAGA
GCAGACCTGGACCCCCTGGACATGCCCTGCACAACCATCCCCTCCGCGCCCCAGGCTGTGA
TTTCCAGTGTCAATGAGACCTCCCTCATGCTGGAGTGGACCCCTCCCCGCGACTCCGGAGG
CCGAGAGGACCTCGTCTACAACATCATCTGCAAGAGCTGTGGCTCGGGCCGGGGTGCCTGC
ACCCGCTGCGGGGACAATGTACAGTACGCACCACGCCAGCTAGGCCTGACCGAGCCACGCA
TTTACATCAGTGACCTGCTGGCCCACACCCAGTACACCTTCGAGATCCAGGCTGTGAACGG
CGTTACTGACCAGAGCCCCTTCTCGCCTCAGTTCGCCTCTGTGAACATCACCACCAACCAG
GCAGCTCCATCGGCAGTGTCCATCATGCATCAGGTGAGCCGCACCGTGGACAGCATTACCC
TGTCGTGGTCCCAGCCAGACCAGCCCAATGGCGTGATCCTGGACTATGAGCTGCAGTACTA
TGAGAAGGAGCTCAGTGAGTACAACGCCACAGCCATAAAAAGCCCCACCAACACGGTCACC
GTGCAGGGCCTCAAAGCCGGCGCCATCTATGTCTTCCAGGTGCGGGCACGCACCGTGGCAG
GCTACGGGCGCTACAGCGGCAAGATGTACTTCCAGACCATGACAGAAGCCGAGTACCAGAC
AAGCATCCAGGAGAAGTTGCCACTCATCATCGGCTCCTCGGCCGCTGGCCTGGTCTTCCTC
ATTGCTGTGGTTGTCATCGCCATCGTGTGTAACAGAAGACGGGGGTTTGAGCGTGCTGACT
CGGAGTACACGGACAAGCTGCAACACTACACCAGTGGCCACATGACCCCAGGCATGAAGAT
CTACATCGATCCTTTCACCTACGAGGACCCCAACGAGGCAGTGCGGGAGTTTGCCAAGGAA
ATTGACATCTCCTGTGTCAAAATTGAGCAGGTGATCGGAGCAGGGGAGTTTGGCGAGGTCT
GCAGTGGCCACCTGAAGCTGCCAGGCAAGAGAGAGATCTTTGTGGCCATCAAGACGCTCAA
[illegible]TGAGCGAAGCCTCCATCATGGGCCAG

TTCGACCATCCCAACGTCATCCACCTGGAGGGTGTCGTGACCAAGAGCACACCTGTGATGA
TCATCACCGAGTTCATGGAGAATGGCTCCCTGGACTCCTTTCTCCGGCAAAACGATGGGCA
GTTCACAGTCATCCAGCTGGTGGGCATGCTTCGGGGCATCGCAGCTGGCATGAAGTACCTG
GCAGACATGAACTATGTTCACCGTGACCTGGCTGCCCGCAACATCCTCGTCAACAGCAACC
TGGTCTGCAAGGTGTCGGACTTTGGGCTCTCACGCTTTCTAGAGGACGATACCTCAGACCC
CACCTACACCAGTGCCCTGGGCGGAAAGATCCCCATCCGCTGGACAGCCCCGGAAGCCATC
CAGTACCGGAAGTTCACCTCGGCCAGTGATGTGTGGAGCTACGGCATTGTCATGTGGGAGG
TGATGTCCTATGGGGAGCGGCCCTACTGGGACATGACCAACCAGGATGTAATCAATGCCAT
TGAGCAGGACTATCGGCTGCCACCGCCCATGGACTGCCCGAGCGCCCTGCACCAACTCATG
CTGGACTGTTGGCAGAAGGACCGCAACCACCGGCCCAAGTTCGGCCAAATTGTCAACACGC
TAGACAAGATGATCCGCAATCCCAACAGCCTCAAAGCCATGGCGCCCCTCTCCTCTGGCAT
CAACCTGCCGCTGCTGGACCGCACGATCCCCGACTACACCAGCTTTAACACGGTGGACGAG
TGGCTGGAGGCCATCAAGATGGGGCAGTACAAGGAGAGCTTCGCCAATGCCGGCTTCACCT
CCTTTGACGTCGTGTCTCAGATGATGATGGAGGACATTCTCCGGGTTGGGCTCACTTTGGC
TGGCCACCAGAAAAAAATCCTGAACAGTATCCAGGTGATGCGGGCGCAGATGAACCAGATT
CAGTCTGTGGAGGTTTGACATTCACCTGCCTCGGCTCACCTCTTCCTCCAAGCCCCGCCCC
CTCTGCCCCACGTGCCGGCCCTCCTGGTGCTCTATCCACTGCAGGGCCAGCCACTCGCCAG
GAGGCCACGGGCCACGGGAAGAACCAAGCGGTGCCAGCCACGAGACGTCACCAAGAAAACA
TGCAACTCAAACGACGGAAAAAAAAAGGGAATGGGAAAAAAGAAAACAGATCCTGGGAGGG
GGCGGGAAATACAAGGAATATTTTTTAAAGAGGATTCTCATAAGGAAAGCAATGACTGTTC
TTGCGGGGGATAAAAAAGGGCTTGGGAGATTCATGCGATGTGTCCAATCGGAGACAAAAGC
AGTTTCTCTCCAACTCCCTCTGGGAAGGTGACCTGGCCAGAGCCAAGAAACACTTTCAGAA
AAACAAATGTGAAGGGGAGAGACAGGGGCCGCCCTTGGCTCCTGTCCCTGCTGCTCCTCTA
GGCCTCACTCAACAACCAAGCGCCTGGAGGACGGGACAGATGGACAGAGAGCCACCCTGAG
AACCCCTCTGGGAAAATCTATTCCTGCCACCACTGGGCAAACAGAAGAATTTTTCTGTCTT
TGGAGAGTATTTTAGAAACTCCAATGAAAGACACTGTTTCTCCTGTTGGCTCACAGGGCTG
AAAGGGGCTTTTGTCCTCCTGGGTCAGGGAGAACGCGGGGACCCCAG

FIGURA 23

CGGGAAGCGCAGCCATGGCTCTGCGGAGGCTGGGGGCCGCGCTGCTGCTGCTGCCGCTGCT
CGCCGCCGTGGAAGAAACGCTAATGGACTCCACTACAGCGACTGCTGAGCTGGGCTGGATG
GTGCATCCTCCATCAGGGTGGGAAGAGGTGAGTGGCTACGATGAGAACATGAACACGATCC
GCACGTACCAGGTGTGCAACGTGTTTGAGTCAAGCCAGAACAACTGGCTACGGACCAAGTT
TATCCGGCGCCGTGGCGCCCACCGCATCCACGTGGAGATGAAGTTTTCGGTGCGTGACTGC
AGCAGCATCCCCAGCGTGCCTGGCTCCTGCAAGGAGACCTTCAACCTCTATTACTATGAGG
CTGACTTTGACTCGGCCACCAAGACCTTCCCCAACTGGATGGAGAATCCATGGGTGAAGGT
GGATACCATTGCAGCCGACGAGAGCTTCTCCCAGGTGGACCTGGGTGGCCGCGTCATGAAA
ATCAACACCGAGGTGCGGAGCTTCGGACCTGTGTCCCGCAGCGGCTTCTACCTGGCCTTCC
AGGACTATGGCGGCTGCATGTCCCTCATCGCCGTGCGTGTCTTCTACCGCAAGTGCCCCCG
CATCATCCAGAATGGCGCCATCTTCCAGGAAACCCTGTCGGGGGCTGAGAGCACATCGCTG
GTGGCTGCCCGGGGCAGCTGCATCGCCAATGCGGAAGAGGTGGATGTACCCATCAAGCTCT
ACTGTAACGGGGACGGCGAGTGGCTGGTGCCCATCGGGCGCTGCATGTGCAAAGCAGGCTT
CGAGGCCGTTGAGAATGGCACCGTCTGCCGAGGTTGTCCATCTGGGACTTTCAAGGCCAAC
CAAGGGGATGAGGCCTGTACCCACTGTCCCATCAACAGCCGGACCACTTCTGAAGGGGCCA
CCAACTGTGTCTGCCGCAATGGCTACTACAGAGCAGACCTGGACCCCCTGGACATGCCCTG
CACAACCATCCCCTCCGCGCCCCAGGCTGTGATTTCCAGTGTCAATGAGACCTCCCTCATG
CTGGAGTGGACCCCTCCCCGCGACTCCGGAGGCCGAGAGGACCTCGTCTACAACATCATCT
GCAAGAGCTGTGGCTCGGGCCGGGGTGCCTGCACCCGCTGCGGGGACAATGTACAGTACGC
ACCACGCCAGCTAGGCCTGACCGAGCCACGCATTTACATCAGTGACCTGCTGGCCCACACC
CAGTACACCTTCGAGATCCAGGCTGTGAACGGCGTTACTGACCAGAGCCCCTTCTCGCCTC
AGTTCGCCTCTGTGAACATCACCACCAACCAGGCAGCTCCATCGGCAGTGTCCATCATGCA
TCAGGTGAGCCGCACCGTGGACAGCATTACCCTGTCGTGGTCCCAGCCGGACCAGCCCAAT
GGCGTGATCCTGGACTATGAGCTGCAGTACTATGAGAAGGAGCTCAGTGAGTACAACGCCA
CAGCCATAAAAAGCCCCACCAACACGGTCACCGTGCAGGGCCTCAAAGCCGGCGCCATCTA
TGTCTTCCAGGTGCGGGCACGCACCGTGGCAGGCTACGGGCGCTACAGCGGCAAGATGTAC
TTCCAGACCATGACAGAAGCCGAGTACCAGACAAGCATCCAGGAGAAGTTGCCACTCATCA
TCGGCTCCTCGGCCGCTGGCCTGGTCTTCCTCATTGCTGTGGTTGTCATCGCCATCGTGTG
TAACAGACGGGGGTTTGAGCGTGCTGACTCGGAGTACACGGACAAGCTGCAACACTACACC
AGTGGCCACATGACCCCAGGCATGAAGATCTACATCGATCCTTTCACCTACGAGGACCCCA
ACGAGGCAGTGCGGGAGTTTGCCAAGGAAATTGACATCTCCTGTGTCAAAATTGAGCAGGT
GATCGGAGCAGGGGAGTTTGGCGAGGTCTGCAGTGGCCACCTGAAGCTGCCAGGCAAGAGA
GAGATCTTTGTGGCCATCAAGACGCTCAAGTCGGGCTACACGGAGAAGCAGCGCCGGGACT
TCCTGAGCGAAGCCTCCATCATGGGCCAGTTCGACCATCCCAACGTCATCCACCTGGAGGG
[illegible]AGTTCATGGAGAATGGCTCCCTG

GACTCCTTTCTCCGGCAAAACGATGGGCAGTTCACAGTCATCCAGCTGGTGGGCATGCTTC
GGGGCATCGCAGCTGGCATGAAGTACCTGGCAGACATGAACTATGTTCACCGTGACCTGGC
TGCCCGCAACATCCTCGTCAACAGCAACCTGGTCTGCAAGGTGTCGGACTTTGGGCTCTCA
CGCTTTCTAGAGGACGATACCTCAGACCCCACCTACACCAGTGCCCTGGGCGGAAAGATCC
CCATCCGCTGGACAGCCCCGGAAGCCATCCAGTACCGGAAGTTCACCTCGGCCAGTGATGT
GTGGAGCTACGGCATTGTCATGTGGGAGGTGATGTCCTATGGGGAGCGGCCCTACTGGGAC
ATGACCAACCAGGATGTAATCAATGCCATTGAGCAGGACTATCGGCTGCCACCGCCCATGG
ACTGCCCGAGCGCCCTGCACCAACTCATGCTGGACTGTTGGCAGAAGGACCGCAACCACCG
GCCCAAGTTCGGCCAAATTGTCAACACGCTAGACAAGATGATCCGCAATCCCAACAGCCTC
AAAGCCATGGCGCCCCTCTCCTCTGGCATCAACCTGCCGCTGCTGGACCGCACGATCCCCG
ACTACACCAGCTTTAACACGGTGGACGAGTGGCTGGAGGCCATCAAGATGGGGCAGTACAA
GGAGAGCTTCGCCAATGCCGGCTTCACCTCCTTTGACGTCGTGTCTCAGATGATGATGGAG
GACATTCTCCGGGTTGGGGTCACTTTGGCTGGCCACCAGAAAAAAATCCTGAACAGTATCC
AGGTGATGCGGGCGCAGATGAACCAGATTCAGTCTGTGGAGGTTTGACATTCACCTGCCTC
GGCTCACCTCTTCCTCCAAGCCCCGCCCCCTCTGCCCCACGTGCCGGCCCTCCTGGTGCTC
TATCCACTGCAGGGCCAGCCACTCGCCAGGAGGCCACGGGCCACGGGAAGAACCAAGCGGT
GCCAGCCACGAGACGTCACCAAGAAAACATGCAACTCAAACGACGGAAAAAAAAAGGGAAT
GGGAAAAAAGAAAACAGATCCTGGGAGGGGGCGGGAAATACAAGGAATATTTTTTAAAGAG
GATTCTCATAAGGAAAGCAATGACTGTTCTTGCGGGGGATAAAAAAGGGCTTGGGAGATTC
ATGCGATGTGTCCAATCGGAGACAAAAGCAGTTTCTCTCCAACTCCCTCTGGGAAGGTGAC
CTGGCCAGAGCCAAGAAACACTTTCAGAAAAACAAATGTGAAGGGGAGAGACAGGGGCCGC
CCTTGGCTCCTGTCCCTGCTGCTCCTCTAGGCCTCACTCAACAACCAAGCGCCTGGAGGAC
GGGACAGATGGACAGACAGCCACCCTGAGAACCCCTCTGGGAAAATCTATTCCTGCCACCA
CTGGGCAAACAGAAGAATTTTTCTGTCTTTGGAGAGTATTTTAGAAACTCCAATGAAAGAC
ACTGTTTCTCCTGTTGGCTCACAGGGCTGAAAGGGGCTTTTGTCCTCCTGGGTCAGGGAGA
ACGCGGGGACCCCAGAAAGGTCAGCCTTCCTGAGGATGGGCAACCCCCAGGTCTGCAGCTC
CAGGTACATATCACGCGCACAGCCTGGCAGCCTGGCCCTCCTGGTGCCCACTCCCGCCAGC
CCCTGCCTCGAGGTCGAC

FIGURA 24

CCGCCAGCTCCTTCGGCAATGAACTTCTCCACCAGCAGCAGCAGCTTCGCCTACGACCGGG
AGTTCCTCCGCACCCTGCCCGGCTTCCTCATCGTGGCCGAGATCGTTCTGGGGCTGCTGGT
ATGGACGCTTATTGCTGGAACTGAGTACTTCCGGGTCCCCGCATTTGGCTGGGTCATGTTT
GTAGCTGTATTTTACTGGGTCCTCACCGTCTTCTTCCTCATTATCTACATAACAATGACCT
ACACCAGGATTCCCCAGGTGCCCTGGACAACAGTGGGCCTGTGCTTTAACGGCAGTGCCTT
CGTCTTGTACCTCTCTGCCGCTGTTGTAGATGCATCTTCCGTCTCCCCTGAGAGGGACAGT
CACAACTTCAACAGCTGGGCGGCCTCATCGTTCTTTGCCTTCCTGGTCAACATCTGCTACG
CTGGAAATACATATTTCAGTTTTATAGCATGGAGATCCAGGACCATACAGTGATTTACCAT
TTTGATAATTAAAAGGAAAAAAAAAGGAAGACTCTCACTGTAAAAACAGCTGTAGGTATAA
TGTATATTCCCAGAGAATTGTATTTAACTAATTAATGTTTTTTATATTCTTAAATTTGCTC
ACAAATTGTGGTTTGTTACAATTAAACTGGATACTTATTTGCAAAGTGTTGTAGCTTATAA
TGAACTCTTAAGTATCTTATTAATGTATTAATGTCTTCATAGATCATATTTTCTTAGACAA
TGTTTAAATAGATAAATTGCTAATATTGAGAATGTGTCAAGTTTGTAAACCTAACTTTTAA
GATGCCAGATTCTTTTTTGATTAAATGTTGCAAAATCCCAAAAAAAAAAAAAAAAAAAAA

FIGURA 25

GCCCTTGGCAATGAACTTCTCCACCAGCAGCAGCAGCTTCGCCTACGACCGGGAGTTCCTC
CGCACCCTGCCCGGCTTCCTCATCGTGGCCGAGATCGTTCTGGGGCTGCTGGTATGGACGC
TTATTGCTGGAACTGAGTACTTCCGGGTCCCCGCATTTGGCTGGGTCATGTTTGTAGCTGT
ATTTTACTGGGTCCTCACCGTCTTCTTCCTCATTATCTACATAACAATGACCTACACCAGG
ATTCCCCAGGTGCCCTGGACAACAGTGGGCCTGTGCTTTAACGGCAGTGCCTTCGTCTTGT
ACCTCTCTGCCGCTGTTGTAGATGCATCTTCCGTCTCCCCTGAGAGGGACAGTCACAACTT
CAACAGCTGGGCGGCCTCATCGTTCTTTGCCTTCCTGGTCACCATCTGCTACGCTGGAAAT
ACATATTTCAGTTTTATAGCATGGAGATCCAGGACCATACAGTGATTTACCAAAGGGC

FIGURA 26

GGCCCTGCCCTGAAGGCTGGTCACTTGCAGAGGTAAACTCCCCTCTTTGACTTCTGGCCAG
GGTTTGTGCTGAGCTGGCTGCAGCCGCTCTCAGCCTCGCTCCGGGCACGTCGGGCAGCCTC
GGGCCCTCCTGCCTGCAGGATCATGCCCACCACCGTGGACGATGTCCTGGAGCATGGAGGG
GAGTTTCACTTTTTCCAGAAGCAAATGTTTTTCCTCTTGGCTCTGCTCTCGGCTACCTTCG
CGCCCATCTACGTGGGCATCGTCTTCCTGGGCTTCACCCCTGACCACCGCTGCCGGAGCCC
CGGAGTGGCCGAGCTGAGTCTGCGCTGCGGCTGGAGTCCTGCAGAGGAACTGAACTACACG
GTGCCGGGCCCAGGACCTGCGGGCGAAGCCTCCCCAAGACAGTGTAGGCGCTACGAGGTGG
ACTGGAACCAGAGCACCTTTGACTGCGTGGACCCCCTGGCCAGCCTGGACACCAACAGGAG
CCGCCTGCCACTGGGCCCCTGCCGGGACGGCTGGGTGTACGAGACGCCTGGCTCGTCCATC
GTCACCGAGTTTAACCTGGTATGTGCCAACTCCTGGATGTTGGACCTATTCCAGTCATCAG
TGAATGTAGGATTCTTTATTGGCTCTATGAGTATCGGCTACATAGCAGACAGGTTTGGCCG
TAAGCTCTGCCTCCTAACTACAGTCCTCATAAATGCTGCAGCTGGAGTTCTCATGGCCATT
TCCCCAACCTATACGTGGATGTTAATTTTTCGCTTAATCCAAGGACTGGTCAGCAAAGCAG
GCTGGTTAATAGGCTACATCCTGATTACAGAATTTGTTGGGCGGAGATATCGGAGAACAGT
GGGGATTTTTTACCAAGTTGCCTATACAGTTGGGCTCCTGGTGCTAGCTGGGGTGGCTTAC
GCACTTCCTCACTGGAGGTGGTTGCAGTTCACAGTTGCTCTGCCCAACTTCTTCTTCTTGC
TCTATTACTGGTGCATACCTGAGTCTCCCAGGTGGCTGATCTCCCAGAATAAGAATGCTGA
AGCCATGAGAATCATTAAGCACATCGCAAAGAAAAATGGAAAATCTCTACCCGCCTCCCTT
CAGCGCCTGAGACTTGAAGAGGAAACTGGCAAGAAATTGAACCCTTCATTTCTTGACTTGG
TCAGAACTCCTCAGATAAGGAAACATACTATGATATTGATGTACAACTGGTTCACGAGCTC
TGTGCTCTACCAGGGCCTCATCATGCACATGGGCCTTGCAGGTGACAATATCTACCTGGAT
TTCTTCTACTCTGCCCTGGTTGAATTCCCAGCTGCCTTCATGATCATCCTCACCATCGACC
GCATCGGACGCCGTTACCCTTGGGCTGCATCAAATATGGTTGCAGGGGCAGCCTGTCTGGC
CTCAGTTTTTATACCTGGTGATCTACAATGGCTAAAAATTATTATCTCATGCTTGGGAAGA
ATGGGGATCACAATGGCCTATGAGATAGTCTGCCTGGTCAATGCTGAGCTGTACCCCACAT
TCATTAGGAATCTTGGCGTCCACATCTGTTCCTCAATGTGTGACATTGGTGGCATCATCAC
GCCATTCCTGGTCTACCGGCTCACTAACATCTGGCTTGAGCTCCCGCTGATGGTTTTCGGC
GTACTTGGCTTGGTTGCTGGAGGTCTGGTGCTGTTGCTTCCAGAAACTAAAGGGAAAGCTT
TGCCTGAGACCATCGAGGAAGCCGAAAATATGCAAAGACCAAGAAAAAATAAAGAAAAGAT
GATTTACCTCCAAGTTCAGAAACTAGACATTCCATTGAACTAAGAAGAGAGACCGTTGCTG
CTGTCATGACCTAGCTTTGATGGCAGCAAGACCAAAAGTAGAAATCCCTGCACTCATCACA
AAGCCCATACAACTCAACCAAACTTACCCCTGAGCCCTATCAACCTAGGTCTACAGCCAGT
GGAGTCTATTGTACACTGTGGAAAAATACCCATGGGACCAGATCCTGCCAAATTCTTCCAG
CTCACTTTATTCTCAGCATTCCTAGGACATTGGACATTGGTTTTCTGGAGGGTTTTTTTTC
[illegible]

AGGGGAACTGTGGGCTAGGCAAANAAAATAGAAAAAGTGTGAAAAACAGTAAAGTTGGGAG
AGGAGCATCTATTTTCTTAAAGAAATAAAACACCNAAAACAAAAAAAAAAAAAAAAAAAAA

FIGURA 27

GCCCTTCTCCTGCCTGCAGGATCATGCCCACCACCGTGGACGATGTCCTGGAGCATGGAGG
GGAGTTTCACTTTTTCCAGAAGCAAATGTTTTTCCTCTTGGCTCTGCTCTCGGCTACCTTC
GCGCCCATCTACGTGGGCATCGTCTTCCTGGGCTTCACCCCTGACCACCGCTGCCGGAGCC
CCGGAGTGGCCGAGCTGAGTCTGCGCTGCGGCTGGAGTCCTGCAGAGGAACTGAACTACAC
GGTGCCGGGCCCAGGACCTGCGGGCGAAGCCTCCCCAAGACAGTGTAGGCGCTACGAGGTG
GACTGGAACCAGAGCACCTTCGACTGCGTGGACCCCCTGGCCAGCCTGGACACCAACAGGA
GCCGCCTGCCACTGGGCCCCTGCCGGGACGGCTGGGTGTACGAGACGCCTGGCTCGTCCAT
CGTCACCGAGTTTAACCTGGTATGTGCCAACTCCTGGATGTTGGACCTATTCCAGTCATCA
GTGAATGTAGGATTCTTTATTGGCTCTATGAGTATCGGCTACATAGCAGACAGGTTTGGCC
GTAAGCTCTGCCTCCTAACTACAGTCCTCATAAATGCTGCAGCTGGAGTTCTCATGGCCAT
TTCCCCAACCTATACGTGGATGTTAATTTTTCGCTTAATCCAAGGACTGGTCAGCAAAGCA
GGCTGGTTAATAGGCTACATCCTGATTACAGAATTTGTTGGGCGGAGATATCGGAGAACAG
TGGGGATTTTTTACCAAGTTGCCTATACAGTTGGGCTCCTGGTGCTAGCTGGGGTGGCTTA
CGCACTTCCTCACTGGAGGTGGTTGCAGTTCACAGTTTCTCTGCCCAACTTCTTCTTCTTG
CTCTATTACTGGTGCATACCTGAGTCTCCCAGGTGGCTGATCTCCCAGAATAAGAATGCTG
AAGCCATGAGAATCATTAAGCACATCGCAAAGAAAAATGGAAAATCTCTACCCGCCTCCCT
TCAGCGCCTGAGACTTGAAGAGGAAACTGGCAAGAAATTGAACCCTTCATTTCTTGACTTG
GTCAGAACTCCTCAGATAAGGAAACATACTATGATATTGATGTACAACTGGTTCACGAGCT
CTGTGCTCTACCAGGGCCTCATCATGCACATGGGCCTTGCAGGTGACAATATCTACCTGGA
TTTCTTCTACTCTGCCCTGGTTGAATTCCCAGCTGCCTTCATGATCATCCTCACCATCGAC
CGCATCGGACGCCGTTACCCTTGGGCTGCATCAAATATGGTTGCAGGGGCAGCCTGTCTGG
CCTCAGTTTTTATACCTGGTGATCTACAATGGCTAAAAATTATTATCTCATGCTTGGGAAG
AATGGGGATCACAATGGCCTATGAGATAGTCTGCCTGGTCAATGCTGAGCTGTACCCCACA
TTCATTAGGAATCTTGGCGTCCACATCTGTTCCTCAATGTGTGACATTGGTGGCATCATCA
CGCCATTCCTGGTCTACCGGCTCACTAACATCTGGCTTGAGCTCCCGCTGATGGTTTTCGG
CGTGCTTGGCTTGGTTGCTGGAGGTCTGGTGCTGTTGCTTCCAGAAACTAAAGGGAAAGCT
TTGCCTGAGACCATCGAGGAAGCCGAAAATATGCAAAGACCAAGAAAAAATAAAGAAAAGA
TGATTTACCTCCAAGTTCAGAAACTAGACATTCCATTGAACTAAGAAGAGAGACCGTTGCT
GCTGTCATGACAAGGGC

FIGURA 28

CGGGCCAGGTTTCCAGGCTCGGCCGCCGCCTCCATCCCAGCACCTGCGGAGGGAGCGCTGA
CCATGGCTCCCTGGCCTGAATTGGGAGATGCCCAGCCCAACCCCGATAAGTACCTCGAAGG
GGCCGCAGGTCAGCAGCCCACTGCCCCTGATAAAAGCAAAGAGACCAACAAAACAGATAAC
ACTGAGGCACCTGTAACCAAGATTGAACTTCTGCCGTCCTACTCCACGGCTACACTGATAG
ATGAGCCCACTGAGGTGGATGACCCCTGGAACCTACCCACTCTTCAGGACTCGGGGATCAA
GTGGTCAGAGAGAGACACCAAAGGGAAGATTCTCTGTTTCTTCCAAGGGATTGGGAGATTG
ATTTTACTTCTCGGATTTCTCTACTTTTTCGTGTGCTCCCTGGATATTCTTAGTAGCGCCT
TCCAGCTGGTTGGAGGAAAAATGGCAGGACAGTTCTTCAGCAACAGCTCTATTATGTCCAA
CCCTTTGTTGGGGCTGGTGATCGGGGTGCTGGTGACCGTCTTGGTGCAGAGCTCCAGCACC
TCAACGTCCATCGTTGTCAGCATGGTGTCCTCTTCATTGCTCACTGTTCGGGCTGCCATCC
CCATTATCATGGGGGCCAACATTGGAACGTCAATCACCAACACTATTGTTGCGCTCATGCA
GGTGGGAGATCGGAGTGAGTTCAGAAGAGCTTTTGCAGGAGCCACTGTCCATGACTTCTTC
AACTGGCTGTCCGTGTTGGTGCTCTTGCCCGTGGAGGTGGCCACCCATTACCTCGAGATCA
TAACCCAGCTTATAGTGGAGAGCTTCCACTTCAAGAATGGAGAAGATGCCCCAGATCTTCT
GAAAGTCATCACTAAGCCCTTCACAAAGCTCATTGTCCAGCTGGATAAAAAAGTTATCAGC
CAAATTGCAATGAACGATGAAAAAGCGAAAAACAAGAGTCTTGTCAAGATTTGGTGCAAAA
CTTTTACCAACAAGACCCAGATTAACGTCACTGTTCCCTCGACTGCTAACTGCACCTCCCC
TTCCCTCTGTTGGACGGATGGCATCCAAAACTGGACCATGAAGAATGTGACCTACAAGGAG
AACATCGCCAAATGCCAGCATATCTTTGTGAATTTCCACCTCCCGGATCTTGCTGTGGGCA
CCATCTTGCTCATACTCTCCCTGCTGGTCCTCTGTGGTTGCCTGATCATGATTGTCAAGAT
CCTGGGCTCTGTGCTCAAGGGGCAGGTCGCCACTGTCATCAAGAAGACCATCAACACTGAT
TTCCCCTTTCCCTTTGCATGGTTGACTGGCTACCTGGCCATCCTCGTCGGGGCAGGCATGA
CCTTCATCGTACAGAGCAGCTCTGTGTTCACGTCGGCCTTGACCCCCCTGATTGGAATCGG
CGTGATAACCATTGAGAGGGCTTATCCACTCACGCTGGGCTCCAACATCGGCACCACCACC
ACCGCCATCCTGGCCGCCTTAGCCAGCCCTGGCAATGCATTGAGGAGTTCACTCCAGATCG
CCCTGTGCCACTTTTTCTTCAACATCTCCGGCATCTTGCTGTGGTACCCGATCCCGTTCAC
TCGCCTGCCCATCCGCATGGCCAAGGGGCTGGGCAACATCTCTGCCAAGTATCGCTGGTTC
GCCGTCTTCTACCTGATCATCTTCTTCTTCCTGATCCCGCTGACGGTGTTTGGCCTCTCGC
TGGCCGGCTGGCGGGTGCTGGTTGGTGTCGGGGTTCCCGTCGTCTTCATCATCATCCTGGT
ACTGTGCCTCCGACTCCTGCAGTCTCGCTGCCCACGCGTCCTGCCGAAGAAACTCCAGAAC
TGGAACTTCCTGCCGCTGTGGATGCGCTCGCTGAAGCCCTGGGATGCCGTCGTCTCCAAGT
TCACCGGCTGCTTCCAGATGCGCTGCTGCTACTGCTGCCGCGTGTGCTGCCGCGCGTGCTG
CTTGCTGTGTGGCTGCCCCAAGTGCTGCCGCTGCAGCAAGTGCTGCGAGGACTTGGAGGAG
GCGCAGGAGGGGCAGGATGTCCCTGTCAAGGCTCCTGAGACCTTTGATAACATAACCATTA
[illegible]CAAAGACCGAATGCACGGCCTTGTA

GGGGACGCCCCAGATTGTCAGGGATGGGGGGATGGTCCTTGAGTTTTGCATGCTCTCCTCC
CTCCCACTTCTGCACCCTTTCACCACCTCGAGGAGATTTGCTCCCCATTAGCGAATGAAAT
TGATGCAGTCCTAAAAAAAAAAA

FIGURA 29

CAGCCCAGCACCTGCGGAGGGAGCGCTGACCATGGCTCCCTGGCCTGAATTGGGAGATGCC
CAGCCCAACCCCGATAAGTACCTCGAAGGGGCCGCAGGTCAGCAGCCCACTGCCCCTGATA
AAAGCAAAGAGACCAACAAAACAGATAACACTGAGGCACCTGTAACCAAGATTGAACTTCT
GCCGTCCTACTCCACGGCTACACTGATAGATGAGCCCACTGAGGTGGATGACCCCTGGAAC
CTACCCACTCTTCAGGACTCGGGGATCAAGTGGTCAGAGAGAGACACCAAAGGGAAGATTC
TCTGTTTCTTCCAAGGGATTGGGAGATTGATTTTACTTCTCGGATTTCTCTACTTTTTCGT
GTGCTCCCTGGATATTCTTAGTAGCGCCTTCCAGCTGGTTGGAGGAAAAATGGCAGGACAG
TTCTTCAGCAACAGCTCTATTATGTCCAACCCTTTGTTGGGGCTGGTGATCGGGGTGCTGG
TGACCGTCTTGGTGCAGAGCTCCAGCACCTCAACGTCCATCGTTGTCAGCATGGTGTCCTC
TTCATTGCTCACTGTTCGGGCTGCCATCCCCATTATCATGGGGGCCAACATTGGAACGTCA
ATCACCAACACTATTGTTGCGCTCATGCAGGTGGGAGATCGGAGTGAGTTCAGAAGAGCTT
TTGCAGGAGCCACTGTCCATGACTTCTTCAACTGGCTGTCCGTGTTGGTGCTCTTGCCCGT
GGAGGTGGCCACCCATTACCTCGAGATCATAACCCAGCTTATAGTGGAGAGCTTCCACTTC
AAGAATGGAGAAGATGCCCCAGATCTTCTGAAAGTCATCACTAAGCCCTTCACAAAGCTCA
TTGTCCAGCTGGATAAAAAAGTTATCAGCCAAATTGCAATGAACGATGAAAAAGCGAAAAA
CAAGAGTCTTGTCAAGATTTGGTGCAAAACTTTTACCAACAAGACCCAGATTAACGTCACT
GTTCCCTCGACTGCTAACTGCACCTCCCCTTCCCTCTGTTGGACGGATGGCATCCAAAACT
GGACCATGAAGAATGTGACCTACAAGGAGAACATCGCCAAATGCCAGCATATCTTTGTGAA
TTTCCACCTCCCGGATCTTGCTGTGGGCACCATCTTGCTCATACTCTCCCTGCTGGTCCTC
TGTGGTTGCCTGATCATGATTGTCAAGATCCTGGGCTCTGTGCTCAAGGGGCAGGTCGCCA
CTGTCATCAAGAAGACCATCAACACTGATTTCCCCTTTCCCTTTGCATGGTTGACTGGCTA
CCTGGCCATCCTCGTCGGGGCAGGCATGACCTTCATCGTACAGAGCAGCTCTGTGTTCACG
TCGGCCTTGACCCCCCTGATTGGAATCGGCGTGATAACCATTGAGAGGGCTTATCCACTCA
CGCTGGGCTCCAACATCGGCACCACCACCACCGCCATCCTGGCCGCCTTAGCCAGCCCTGG
CAATGCATTGAGGAGTTCACTCCAGATCGCCCTGTGCCACTTTTTCTTCAACATCTCCGGC
ATCTTGCTGTGGTACCCGATCCCGTTCACTCGCCTGCCCATCCGCATGGCCAAGGGGCTGG
GCAACATCTCTGCCAAGTATCGCTGGTTCGCCGTCTTCTACCTGATCATCTTCTTCTTCCT
GATCCCGCTGACGGTGTTTGGCCTCTCGCTGGCCGGCTGGCGGGTGCTGGTTGGTGTCGGG
GTTCCCGTCGTCTTCATCATCATCCTGGTACTGTGCCTCCGACTCCTGCAGTCTCGCTGCC
CACGCGTCCTGCCGAAGAAACTCCAGAACTGGAACTTCCTGCCGCTGTGGATGCGCTCGCT
GAAGCCCTGGGATGCCGTCGTCTCCAAGTTCACCGGCTGCTTCCAGATGCGCTGCTGCTGC
TGCTGCCGCGTGTGCTGCCGCGCGTGCTGCTTGCTGTGTGGCTGCCCCAAGTGCTGCCGCT
GCAGCAAGTGCTGCGAGGACTTGGAGGAGGCGCAGGAGGGGCAGGATGTCCCTGTCAAGGC
TCCTGAGACCTTTGATAACATAACCATTAGCAGAGAGGCTCAGGGTGAGGTCCCTGCCTCG
[illegible]CAGATTGTCAGGGATGGGGGGA

TGGTCCTTGAGTTTTGCATGCTCTCCTCCCTCCCACTTCTGCACCCTTTCACCACCTCGAG
GAGATTTGCTCCCCATTAGCGAATGAAATTGATGCAGTCCTACCTAACTCGATTCCCTTTG
GCTTGGTGGGTAGGCCTGCAGGGCACTTTTATTCCAACCCATGGCCTCCATGACTTTTTCA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURA 30

GGCGTGGGAGCTGCCTAGGCCGGGCCCTGCCAGGGAGCAAGTCTGCATCGCAGAGGCCAGG
GCAGAGTGGCTCCTCACAGCCTGAAGCTCATCCTTCTGCACGGGCCAGCCAGGCCAGCACA
GAGGCACCAGGGCAGCAGTGCACACAGGTCCCCGGGGACCCCACCATGTGGAGCGGATGGT
GGCTGTGGCCCCTTGTGGCCGTCTGCACTGCAGACTTCTTTCGGGACGAGGCAGAGAGGAT
CATGAGGGACTCCCCTGTCATTGATGGGCACAATGACCTCCCCTGGCAGCTGCTGGATATG
TTCAACAACCGGCTGCAGGACGAGAGGGCCAACCTGACCACCTTGGCCGGCACACACACCA
ACATCCCCAAGCTGAGGGCCGGCTTTGTGGGAGGCCAGTTCTGGTCCGTGTACACGCCCTG
CGACACCCAGAACAAAGACGCCGTGCGGAGGACGCTGGAGCAGATGGACGTGGTCCACCGC
ATGTGCCGGATGTACCCGGAGACCTTCCTGTATGTCACCAGCAGTGCAGGCATTCGGCAGG
CCTTCCGGGAAGGGAAGGTGGCCAGCCTGATCGGCGTGGAGGGCGGCCACTCCATTGACAG
CAGTTTGGGCGTCCTGCGGGCACTCTATCAGCTGGGCATGCGGTACCTGACCCTCACCCAC
AGCTGCAACACGCCCTGGGCTGACAACTGGCTGGTGGACACGGGAGACAGCGAGCCCCAGA
GCCAAGGCTTGTCACCCTTTGGGCAGCGTGTGGTGAAGGAGCTGAACCGTCTGGGGGTCCT
CATCGACTTGGCTCACGTGTCTGTGGCCACCATGAAGGCCACCCTGCAGCTGTCCAGAGCC
CCGGTCATCTTCAGCCACTCCTCGGCCTACAGCGTGTGCGCAAGCCGGCGCAACGTGCCTG
ACGACGTCCTGAGGCTGGTGAAACAGACAGACAGCCTGGTGATGGTGAACTTCTACAACAA
TTACATTTCCTGCACCAACAAGGCCAACCTGTCCCAAGTGGCCGACCATCTGGATCACATC
AAGGAGGTGGCAGGAGCCAGAGCCGTGGGTTTTGGTGGGGACTTTGATGGTGTTCCAAGGG
TCCCTGAGGGGCTGGAGGACGTCTCCAAGTATCCAGACCTGATCGCTGAGCTGCTCAGGAG
GAACTGGACGGAGGCGGAGGTCAAGGGCGCACTGGCTGACAACCTGCTGAGGGTCTTCCAG
GCTGTGGAACAGGCCAGCAACCTCACACAGGCTCCCGAGGAGGAGCCCATCCCGCTGGACC
AGCTGGGTGGCTCCTGCAGGACCCATTACGGCTACTCCTCTGGGGCTTCCAGCCTCCATCG
CCACTGGGGGCTCCTGCTGGCCTCCCTCGCTCCCCTGGTCCTCTGTCTGTCTCTCCTGTGA
AACCTGGGAGACCAGAGTCCCCTTTAGGGTTCCCGGAGCTCCGGGAAGACCCGCCCATCCC
AGGACTCCAGATGCCAGGAGCCCTGCTGCCCACATGCAAGGACCAGCATCTCCTGAGAGGA
CGCCTGGGCTTACCTGGGGGGCAGGATGCCTGGGGACAGTTCAGGACACACACACAGTAGG
CCCGCAATAAAAGCAACACCCCTTCAAAAAAAAAAAAAAAAAAAAA

FIGURA 31

CATCTAAAGCCTCCTCAGCCTTCTGAGTCAGCCTGAAAGGAACAGGCCGAACTGCTGTATG
GGCTCTACTGCCAGTGTGACCTCACCCTCTCCAGTCACCCCTCCTCAGTTCCAGCTATGAG
TTCCTGCAACTTCACACATGCCACCTTTGTGCTTATTGGTATCCCAGGATTAGAGAAAGCC
CATTTCTGGGTTGGCTTCCCCCTCCTTTCCATGTATGTAGTGGCAATGTTTGGAAACTGCA
TCGTGGTCTTCATCGTAAGGACGGAACGCAGCCTGCACGCTCCGATGTACCTCTTTCTCTG
CATGCTTGCAGCCATTGACCTGGCCTTATCCACATCCACCATGCCTAAGATCCTTGCCCTT
TTCTGGTTTGATTCCCGAGAGATTAGCTTTGAGGCCTGTCTTACCCAGATGTTCTTTATTC
ATGCCCTCTCAGCCATTGAATCCACCATCCTGCTGGCCATGGCCTTTGACCGTTATGTGGC
CATCTGCCACCCACTGCGCCATGCTGCAGTGCTCAACAATACAGTAACAGCCCAGATTGGC
ATCGTGGCTGTGGTCCGCGGATCCCTCTTTTTTTTCCCACTGCCTCTGCTGATCAAGCGGC
TGGCCTTCTGCCACTCCAATGTCCTCTCGCACTCCTATTGTGTCCACCAGGATGTAATGAA
GTTGGCCTATGCAGACACTTTGCCCAATGTGGTATATGGTCTTACTGCCATTCTGCTGGTC
ATGGGCGTGGACGTAATGTTCATCTCCTTGTCCTATTTTCTGATAATACGAACGGTTCTGC
AACTGCCTTCCAAGTCAGAGCGGGCCAAGGCCTTTGGAACCTGTGTGTCACACATTGGTGT
GGTACTCGCCTTCTATGTGCCACTTATTGGCCTCTCAGTTGTACACCGCTTTGGAAACAGC
CTTCATCCCATTGTGCGTGTTGTCATGGGTGACATCTACCTGCTGCTGCCTCCTGTCATCA
ATCCCATCATCTATGGTGCCAAAACCAAACAGATCAGAACACGGGTGCTGGCTATGTTCAA
GATCAGCTGTGACAAGGACTTGCAGGCTGTGGGAGGCAAGTGACCCTTAACACTACACTTC
TCCTTATCTTTATTGGCTTGATAAACATAATTATTTCTAACACTAGCTTATTTCCAGTTGC
CCATAAGCACATCAGTACTTTTCTCTGGCTGGAATAGTAAACTAAAGTATGGTACATCTAC
CTAAAGGACTATTATGTGGAATAATACATACTAATGAAGTATTACATGATTTAAAGACTAC
AATAAAACCAAACATGCTTATAACATTAAGAAAAACAATAAAGATACATGATTGAAACCAA
GTTGAAAAATAGCATATGCCTTGGAGGAAATGTGCTCAAATTACTAATGATTTAGTGTTGT
CCCTACTTTCTCTCTCTTTTTTCTTTCTTTTTTTTTTATTATGGTTAGCTGTCACATACAA
CTTTTTTTTTTTTTTGAGATGGGGTCTCCAGCCTGGGCAACAGAGCAAGACCCTGTCTCAAA
GCATAAAATGGAATAACATATCAAATGAAACAGGGAAAATGAAGCTGACAATTTATGGGAG
CCA

FIGURA 32A

CACACATACGCACGCACGATCTCACTTCGATCTATACACTGGAGGATTAAAACAAACAAAC
AAAAAAAACATTTCCTTCGCTCCCCCTCCCTCTCCACTCTGAGAAGCAGAGGAGCCGCACG
GCGAGGGGCCGCAGACCGTCTGGAAATGCGAATCCTAAAGCGTTTCCTCGCTTGCATTCAG
CTCCTCTGTGTTTGCCGCCTGGATTGGGCTAATGGATACTACAGACAACAGAGAAAACTTG
TTGAAGAGATTGGCTGGTCCTATACAGGAGCACTGAATCAAAAAAATTGGGGAAAGAAATA
TCCAACATGTAATAGCCCAAAACAATCTCCTATCAATATTGATGAAGATCTTACACAAGTA
AATGTGAATCTTAAGAAACTTAAATTTCAGGGTTGGGATAAAACATCATTGGAAAACACAT
TCATTCATAACACTGGGAAAACAGTGGAAATTAATCTCACTAATGACTACCGTGTCAGCGG
AGGAGTTTCAGAAATGGTGTTTAAAGCAAGCAAGATAACTTTTCACTGGGGAAAATGCAAT
ATGTCATCTGATGGATCAGAGCATAGTTTAGAAGGACAAAAATTTCCACTTGAGATGCAAA
TCTACTGCTTTGATGCGGACCGATTTTCAAGTTTTGAGGAAGCAGTCAAAGGAAAAGGGAA
GTTAAGAGCTTTATCCATTTTGTTTGAGGTTGGGACAGAAGAAAATTTGGATTTCAAAGCG
ATTATTGATGGAGTCGAAAGTGTTAGTCGTTTTGGGAAGCAGGCTGCTTTAGATCCATTCA
TACTGTTGAACCTTCTGCCAAACTCAACTGACAAGTATTACATTTACAATGGCTCATTGAC
ATCTCCTCCCTGCACAGACACAGTTGACTGGATTGTTTTTAAAGATACAGTTAGCATCTCT
GAAAGCCAGTTGGCTGTTTTTTGTGAAGTTCTTACAATGCAACAATCTGGTTATGTCATGC
TGATGGACTACTTACAAAACAATTTTCGAGAGCAACAGTACAAGTTCTCTAGACAGGTGTT
TTCCTCATACACTGGAAAGGAAGAGATTCATGAAGCAGTTTGTAGTTCAGAACCAGAAAAT
GTTCAGGCTGACCCAGAGAATTATACCAGCCTTCTTGTTACATGGGAAAGACCTCGAGTCG
TTTATGATACCATGATTGAGAAGTTTGCAGTTTTGTACCAGCAGTTGGATGGAGAGGACCA
AACCAAGCATGAATTTTTGACAGATGGCTATCAAGACTTGGGTGCTATTCTCAATAATTTG
CTACCCAATATGAGTTATGTTCTTCAGATAGTAGCCATATGCACTAATGGCTTATATGGAA
AATACAGCGACCAACTGATTGTCGACATGCCTACTGATAATCCTGAACTTGATCTTTTCCC
TGAATTAATTGGAACTGAAGAAATAATCAAGGAGGAGGAAGAGGGAAAAGACATTGAAGAA
GGCGCTATTGTGAATCCTGGTAGAGACAGTGCTACAAACCAAATCAGGAAAAAGGAACCCC
AGATTTCTACCACAACACACTACAATCGCATAGGGACGAAATACAATGAAGCCAAGACTAA
CCGATCCCCAACAAGAGGAAGTGAATTCTCTGGAAAGGGTGATGTTCCCAATACATCTTTA
AATTCCACTTCCCAACCAGTCACTAAATTAGCCACAGAAAAAGATATTTCCTTGACTTCTC
AGACTGTGACTGAACTGCCACCTCACACTGTGGAAGGTACTTCAGCCTCTTTAAATGATGG
CTCTAAAACTGTTCTTAGATCTCCACATATGAACTTGTCGGGGACTGCAGAATCCTTAAAT
ACAGTTTCTATAACAGAATATGAGGAGGAGAGTTTATTGACCAGTTTCAAGCTTGATACTG
GAGCTGAAGATTCTTCAGGCTCCAGTCCCGCAACTTCTGCTATCCCATTCATCTCTGAGAA
CATATCCCAAGGGTATATATTTTCCTCCGAAAACCCAGAGACAATAACATATGATGTCCTT
ATACCAGAATCTGCTAGAAATGCTTCCGAAGATTCAACTTCATCAGGTTCAGAAGAATCAC
TAAAGGATCCTTTCTATGGAGGGAAATGTGTGGTTTCCTAGCTCTACAGACATAACAGCACA

GCCCGATGTTGGATCAGGCAGAGAGAGCTTTCTCCAGACTAATTACACTGAGATACGTGTT
GATGAATCTGAGAAGACAACCAAGTCCTTTTCTGCAGGCCCAGTGATGTCACAGGGTCCCT
CAGTTACAGATCTGGAAATGCCACATTATTCTACCTTTGCCTACTTCCCAACTGAGGTAAC
ACCTCATGCTTTTACCCCATCCTCCAGACAACAGGATTTGGTCTCCACGGTCAACGTGGTA
TACTCGCAGACAACCCAACCGGTATACAATGGTGAGACACCTCTTCAACCTTCCTACAGTA
GTGAAGTCTTTCCTCTAGTCACCCCTTTGTTGCTTGACAATCAGATCCTCAACACTACCCC
TGCTGCTTCAAGTAGTGATTCGGCCTTGCATGCTACGCCTGTATTTCCCAGTGTCGATGTG
TCATTTGAATCCATCCTGTCTTCCTATGATGGTGCACCTTTGCTTCCATTTTCCTCTGCTT
CCTTCAGTAGTGAATTGTTTCGCCATCTGCATACAGTTTCTCAAATCCTTCCACAAGTTAC
TTCAGCTACCGAGAGTGATAAGGTGCCCTTGCATGCTTCTCTGCCAGTGGCTGGGGGTGAT
TTGCTATTAGAGCCCAGCCTTGCTCAGTATTCTGATGTGCTGTCCACTACTCATGCTGCTT
CAGAGACGCTGGAATTTGGTAGTGAATCTGGTGTTCTTTATAAAACGCTTATGTTTTCTCA
AGTTGAACCACCCAGCAGTGATGCCATGATGCATGCACGTTCTTCAGGGCCTGAACCTTCT
TATGCCTTGTCTGATAATGAGGGCTCCCAACACATCTTCACTGTTTCTTACAGTTCTGCAA
TACCTGTGCATGATTCTGTGGGTGTAACTTATCAGGGTTCCTTATTTAGCGGCCCTAGCCA
TATACCAATACCTAAGTCTTCGTTAATAACCCCAACTGCATCATTACTGCAGCCTACTCAT
GCCCTCTCTGGTGATGGGGAATGGTCTGGAGCCTCTTCTGATAGTGAATTTCTTTTACCTG
ACACAGATGGGCTGACAGCCCTTAACATTTCTTCACCTGTTTCTGTAGCTGAATTTACATA
TACAACATCTGTGTTTGGTGATGATAATAAGGCGCTTTCTAAAAGTGAAATAATATATGGA
AATGAGACTGAACTGCAAATTCCTTCTTTCAATGAGATGGTTTACCCTTCTGAAAGCACAG
TCATGCCCAACATGTATGATAATGTAAATAAGTTGAATGCGTCTTTACAAGAAACCTCTGT
TTCCATTTCTAGCACCAAGGGCATGTTTCCAGGGTCCCTTGCTCATACCACCACTAAGGTT
TTTGATCATGAGATTAGTCAAGTTCCAGAAAATAACTTTTCAGTTCAACCTACACATACTG
TCTCTCAAGCATCTGGTGACACTTCGCTTAAACCTGTGCTTAGTGCAAACTCAGAGCCAGC
ATCCTCTGACCCTGCTTCTAGTGAAATGTTATCTCCTTCAACTCAGCTCTTATTTTATGAG
ACCTCAGCTTCTTTTAGTACTGAAGTATTGCTACAACCTTCCTTTCAGGCTTCTGATGTTG
ACACCTTGCTTAAAACTGTTCTTCCAGCTGTGCCCAGTGATCCAATATTGGTTGAAACCCC
CAAAGTTGATAAAATTAGTTCTACAATGTTGCATCTCATTGTATCAAATTCTGCTTCAAGT
GAAAACATGCTGCACTCTACATCTGTACCAGTTTTTGATGTGTCGCCTACTTCTCATATGC
ACTCTGCTTCACTTCAAGGTTTGACCATTTCCTATGCAAGTGAGAAATATGAACCAGTTTT
GTTAAAAAGTGAAAGTTCCCACCAAGTGGTACCTTCTTTGTACAGTAATGATGAGTTGTTC
CAAACGGCCAATTTGGAGATTAACCAGGCCCATCCCCCAAAAGGAAGGCATGTATTTGCTA
CACCTGTTTTATCAATTGATGAACCATTAAATACACTAATAAATAAGCTTATACATTCCGA
TGAAATTTTAACCTCCACCAAAAGTTCTGTTACTGGTAAGGTATTTGCTGGTATTCCAACA
GTTGCTTCTGATACATTTGTATCTACTGATCATTCTGTTCCTATAGGAAATGGGCATGTTG
CCATTACAGCTGTTTCTCCCCACAGAGATGGTTCTGTAACCTCAACAAAGTTGCTGTTTCC

GGTGAAGATGGTGACACTGATGATGATGGTGATGATGATGATGACAGAGATAGTGATGGCT
TATCCATTCATAAGTGTATGTCATGCTCATCCTATAGAGAATCACAGGAAAAGGTAATGAA

FIGURA 32B

TGATTCAGACACCCACGAAAACAGTCTTATGGATCAGAATAATCCAATCTCATACTCACTA
TCTGAGAATTCTGAAGAAGATAATAGAGTCACAAGTGTATCCTCAGACAGTCAAACTGGTA
TGGACAGAAGTCCTGGTAAATCACCATCAGCAAATGGGCTATCCCAAAAGCACAATGATGG
AAAAGAGGAAAATGACATTCAGACTGGTAGTGCTCTGCTTCCTCTCAGCCCTGAATCTAAA
GCATGGGCAGTTCTGACAAGTGATGAAGAAAGTGGATCAGGGCAAGGTACCTCAGATAGCC
TTAATGAGAATGAGACTTCCACAGATTTCAGTTTTGCAGACACTAATGAAAAAGATGCTGA
TGGGATCCTGGCAGCAGGTGACTCAGAAATAACTCCTGGATTCCCACAGTCCCCAACATCA
TCTGTTACTAGCGAGAACTCAGAAGTGTTCCACGTTTCAGAGGCAGAGGCCAGTAATAGTA
GCCATGAGTCTCGTATTGGTCTAGCTGAGGGGTTGGAATCCGAGAAGAAGGCAGTTATACC
CCTTGTGATCGTGTCAGCCCTGACTTTTATCTGTCTAGTGGTTCTTGTGGGTATTCTCATC
TACTGGAGGAAATGCTTCCAGACTGCACACTTTTACTTAGAGGACAGTACATCCCCTAGAG
TTATATCCACACCTCCAACACCTATCTTTCCAATTTCAGATGATGTCGGAGCAATTCCAAT
AAAGCACTTTCCAAAGCATGTTGCAGATTTACATGCAAGTAGTGGGTTTACTGAAGAATTT
GAGACACTGAAAGAGTTTTACCAGGAAGTGCAGAGCTGTACTGTTGACTTAGGTATTACAG
CAGACAGCTCCAACCACCCAGACAACAAGCACAAGAATCGATACATAAATATCGTTGCCTA
TGATCATAGCAGGGTTAAGCTAGCACAGCTTGCTGAAAAGGATGGCAAACTGACTGATTAT
ATCAATGCCAATTATGTTGATGGCTACAACAGACCAAAAGCTTATATTGCTGCCCAAGGCC
CACTGAAATCCACAGCTGAAGATTTCTGGAGAATGATATGGGAACATAATGTGGAAGTTAT
TGTCATGATAACAAACCTCGTGGAGAAAGGAAGGAGAAAATGTGATCAGTACTGGCCTGCC
GATGGGAGTGAGGAGTACGGGAACTTTCTGGTCACTCAGAAGAGTGTGCAAGTGCTTGCCT
ATTATACTGTGAGGAATTTTACTCTAAGAAACACAAAAATAAAAAAGGGCTCCCAGAAAGG
AAGACCCAGTGGACGTGTGGTCACACAGTATCACTACACGCAGTGGCCTGACATGGGAGTA
CCAGAGTACTCCCTGCCAGTGCTGACCTTTGTGAGAAAGGCAGCCTATGCCAAGCGCCATG
CAGTGGGGCCTGTTGTCGTCCACTGCAGTGCTGGAGTTGGAAGAACAGGCACATATATTGT
GCTAGACAGTATGTTGCAGCAGATTCAACACGAAGGAACTGTCAACATATTTGGCTTCTTA
AAACACATCCGTTCACAAAGAAATTATTTGGTACAAACTGAGGAGCAATATGTCTTCATTC
ATGATACACTGGTTGAGGCCATACTTAGTAAAGAAACTGAGGTGCTGGACAGTCATATTCA
TGCCTATGTTAATGCACTCCTCATTCCTGGACCAGCAGGCAAAACAAAGCTAGAGAAACAA
TTCCAGCTCCTGAGCCAGTCAAATATACAGCAGAGTGACTATTCTGCAGCCCTAAAGCAAT
GCAACAGGGAAAAGAATCGAACTTCTTCTATCATCCCTGTGGAAAGATCAAGGGTTGGCAT
TTCATCCCTGAGTGGAGAAGGCACAGACTACATCAATGCCTCCTATATCATGGGCTATTAC
CAGAGCAATGAATTCATCATTACCCAGCACCCTCTCCTTCATACCATCAAGGATTTCTGGA
GGATGATATGGGACCATAATGCCCAACTGGTGGTTATGATTCCTGATGGCCAAAACATGGC
ACAAGATGAATTTTCTTTACTGGCCAAATAAAGATGAGCCTATAAATTGTGAGAGCTTTAAG

GTCACTCTTATGGCTGAAGAACACAAATGTCTATCTAATGAGGAAAAACTTATAATTCAGG
ACTTTATCTTAGAAGCTACACAGGATGATTATGTACTTGAAGTGAGGCACTTTCAGTGTCC
TAAATGGCCAAATCCAGATAGCCCCATTAGTAAAACTTTTGAACTTATAAGTGTTATAAAA
GAAGAAGCTGCCAATAGGGATGGGCCTATGATTGTTCATGATGAGCATGGAGGAGTGACGG
CAGGAACTTTCTGTGCTCTGACAACCCTTATGCACCAACTAGAAAAAGAAAATTCCGTGGA
TGTTTACCAGGTAGCCAAGATGATCAATCTGATGAGGCCAGGAGTCTTTGCTGACATTGAG
CAGTATCAGTTTCTCTACAAAGTGATCCTCAGCCTTGTGAGCACAAGGCAGGAAGAGAATC
CATCCACCTCTCTGGACAGTAATGGTGCAGCATTGCCTGATGGAAATATAGCTGAGAGCTT
AGAGTCTTTAGTTTAACACAGAAAGGGGTGGGGGGACTCACATCTGAGCATTGTTTTCCTC
TTCCTAAAATTAGGCAGGAAAATCAGTCTAGTTCTGTTATCTGTTGATTTCCCATCACCTG
ACAGTAACTTTCATGACATAGGATTCTGCCGCCAAATTTATATCATTAACAATGTGTGCCT
TTTTGCAAGACTTGTAATTTACTTATTATGTTTGAACTAAAATGATTGAATTTTACAGTAT
TTCTAAGAATGGAATTGTGGTATTTTTTTCTGTATTGATTTTAACAGAAAATTTCAATTTA
TAGAGGTTAGGAATTCCAAACTACAGAAAATGTTTGTTTTTAGTGTCAAATTTTTAGCTGT
ATTTGTAGCAATTATCAGGTTTGCTAGAAATATAACTTTTAATACAGTAGCCTGTAAATAA
AACACTCTTCCATATGATATTCAACATTTTACAACTGCAGTATTCACCTAAAGTAGAAATA
ATCTGTTACTTATTGTAAATACTGCCCTAGTGTCTCCATGGACCAAATTTATATTTATAAT
TGTAGATTTTTATATTTTACTACTGAGTCAAGTTTTCTAGTTCTGTGTAATTGTTTAGTTT
AATGACGTAGTTCATTAGCTGGTCTTACTCTACCAGTTTTCTGACATTGTATTGTGTTACC
TAAGTCATTAACTTTGTTTCAGCATGTAATTTTAACTTTTGTGGAAAATAGAAATACCTTC
ATTTTGAAAGAAGTTTTTATGAGAATAACACCTTACCAAACATTGTTCAAATGGTTTTTAT
CCAAGGAATTGCAAAAATAAATATAAATATTGCCATTAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAA

FIGURA 33

ATTGTCTGGGAATTGCAGCCGCGGGGCGGGCGGCGGCGGCGGCGGCGGCCGGGACCCA
GCGGGCCAGGTGGGGACGGCGCGGAGCGGGTGCGGGAGATGCCGTGCGGGACTGGGGCCAC
CTGAGCCGCCCGCCTCGTCCCCGCCTTCTGTGGGAAGGATGTGCGCGCGGATGGCCGGTCG
CACAAGAGCGGCCCCTCGGGGGCCCTACGGCCCCTGGCTCTGCCTCCTGGTGGCCCTCGCC
CTGGACGTCGTGAGAGTGGACTGTGGCCAGGCTCCCCTGGACCCTGTCTACCTGCATGTGA
CAGCCGCCCGCCCAGCCCAGCCCACACTCTGGACTGCCAAGCTAGACCGCTTCAAGGGCTC
CAGGCACCACACCACCCTCATCACCTGCCACCGTGCTGGGCTCACAGAGCCAGATTCCAGC
AGTCCCCTTGAACTGTCTGAGTTCCTATGGGTGGACTTTGTGGTGGAGAATAGCACTGGTG
GGGGCGTAGCGGTCACTCGCCCCGTCACGTGGCAGCTGGAGTACCCAGGCCAGGCCCCTGA
AGCAGAGAAGGACAAAATGGTGTGGGAAATCCTGGTGTCTGAGCGGGACATCAGAGCCCTT
ATCCCACTGGCCAAGGCTGAGGAGCTGGTGAATACAGCACCACTGACTGGAGTGCCCCAGC
ATGTCCCCGTGCGCCTTGTCACTGTGGACGGCGGGGGGGCCTTGGTGGAGGTGACAGAGCA
TGTCGGCTGCGAGTCTGCCAACACACAGGTCCTGCAGGTGTCTGAGGCCTGTGATGCCGTG
TTCGTGGCTGGCAAGGAGAGCCGGGGCGCCCGGGGGGTGCGAGTGGACTTCTGGTGGCGCC
GGCTCCGCGCCTCGCTGCGGCTGACCGTGTGGGCCCCGCTGCTACCGCTGCGTATCGAGCT
CACCGACACCACCCTCGAGCAGGTCCGCGGCTGGAGGGTACCTGGCCCTGCTGAAGGGCCT
GCGGAACCCGCTGCAGAGGCGTCAGATGAGGCCGAGCGGCGCGCCCGTGGCTGCCACCTGC
AGTACCAGCGGGCCGGTGTGCGCTTCCTCGCCCCCTTCGCGGCCCACCCGCTGGACGGCGG
CCGCCGCCTCACGCACCTGCTTGGCCCCGACTGGCTGCTAGACGTGTCCCACCTCGTGGCG
CCACACGCCCGCGTGCTGGACTCGCGTGTAGCCTCTCTGGAGGGTGGCCGTGTCGTGGTGG
GCCGGGAGCCCGGTGTCACCTCCATTGAGGTGCGTTCCCCACTGTCTGACTCCATCCTGGG
GGAGCAGGCGCTGGCTGTGACGGACGACAAGGTCTCAGTGCTGGAGCTGAGGGTGCAGCCA
GTGATGGGCATCTCGCTGACCTTGAGCCGGGGCACTGCCCACCCCGGGGAGGTCACAGCTA
CGTGCTGGGCACAGTCAGCCCTTCCCGCCCCAAAGCAGGAGGTGGCCCTCTCCCTATGGCT
GTCCTTCTCTGATCACACTGTGGCCCCAGCTGAGCTCTACGACCGCCGTGACCTGGGACTG
TCCGTCTCAGCCGAGGAGCCTGGTGCCATCCTGCCAGCTGAGGAGCAGGGTGCCCAGCTCG
GGGTGGTGGTGAGTGGGGCAGGCGCCGAGGGGCTGCCGCTGCATGTGGCTCTGCACCCGCC
CGAGCCCTGCCGCCGGGGCCGCCACCGTGTGCCTCTGGCCTCTGGCACCGCCTGGCTGGGG
CTGCCCCCTGCCTCCACTCCAGCCCCTGCTCTCCCATCCAGCCCTGCTTGGAGCCCACCAG
CCACAGAAGCCACCATGGGTGGTAAACGGCAGGTGGCAGGCAGTGTCGGGGGCAACACAGG
TGTGAGGGGCAAGTTTGAGCGGGCAGAGGAGGAGGCCAGGAAGGAGGAGACCGAAGCCAGG
GAGGAGGAGGAGGAAGAGGAGGAGGAGATGGTCCCTGCCCCTCAGCATGTCACTGAGCTAG
AGCTGGGCATGTACGCCCTGCTGGGAGTCTTCTGCGTGGCCATCTTCATCTTCTTGGTCAA
TGGTGTGGTCTTCGTCCTGCGCTATCAGCGCAAAGAACCTCCCGACAGTGCCACTGACCCC

AGCTGGACCGGCAGTCCCCTGGCCCGCCCAAGGGGGAGGGGAGCTGCCCCTGTGAGAGTGG
GGGAGGAGGGGAGGCCCCTACCCTGGCCCCTGGCCCTCCTGGGGGCACCACCAGCTCCTCA
AGCACCCTGGCCCGAAAGGAGGCTGGGGGGCGGCGGAAGCGAGTAGAGTTTGTGACATTTG
TGCCAGCCCCTCCAGCCCAGTCACCTGAGGAGCCTGTAGGGGCCCCTGCTGTGCAGTCCAT
CCTTGTGGCAGGCGAGGAGGACATCCGCTGGGTGTGTGAGGACATGGGGCTGAAGGACCCT
GAGGAGCTTCGCAACTACATGGAGAGGATCCGGGGCAGCTCCTGACCCTCCACAGCCACCT
GGTCAGCCACCAGCTGGGGCAACGAGGGTGGAGGTCCCACTGAGCCTCTCGCCTGCCCCCG
CCACTCGTCTGGTGCTTGTTGATCCAAGTCCCCTGCCTGGTCCCCCACAAGGACTCCCATC
CAGGCCCCCTCTGCCCTGCCCCTTGTCATGGACCATGGTCGTGAGGAAGGGCTCATGCCCC
TTATTTATGGGAACCATCTCATTCTAACAGAATAAACCGAGAAGGAAACCAGAAAAAAAAA
AAAAAAAAAA

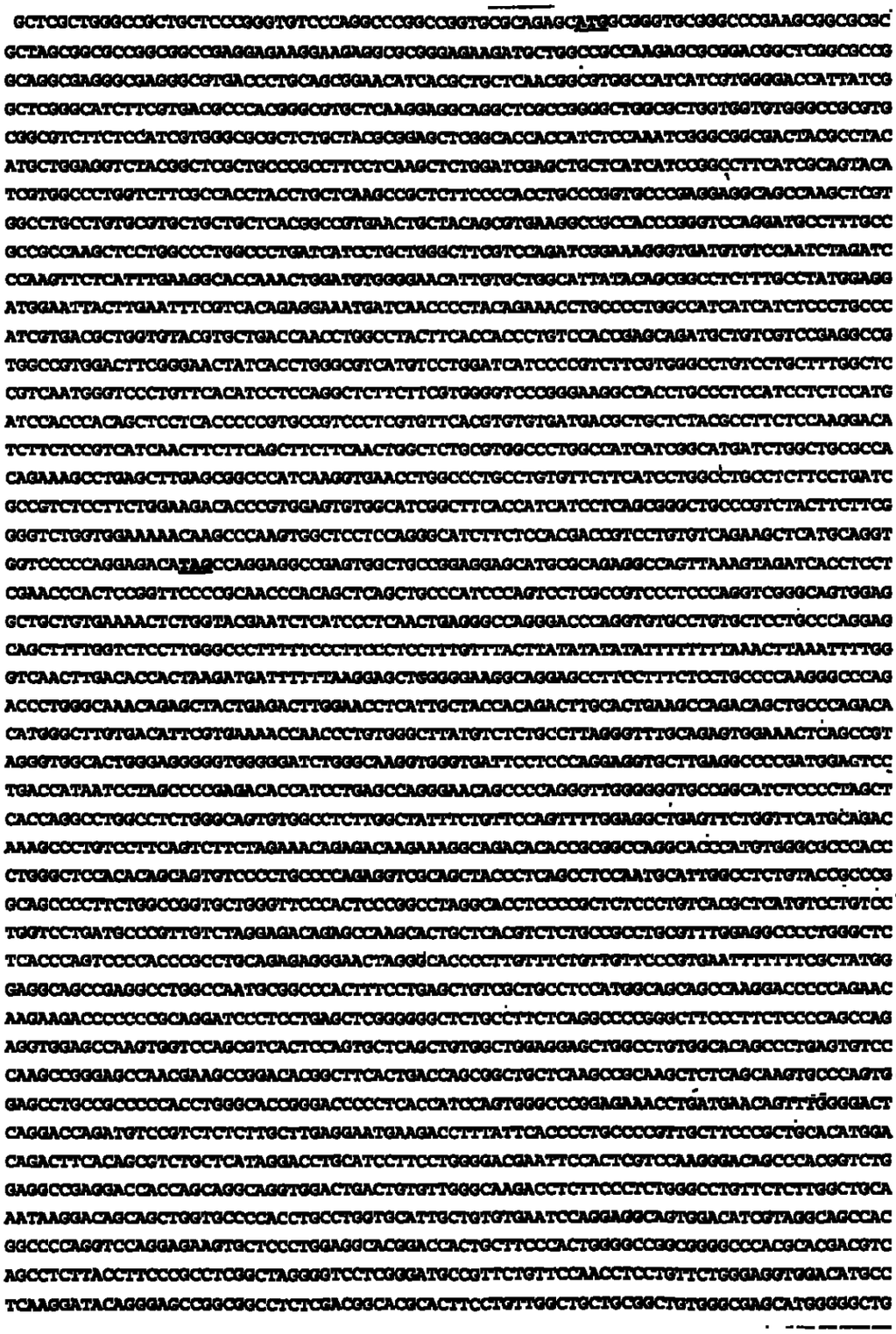
FIGURA 34

ACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGGTCGGGGCCAGAGCGCAGGTGTACCTGG
CGGCCGTGCTGGAGCACCTGACCGCCGAGATCCTGGAGCTGGCTGGCAACCCGGCCCGCGA
CAAGAAGACCCGCATCATCCTGCGCCACCTGTAGCTGGCCATTCGCAACGGCGAGGAGCTT
AACAAGCTGCTGGGCGAAGTCACCATCGCGCAGGGCGGTGTCCTGCCCAACATTCAGGGCG
TGCTTCTGCCCCAGAAGACCAAGAGCCACCACAAGGCCAAGGGTGAAAACCATTCACTAGG
AGAGGAGAAACACAATGGCCACCAAGACAGAGTTGAGTCCCACAGCAAGGGAGAGCAAGAA
CGCACAAGATATGCAAGTGGATGAGACACTGATCCCCAGGAAAGGTCCAAGTTTATGTTCT
GCTCGCTATGGAATAGCCCTCGTCTTACATTTCTGCAATTTCACAACGATAGCACAAAATG
TCATCATGAACATCACCATGGTAGCCATGGTCAACAGCACAAGCCCTCAATCCCAGCTCAA
TGATTCCTCTGAGGTGCTGCCTGTTGACTCATTTGGTGGCCTAAGTAAAGCCCCAAAGAGT
CTTCCTGCAAAGTCCTCAATACTTGGGGGTCAGTTTGCAATTTGGGAAAAGTGGGGCCCTC
CACAAGAACGAAGCAGACTCTGCAGCATTGCTTTATCAGGAATGTTACTGGGATGCTTTAC
TGCCATCCTCATAGGTGGCTTCATTAGTGAAACCCTTGGGTGGCCCTTTGTCTTCTATATC
TTTGGAGGTGTTGGCTGTGTCTGCTGCCTTCTCTGGTTTGTTGTGATTTATGATGACCCCT
TTTCCTATCCATGGATAAGCACCTCAGAAAAAGAATACATCATATCCTCCTTGAAACAACA
GGTCGGGTCTTCTAAGCAGCCTCTTCCCATCAAAGCTATGCTCAGATCTCTACCCATTTGG
TCCATATGTTTAGGCTGTTTCAGCCATCAATGGTTAGTTAGCACAATGGTTGTATACATAC
CAACTTACATCAGCTCTGTGTACCATGTTAACATCAGAGACAATGGACTTCTATCTGCCCT
TCCTTTTATTGTTGCCTGGGTCATAGGCATGGTGGGAGGCTATCTGGCAGATTTCCTTCTA
ACCAAAAAGTTTAGACTCATCACTGTGAGGAAAATTGCCACAATTTTAGGAAGTCTCCCCT
CTTCAGCACTCATTGTGTCTCTGCCTTACCTCAATTCCGGCTATATCACAGCAACTGCCTT
GCTGACGCTCTCTTGCGGATTAAGCACATTGTGTCAGTCAGGGATTTATATCAATGTCTTA
GATATTGCTCCAAGGTATTCCAGTTTTCTCATGGGAGCATCAAGAGGATTTTCGAGCATAG
CACCTGTCATTGTACCCACTGTCAGCGGATTTCTTCTTAGTCAGGACCCTGAGTTTGGGTG
GAGGAATGTCTTCTTCTTGCTGTTTGCCGTTAACCTGTTAGGACTACTCTTCTACCTCATA
TTTGGAGAAGCAGATGTCCAAGAATGGGCTAAAGAGAGAAAACTCACTCGTTTATGAAGTT
ATCCCACCTTGGATGGAAAAGTCATTAGGCACCGTATTGCATAAAATAGAAGGCTTCCGTG
ATGAAAATACCAGTGAAAAGATTTTTTTTTCCTGTGGCTCTTTTCAATTATGAGATCAGTT
CATTATTTTATTCAGACTTTTTTTTGAGAGAAATGTAAGATGAATAAAAATTCAAATAAAA
TGATAACTAAGAAAAAAAAAAAAAA

FIGURA 35

GTGAAAACCATTCACTAGGAGAGGAGAAACACAATGGCCACCAAGACAGAGTTGAGTCCCA
CAGCAAGGGAGAGCAAGAACGCACAAGATATGCAAGTGGATGAGACACTGATCCCCAGGAA
AGTTCCAAGTTTATGTTCTGCTCGCTATGGAATAGCCCTCGTCTTACATTTCTGCAATTTC
ACAACGATAGCACAAAATGTCATCATGAACATCACCATGGTAGCCATGGTCAACAGCACAA
GCCCTCAATCCCAGCTCAATGATTCCTCTGAGGTGCTGCCTGTTGACTCATTTGGTGGCCT
AAGTAAAGCCCCAAAGAGTCTTCCTACAAAGTCCTCAATACTTGGGGGTCAGTTTGCAATT
TGGGAAAAGTGGGGCCCTCCACAAGAACGAAGCAGACTCTGCAGCATTGCTTTATCAGGAA
TGTTACTGGGATGCTTTACTGCCATCCTCATAGGTGGCTTCATTAGTGAAACCCTTGGGTG
GCCCTTTGTCTTCTATATCTTTGGAGGTGTTGGCTGTGTCTGCTGCCTTCTCTGGTTTGTT
GTGATTTATGATGACCCCGTTTCCTATCCATGGATAAGCACCTCAGAAAAAGAATACATCA
TATCCTCCTTGAAACAACAGGTCAGGTCTTCTAAGCAGCCTCTTCCCATCAAAGCTATGCT
CAGATCTCTACCCATTTGGTCCATATGTTTAGGCTGTTTCAGCCATCAATGGTTAGTTAGC
ACAATGGTTGTATACATACCAACTTACATCAGCTCTGTGTACCATGTTAACATCAGAGACA
ATGGACTTCTATCTGCCCTTCCTTTTATTGTTGCCTGGGTCATAGGCATGGTGGGAGGCTA
TCTGGCAGATTTCCTTCTAACCAAAAAGTTTAGACTCATCACTGTGAGGAAAATTGCCACA
ATTTTAGGAAGTCTCCCCTCTTCAGCACTCATTGTGTCTCTGCCTTACCTCAATTCCGGCT
ATATCACAGCAACTGCCTTGCTGACGCTCTCTTGCGGATTAAGCACATTGTGTCAGTCAGG
GATTTATATCAATGTCTTAGATATTGCTCCAAGGTATTCCAGTTTTCTCATGGGAGCATCA
AGAGGATTTTCGAGCATAGCACCTGTCATTGTACCCACTGTCAGCGGATTTCTTCTTAGTC
AGGACCCTGAGTTTGGGTGGAGGAATGTCTTCTTCTTGCTGTTTGCCGTTAACCTGTTAGG
ACTACTCTTCTACCTCATATTTGGAGAAGCAGATGTCCAAGAATGGGCTAAAGAGAGAAAA
CTCACTCGTTTATGAAGTTATCCCACCTTGGATGGAAAAGTCATTAGGCACCGTATTGCAT
AAAATAGAAGGCTTCCGTGATGAAAATACCAGTGAAAAGATTTTTTTTTCCTGTGGCTCTT
TTCAATTATGAGATCAGTTCATTATTTTATTCAGACTTTTTTTTGAGAGAAATGTAAGATG
AATAAAAATTCAAATAAAATGATAACTAAGAAAAAAAAAAAAAA

FIGURA 36



GGGCGTCTGAGCTGGGGTGGGAGCTCCTCCGCAGAAGGTTGGTGGGGGGTCCAGTCTGTGATCCTTGGTGCTGTGTGCCCCA
CTCCAGCCTGGGGACCCCACTTCAGAAGGTAGGGGCCGTGTCCCGCGGTGCTGACTGAGGCCTGCTTCCCCCTCCCCCTCCT
GCTGTGCTGGAATTCCACAGGGACCAGGGCCACCGCAGGGGACTGTCTCAGAAGACTTGATTTTTCCGTCCCTTTTTCTCCA
CACTCCACTGACAAACGTCCCCAGCGGTTTCCACTTGTGGGCTTCAGGTGTTTTCAAGCACAACCCACCACAACAAGCAAGT
GCATTTTCAGTCGTTGTGCTTTTTTGTTTTGTGCTAACGTCTTACTAATTTAAAGATGCTGTCGGCACCATGTTTATTTATT
TCCAGTGGTCATGCTCAGCCTTGCTGCTCTGCGTGGCGCAGGTGCCATGCCTGCTCCCTGTCTGTGTCCCAGCCACGCAGGG
CCATCCACTGTGACGTCGGCCGACCAGGCTGGACACCCTCTGCCGAGTAATGACGTGTGTGGCTGGGACCTTCTTTATTCTG
TGTTAATGGCTAACCTGTTACACTGGGCTGGGTTGGGTAGGGTGTTCTGGCTTTTTTGTGGGGTTTTTATTTTTAAAGAAAC
ACTCAATCATCCTAG

FIGURA 37

ATTCGCCCTTTGCTCCCGGGTGTCCCAGGCCCGGCCGGTGCGCAGAGCATGGCGGGTGCGG
GCCCGAAGCGGCGCGCGCTAGCGGCGCCGGCGGCCGAGGAGAAGGAAGAGGCGCGGGAGAA
GATGCTGGCCGCCAAGAGCGCGGACGGCTCGGCGCCGGCAGGCGAGGGCGAGGGCGTGACC
CTGCAGCGGAACATCACGCTGCTCAACGGCGTGGCCATCATCGTGGGGACCATTATCGGCT
CGGGCATCTTCGTGACGCCCACGGGCGTGCTCAAGGAGGCAGGCTCGCCGGGGCTGGCGCT
GGTGGTGTGGGCCGCGTGCGGCGTCTTCTCCATCGTGGGCGCGCTCTGCTACGCGGAGCTC
GGCACCACCATCTCCAAATCGGGCGGCGACTACGCCTACATGCTGGAGGTCTACGGCTCGC
TGCCCGCCTTCCTCAAGCTCTGGATCGAGCTGCTCATCATCCGGCCTTCATCGCAGTACAT
CGTGGCCCTGGTCTTCGCCACCTACCTGCTCAAGCCGCTCTTCCCCACCTGCCCGGTGCCC
GAGGAGGCAGCCAAGCTCGTGGCCTGCCTCTGCGTGCTGCTGCTCACGGCCGTGAACTGCT
ACAGCGTGAAGGCCGCCACCCGGGTCCAGGATGCCTTTGCCGCCGCCAAGCTCCTGGCCCT
GGCCCTGATCATCCTGCTGGGCTTCGTCCAGATCGGGAAGGGTGATGTGTCCAATCTAGAT
CCCAACTTCTCATTTGAAGGCACCAAACTGGATGTGGGGAACATTGTGCTGGCATTATACA
GCGGCCTCTTTGCCTATGGAGGATGGAATTACTTGAATTTCGTCACAGAGGAAATGATCAA
CCCCTACAGAAACCTGCCCCTGGCCATCATCATCTCCCTGCCCATCGTGACGCTGGTGTAC
GTGCTGACCAACCTGGCCTACTTCACCACCCTGTCCACCGAGCAGATGCTGTCGTCCGAGG
CCGTGGCCGTGGACTTCGGGAACTATCACCTGGGCGTCATGTCCTGGATCATCCCCGTCTT
CGTGGGCCTGTCCTGCTTCGGCTCCGTCAATGGGTCCCTGTTCACATCCTCCAGGCTCTTC
TTCGTGGGGTCCCGGGAAGGCCACCTGCCCTCCATCCTCTCCATGATCCACCCACAGCTCC
TCACCCCCGTGCCGTCCCTCGTGTTCACGTGTGTGATGACGCTGCTCTACGCCTTCTCCAA
GGACATCTTCTCCGTCATCAACTTCTTCAGCTTCTTCAACTGGCTCTGCGTGGCCCTGGCC
ATCATCGGCATGATCTGGCTGCGCCACAGAAAGCCTGAGCTTGAGCGGCCCATCAAGGTGA
ACCTGGCCCTGCCTGTGTTCTTCATCCTGGCCTGCCTCTTCCTGATCGCCGTCTCCTTCTG
GAAGACACCCGTGGAGTGTGGCATCGGCTTCACCATCATCCTCAGCGGGCTGCCCGTCTAC
TTCTTCGGGGTCTGGTGGAAAAACAAGCCCAAGTGGCTCCTCCAGGGCATCTTCTCCACGA
CCGTCCTGTGTCAGAAGCTCATGCAGGTGGTCCCCCAGGAGACATAGCCAGGAGGCCGAAG
GGCG

FIGURA 38

CCAGGCTCTCCACCCCCACTTCCCAATTGAGGAAACCGAGGCAGAGGAGGCTCAGGTGTGG
CCAATCACCCTGCACATCAGAGTTACCCTGGGCAGGGCCCACTGAGACCTGGGAGGGGCCA
CTCGGGACCTGGAGGGCTGGGGGCTGCCCGGGCGTTAGGGGTAAAGCTCCCTACCCAACTG
CGCAGAAGGCCTCAGAGGCCTGGGGGCTGGGCTTCCCCTTTCACATCGCCCTTTAGAGGCC
CACGTGTGGGCATTGGCCCGCGATCTGAAAGGGGCTGTCCTGTTCCTCATGGGCGCTGCCA
GCGCCACGCACTCCTCTTTCTGCCTGGCCGGCCACTCCCGTCTGCTGTGACGCGCGGACAG
AGAGCTACCGGTGGACCCACGGTGCCTCCCTCCCTGGGATCTACACAGACCATGGCCTTGC
CAACGGCTCGACCCCTGTTGGGGTCCTGTGGGACCCCCGCCCTCGGCAGCCTCCTGTTCCT
GCTCTTCAGCGTCGGATGGGTGCAGCCCTCGAGGACCCTGGCTGGAGAGACAGGGCAGGAG
GCTGCACCCCTGGACGGAGTCCTGGCCAACCCACCTAACATTTCCAGCCTCTCCCCTCGCC
AACTCCTTGGCTTCCCGTGTGCGGAGGTGTCCGGCCTGAGCACGGAGCGTGTCCGGGAGCT
GGCTGTGGCCTTGGCACAGAAGAATGTCAAGCTCTCAACAGAGCAGCTGCGCTGTCTGGCT
CACCGGCTCTCTGAGCCCCCCGAGGACCTGGACGCCCTCCCATTGGACCTGCTGCTATTCC
TCAACCCAGATGCGTTCTCGGGGCCCCAGGCCTGCACCCGTTTCTTCTCCCGCATCACGAA
GGCCAATGTGGACCTGCTCCCGAGGGGGGCTCCCGAGCGACAGCGGCTGCTGCCTGCGGCT
CTGGCCTGCTGGGGTGTGCGGGGGTCTCTGCTGAGCGAGGCTGATGTGCGGGCTCTGGGAG
GCCTGGCTTGCGACCTGCCTGGGCGCTTTGTGGCCGAGTCGGCCGAAGTGCTGCTACCCCG
GCTGGTGAGCTGCCCGGGACCCCTGGACCAGGACCAGCAGGAGGCAGCCAGGGCGGCTCTG
CAGGGCGGGGGACCCCCCTACGGCCCCCCGTCGACATGGTCTGTCTCCACGATGGACGCTC
TGCGGGGCCTGCTGCCCGTGCTGGGCCAGCCCATCATCCGCAGCATCCCGCAGGGCATCGT
GGCCGCGTGGCGGCAACGCTCCTCTCGGGACCCATCCTGGCGGCAGCCTGAACGGACCATC
CTCCGGCCGCGGTTCCGGCGGGAAGTGGAGAAGACAGCCTGTCCTTCAGGCAAGAAGGCCC
GCGAGATAGACGAGAGCCTCATCTTCTACAAGAAGTGGGAGCTGGAAGCCTGCGTGGATGC
GGCCCTGCTGGCCACCCAGATGGACCGCGTGAACGCCATCCCCTTCACCTACGAGCAGCTG
GACGTCCTAAAGCATAAACTGGATGAGCTCTACCCACAAGGTTACCCCGAGTCTGTGATCC
AGCACCTGGGCTACCTCTTCCTCAAGATGAGCCCTGAGGACATTCGCAAGTGGAATGTGAC
GTCCCTGGAGACCCTGAAGGCTTTGCTTGAAGTCAACAAAGGGCACGAAATGAGTCCTCAG
GTGGCCACCCTGATCGACCGCTTTGTGAAGGGAAGGGGCCAGCTAGACAAAGACACCCTAG
ACACCCTGACCGCCTTCTACCCTGGGTACCTGTGCTCCCTCAGCCCCGAGGAGCTGAGCTC
CGTGCCCCCCAGCAGCATCTGGGCGGTCAGGCCCCAGGACCTGGACACGTGTGACCCAAGG
CAGCTGGACGTCCTCTATCCCAAGGCCCGCCTTGCTTTCCAGAACATGAACGGGTCCGAAT
ACTTCGTGAAGATCCAGTCCTTCCTGGGTGGGGCCCCCACGGAGGATTTGAAGGCGCTCAG
TCAGCAGAATGTGAGCATGGACTTGGCCACGTTCATGAAGCTGCGGACGGATGCGGTGCTG
CCGTTGACTGTGGCTGAGGTGCAGAAACTTCTGGGACCCCACGTGGAGGGCCTGAAGGCGG
[illegible]

GCTGGGGCTGGGGCTACAGGGCGGCATCCCCAACGGCTACCTGGTCCTAGACCTCAGCGTG
CAAGAGGCCCTCTCGGGGACGCCCTGCCTCCTAGGACCTGGACCTGTTCTCACCGTCCTGG
CACTGCTCCTAGCCTCCACCCTGGCCTGAGGGCCCCACTCCCTTGCTGGCCCCAGCCCTGC
TGGGGATCCCCGCCTGGCCAGGAGCAGGCACGGGTGATCCCCGTTCCACCCCAAGAGAACT
CGCGCTCAGTAAACGGGAACATGCCCCCTGCAGACACGTAAAAAAAAAAAAAAAAAA

FIGURA 39

TGGGATCTACACAGACCATGGCCTTGCCAACGGCTCGACCCCTGTTGGGGTCCTGTGGGAC
CCCCGCCCTCGGCAGC
CTCCTGTTCCTGCTCTTCAGCCTCGGATGGGTGCAGCCCTCGAGGACCCTGGCTGGAGAGA
CAGGGCAGGAGGCTGCGCCCCTGGACGGAGTCCTGGCCAACCCACCTAACATTTCCAGCCT
CTCCCCTCGCCAACTCCTTGGCTTCCCGTGTGCGGAGGTGTCCGGCCTGAGCACGGAGCGT
GTCCGGGAGCTGGCTGTGGCCTTGGCACAGAAGAATGTCAAGCTCTCAACAGAGCAGCTGC
GCTGTCTGGCTCACCGGCTCTCTGAGCCCCCCGAGGACCTGGACGCCCTCCCATTGGACCT
GCTGCTATTCCTCAACCCAGATGCGTTCTCGGGGCCCCAGGCCTGCACCCGTTTCTTCTCC
CGCATCACGAAGGCCAATGTGGACCTGCTCCCGAGGGGGGCTCCCGAGCGACAGCGGCTGC
TGCCTGCGGCTCTGGCCTGCTGGGGTGTGCGGGGGTCTCTGCTGAGCGAGGCTGATGTGCG
GGCTCTGGGAGGCCTGGCTTGCGACCTGCCTGGGCGCTTTGTGGCCGAGTCGGCCGAAGTG
CTGCTACCCCGGCTGGTGAGCTGCCCGGGACCCCTGGACCAGGACCAGCAGGAGGCAGCCA
GGGCGGCTCTGCAGGGCGGGGGACCCCCCTACGGCCCCCCGTCGACATGGTCTGTCTCCAC
GATGGACGCTCTGCGGGGCCTGCTGCCCGTGCTGGGCCAGCCCATCATCCGCAGCATCCCG
CAGGGCATCGTGGCCGCGTGGCGGCAACGCTCCTCTCGGGACCCATCCTGGCGGCAGCCTG
AACGGACCATCCTCCGGCCGCGGTTCCGGCGGGAAGTGGAGAAGACAGCCTGTCCTTCAGG
CAAGAAGGCCCGCGAGATAGACGAGAGCCTCATCTTCTACAAGAAGTGGGAGCTGGAAGCC
TGCGTGGATGCGGCCCTGCTGGCCACCCAGATGGACCGCGTGAACGCCATCCCCTTCACCT
ACGAGCAGCTGGACGTCCTAAAGCATAAACTGGATGAGCTCTACCCACAAGGTTACCCCGA
GTCTGTGATCCAGCACCTGGGCTACCTCTTCCTCAAGATGAGCCCTGAGGACATTCGCAAG
TGGAATGTGACGTCCCTGGAGACCCTGAAGGCTTTGCTTGAAGTCAACAAAGGGCACGAAA
TGAGTCCTCAGGTGGCCACCCTGATCGACCGCTTTGTGAAGGGAAGGGGCCAGCTAGACAA
AGACACCCTAGACACCCTGACCGCCTTCTACCCTGGGTACCTGTGCTCCCTCAGCCCCGAG
GAGCTGAGCTCCGTGCCCCCCAGCAGCATCTGGGCGGTCAGGCCCCAGGACCTGGACACGT
GTGACCCAAGGCAGCTGGACGTCCTCTATCCCAAGGCCCGCCTTGCTTTCCAGAACATGAA
CGGGTCCGAATACTTCGTGAAGATCCAGTCCTTCCTGGGTGGGGCCCCCACGGAGGATTTG
AAGGCGCTCAGTCAGCAGAATGTGAGCATGGACTTGGCCACGTTCATGAAGCTGCGGACGG
ATGCGGTGCTGCCGTTGACTGTGGCTGAGGTGCAGAAACTTCTGGGACCCCACGTGGAGGG
CCTGAAGGCGGAGGAGCGGCACCGCCCGGTGCGGGACTGGATCCTACGGCAGCGGCAGGAC
GACCTGGACACGCTGGGGCTGGGGCTACAGGGCGGCATCCCCAACGGCTACCTGGTCCTAG
ACCTCAGCGTGCAAGAGGCCCTCTCGGGGACGCCCTGCCTCCTAGGACCTGGACCTGTTCT
CACCGTCCTGGCACTGCTCCTAGCCTCCACCCTGGCCTGAGGGCCCCACTCCCTTGCTGGC
CCCAGCCCTGCTGGGGATCCCCGCCTGGCCAGGAGCAGGCACGGGTGATCCCCGTTCCACC
CCAAGAGAACT

FIGURA 40

ATAACAGCATGAAGTGCCGTGGAACTGGAATAGGCGTGTCCTCTCCCTCGACCCTCCCCCT
CCTTGTCCCTCTGCTCACCCCTCGCTCGTTCCCTCCCTCCGGCGAGGGCCGCCTTTATAAC
AACTGCTCAGAGTGCGAGGGCGGGATAGCTGTCCAAGGTCTCCCCCAGCACTGAGGAGCTC
GCCTGCTGCCCTCTTGCGCGCGGGAAGCAGCACCAAGTTCACGGCCAACGCCTTGGCACTA
GGGTCCAGAATGGCTACAACAGTCCCTGATGGTTGCCGCAATGGCCTGAAATCCAAGTACT
ACAGACTTTGTGATAAGGCTGAAGCTTGGGGCATCGTCCTAGAAACGGTGGCCACAGCCGG
GGTTGTGACCTCGGTGGCCTTCATGCTCACTCTCCCGATCCTCGTCTGCAAGGTGCAGGAC
TCCAACAGGCGAAAAATGCTGCCTACTCAGTTTCTCTTCCTCCTGGGTGTGTTGGGCATCT
TTGGCCTCACCTTCGCCTTCATCATCGGACTGGACGGGAGCACAGGGCCCACACGCTTCTT
CCTCTTTGGGATCCTCTTTTCCATCTGCTTCTCCTGCCTGCTGGCTCATGCTGTCAGTCTG
ACCAAGCTCGTCCGGGGGAGGAAGCCCCTTTCCCTGTTGGTGATTCTGGGTCTGGCCGTGG
GCTTCAGCCTAGTCCAGGATGTTATCGCTATTGAATATATTGTCCTGACCATGAATAGGAC
CAACGTCAATGTCTTTTCTGAGCTTTCCGCTCCTCGTCGCAATGAAGACTTTGTCCTCCTG
CTCACCTACGTCCTCTTCTTGATGGCGCTGACCTTCCTCATGTCCTCCTTCACCTTCTGTG
GTTCCTTCACGGGCTGGAAGAGACATGGGGCCCACATCTACCTCACGATGCTCCTCTCCAT
TGCCATCTGGGTGGCCTGGATCACCCTGCTCATGCTTCCTGACTTTGACCGCAGGTGGGAT
GACACCATCCTCAGCTCCGCCTTGGCTGCCAATGGCTGGGTGTTCCTGTTGGCTTATGTTA
GTCCCGAGTTTTGGCTGCTCACAAAGCAACGAAACCCCATGGATTATCCTGTTGAGGATGC
TTTCTGTAAACCTCAACTCGTGAAGAAGAGCTATGGTGTGGAGAACAGAGCCTACTCTCAA
GAGGAAATCACTCAAGGTTTTGAAGAGACAGGGGACACGCTCTATGCCCCCTATTCCACAC
ATTTTCAGCTGCAGAACCAGCCTCCCCAAAAGGAATTCTCCATCCCACGGGCCCACGCTTG
GCCGAGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGGCAGCTAACTCTGTCCTGAAGAG
TGGGACAAATGCAGCCGGGCGGCAGATCTAGCGGGAGCTCAAAGGGATGTGGGCGAAATCT
TGAGTCTTCTGAGAAAACTGTACAAGACACTACGGGAACAGTTTGCCTCCCTCCCAGCCTC
AACCACAATTCTTCCATGCTGGGGCTGATGTGGGCTAGTAAGACTCCAGTTCTTAGAGGCG
CTGTAGTATTTTTTTTTTTTTGTCTCATCCTTTGGATACTTCTTTTAAGTGGGAGTCTCAG
GCAACTCAAGTTTAGACCCTTACTCTTTTTGTTTGTTTTTTGAAACAGGATCTTGCTCTGT
CACCCAGGCTTGAGTGCAGTGGTGCGATCACAGCCCAGTGCAGCCTCGACCACCTGTGCTC
AAGCAATCCTCCCATCTCCATCTCCCAAAGTGCTGGGATGACAGGCGTGAGCCACAGCTCC
CAGCCTAGGCCCTTAATCTTGCTGTTATTTTCCATGGACTAAAGGTCTGGTCATCTGAGCT
CACGCTGGCTCACACAGCTCTAGGGGCCTGCTCCTCTAACTCACAGTGGGTTTTGTGAGGC
TCTGTGGCCCAGAGCAGACCTGCATATCTGAGCAAAAATAGCAAAAGCCTCTCTCAGCCCA
CTGGCCTGAATCTACACTGGAAGCCAACTTGCTGGCACCCCCGCTCCCCAACCCTTCTTGC
CTGGGTAGGAGAGGCTAAAGATCACCCTAAATTTACTCATCTCTCTAGTGCTGCCTCACAT

FIGURA 41

AGAGACTCAAGATGATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTGTTAA
CCCTATAAACGCCAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGGTCGG
GACCAAGGCCCAAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATACTTCA
GCACTTGTAAGAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTTTTATATGA
ATGTTGCCCTGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCCCATT
GACCATGTTTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCTGACG
CCTCAAAACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGAGTAA
TGAGGCTTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAATGTT
GAATTACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGACTTAA
AAAATGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATTATCC
TAATGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAACAAAT
GGTGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGACTTCA
TTGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATATTGGA
GGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGAGAAA
CTTCCACGAGGTGTCCTAGAAAGGTTCATGGGAGACAAAGTGGCTTCCGAAGCTCTTATGA
AGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCTTTGA
GACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAATGGA
ATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGATTGATC
AGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCAAACCAC
CTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGGAGAATAC
ACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGTTCAGCGCC
TCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAATGAGCTTTA
CAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGTATATCGTACA
GCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAGAAACGGTGCGA
TTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGAAAAGTTAAAACA
AGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTTGAAAGAGCTCCTG
ACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTTTAAGGGAATGACTA
GTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAACATCATTCTTTATCA
CCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGTTACTAACATTTTAAAG
ACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATACACTTCTGGTGAATGAAT
TGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAATTCATGTTGTAGATAAACT
CCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAAATTA
ATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTGACTG
TCTATACAACTAAAATTATAACCAAAGTTGTGGAACCAAAAATTAAAGTGATTGAAGGCAG
TCTTTCAGCCTTATTATCAAAACTGAAGGACCCACACTAACAAAAGTCAAAATTGAAGGTGAA

CCTGAATTCAGACTGATTAAAGAAGGTGAAACAATAACTGAAGTGATCCATGGAGAGCCAA
TTATTAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGAC
ACGAGAAGAACGAATCATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGT
GGAGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGAGGTCACCAAGGTCACCAAAT
TCATTGAAGGTGGTGATGGTCATTTATTTGAAGATGAAGAAATTAAAAGACTGCTTCAGGG
AGACACACCCGTGAGGAAGTTGCAAGCCAACAAAAAAGTTCAAGGTTCTAGAAGACGATTA
AGGGAAGGTCGTTCTCAGTGAAAATCCAAAAACCAGAAAAAAATGTTTATACAACCCTAAG
TCAATAACCTGACCTTAGAAAATTGTGAGAGCCAAGTTGACTTCAGGAACTGAAACATCAG
CACAAAGAAGCAATCATCAAATAATTCTGAACACAAATTTAATATTTTTTTTTCTGAATGA
GAAACATGAGGGAAATTGTGGAGTTAGCCTCCTGTGGTAAAGGAATTGAAGAAAATATAAC
ACCTTACACCCTTTTTCATCTTGACATTAAAAGTTCTGGCTAACTTTGGAATCCATTAGAG
AAAAATCCTTGTCACCAGATTCATTACAATTCAAATCGAAGAGTTGTGAACTGTTATCCCA
TTGAAAAGACCGAGCCTTGTATGTATGTTATGGATACATAAAATGCACGCAAGCCATTATC
TCTCCATGGGAAGCTAAGTTATAAAAATAGGTGCTTGGTGTACAAAACTTTTTATATCAAA
AGGCTTTGCACATTTCTATATGAGTGGGTTTACTGGTAAATTATGTTATTTTTTACAACTA
ATTTTGTACTCTCAGAATGTTTGTCATATGCTTCTTGCAATGCATATTTTTTAATCTCAAA
CGTTTCAATAAAACCATTTTTCAGATATAAAGAGAATTACTTCAAATTGAGTAATTCAGAA
AAACTCAAGATTTAAGTTAAAAAGTGGTTTGGACTTGGGAA

FIGURA 42

CCCTTGAGACTCAAGATGATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTG
TTAACCCTATAAACGCCAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGG
TCGGGACCAAGGCCCAAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATAC
TTCAGCACTTGTAAGAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTGTTAT
ATGAATGTTGCCCTGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCC
CATTGACCATGTTTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCT
GACGCCTCAAAACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGA
GTAATGAGGCTTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAA
TGTTGAATTACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGAC
TTAAAAAATGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATT
ATCCTAATGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAAC
AAATGGTGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGAC
TTCATTGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATAT
TGGAGGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGA
GAAACTTCCACGAGGTGTCCTAGAAAGGATCATGGGAGACAAAGTGGCTTCCGAAGCTCTT
ATGAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCT
TTGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
TGGAATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGATT
GATCAGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCAAA
CCACCTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGGAGA
ATACACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGATCAG
CGCCTCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAATGAGC
TTTACAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGTATATCG
TACAGCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAGAAACGGT
GCGATTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGAAAAGTTAA
AACAAGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTTGAAAGAGCT
CCTGACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTTTAAGGGAATG
ACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAACATCATTCTTT
ATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGTTACTAACATTTT
AAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATACACTTCTGGTGAAT
GAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAATTCATGTTGTAGATA
AACTCCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAA
ATTAATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTG
ACTGTCTATAGACCCACACTAACAAAAGTCAAAATTGAAGGTGAACCTGAATTCAGACTGA
TTAAAGAAGCTTAAAGAATAACTGAAGTGATCCATGGAGAGCCAATTATTAAAAAATACAC

CAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGAATC
ATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAA
CTCTGAAGAAATTGTTACAAGAAGAGGTCACCAAGGTCACCAAATTCATTGAAGGTGGTGA
TGGTCATTTATTTGAAGATGAAGAAATTAAAAGACTGCTTCAGGGAGACACACCCGTGAGG
AAGTTGCAAGCCAACAAAAAAGTTCAAGGATCTAGAAGACGATTAAGGGAAGGTCGTTCTC
AGTGAAGGGC

FIGURA 43

CCCTTGAGACTCAAGATGATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTG
TTAACCCTATAAACGCCAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGG
TCGGGACCAAGGCCCAAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATAC
TTCAGCACTTGTAAGAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTGTTAT
ATGAATGTTGCCCTGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCC
CATTGACCATGTTTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCT
GACGCCTCAAAACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGA
GTAATGAGGCTTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAA
TGTTGAATTACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGAC
TTAAAAAATGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATT
ATCCTAATGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAAC
AAATGGTGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGAC
TTCATTGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATAT
TGGAGGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGA
GAAACTTCCACGAGGTGTCCTAGAAAGGATCATGGGAGACAAAGTGGCTTCCGAAGCTCTT
ATGAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCT
TTGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
TGGAATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGATT
GATCAGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCAAA
CCACCTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGGAGA
ATACACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGATCAG
CGCCTCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAATGAGC
TTTACAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGTATATCG
TACAGCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAGAAACGGT
GCGATTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGAAAAGTTAA
AACAAGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTTGAAAGAGCT
CCTGACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTTTAAGGGAATG
ACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAACATCATTCTTT
ATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGTTACTAACATTTT
AAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATACACTTCTGGTGAAT
GAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAATTCATGTTGTAGATA
AACTCCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAA
ATTAATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTG
ACTGTCTATAAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAA
TTACTGTAAAAAGAGGCACGTAGAAGAACGAATCATTACAGGTCCTGAAATAAAATACACTAG

GATTTCTACTGGAGGTGGAGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGACACA
CCCGTGAGGAAGTTGCAAGCCAACAAAAAAGTTCAAGGATCTAGAAGACGATTAAGGGAAG
GTCGTTCTCAGTGAAGGGC

FIGURA 44

CCCTTGAGACTCAAGATGATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTG
TTAACCCTATAAACGCCAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGG
TCGGGACCAAGGCCCAAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATAC
TTCAGCACTTGTAAGAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTGTTAT
ATGAATGTTGCCCTGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCC
CATTGACCATGTTTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCT
GACGCCTCAAAACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGA
GTAATGAGGCTTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAA
TGTTGAATTACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGAC
TTAAAAAATGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATT
ATCCTAATGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAAC
AAATGGTGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGAC
TTCATTGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATAT
TGGAGGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGA
GAAACTTCCACGAGGTGTCCTAGAAAGGATCATGGGAGACAAAGTGGCTTCCGAAGCTCTT
ATGAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCT
TTGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
TGGAATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGATT
GATCAGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCAAA
CCACCTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGGAGA
ATACACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGATCAG
CGCCTCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAATGAGC
TTTACAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGTATATCG
TACAGCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAGAAACGGT
GCGATTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGAAAAGTTAA
AACAAGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTTGAAAGAGCT
CCTGACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTTTAAGGGAATG
ACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAACATCATTCTTT
ATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGTTACTAACATTTT
AAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATACACTTCTGGTGAAT
GAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAATTCATGTTGTAGATA
AACTCCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAA
ATTAATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTG
ACTGTCTATAGACCCACACTAACAAAAGTCAAAATTGAAGGTGAACCTGAATTCAGACTGA
TTTAAGCAAGCTTCAAACAATTACTTGAAGTGATCCATGGAGAGCCAATTATTAAAAAATACAC

CAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGAATC
ATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAA
CTCTGAAGAAATTGTTACAAGAAGACACACCCGTGAGGAAGTTGCAAGCCAACAAAAAAGT
TCAAGGATCTAGAAGACGATTAAGGGAAGGTCGTTCTCAGTGAAGGGC

FIGURA 45

CCCTTGAGACTCAAGATGATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTG
TTAACCCTATAAACGCCAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGG
TCGGGACCAAGGCCCAAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATAC
TTCAGCACTTGTAAGAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTGTTAT
ATGAATGTTGCCCTGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCC
CATTGACCATGTTTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCT
GACGCCTCAAAACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGA
GTAATGAGGCTTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAA
TGTTGAATTACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGAC
TTAAAAAATGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATT
ATCCTAATGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAAC
AAATGGTGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGAC
TTCATTGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATAT
TGGAGGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGA
GAAACTTCCACGAGGTGTCCTAGAAAGGATCATGGGAGACAAAGTGGCTTCCGAAGCTCTT
ATGAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCT
TTGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
TGGAATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGATT
GATCAGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCAAA
CCACCTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGGAGA
ATACACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGATCAG
CGCCTCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAATGAGC
TTTACAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGTATATCG
TACAGCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAGAAACGGT
GCGATTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGAAAAGTTAA
AACAAGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTTGAAAGAGCT
CCTGACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTTTAAGGGAATG
ACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAACATCATTCTTT
ATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGTTACTAACATTTT
AAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATACACTTCTGGTGAAT
GAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAATTCATGTTGTAGATA
AACTCCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAA
ATTAATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTG
ACTGTCTATAAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAA
TAACTTGAAAAAGAGACACGAGAAGAACGAATCATTACAGGTCCTGAAATAAAATACACTAG

GATTTCTACTGGAGGTGGAGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGAGGTC
ACCAAGGTCACCAAATTCATTGAAGGTGGTGATGGTCATTTATTTGAAGATGAAGAAATTA
AAAGACTGCTTCAGGGAGACACACCCGTGAGGAAGTTGCAAGCCAACAAAAAAGTTCAAGG
ATCTAGAAGACGATTAAGGGAAGGTCGTTCTCAGTGAAGGGC

FIGURA 46

ACACGGACCAAGGAGTCTAACACGTGCGCGAGTCGGGGGCTCGCACGAAAGCCGCCGTGGC
GCAATGAAGGTGAAGGCCGGCGCGCTCGCCGGCCGAGGTGGGATCCCGAGGCCTCTCCAGT
CCGCCGAGGGCGCACCACCGGCCCGTCTCGCCCGCCGCGCCGGGGAGGTGGAGCACGAGCG
CACGTGTTAGGACCCGAAAGATGGTGAACTATGCCTGGGCAGGGCGAAGCCAGAGGAAACT
CTGGTGGAGGTCCGTAGCGGTCCTGACGTGCAAATCGGTCGTCCGACCTGGGTATAGGGGC
GGGCTCCAGGCGAGGCGGTCGACGCTCCTGAAAACTTGCGCGCGCGCTCGCGCCACTGCGC
CCGGAGCGATGAAGATGGTCGCGCCCTGGACGCGGTTCTACTCCAACAGCTGCTGCTTGTG
CTGCCATGTCCGCACCGGCACCATCCTGCTCGGCGTCTGGTATCTGATCATCAATGCTGTG
GTACTGTTGATTTTATTGAGTGCCCTGGCTGATCCGGATCAGTATAACTTTTCAAGTTCTG
AACTGGGAGGTGACTTTGAGTTCATGGATGATGCCAACATGTGCATTGCCATTGCGATTTC
TCTTCTCATGATCCTGATATGTGCTATGGCTACTTACGGAGCGTACAAGCAACGCGCAGCC
TGGATCATCCCATTCTTCTGTTACCAGATCTTTGACTTTGCCCTGAACATGTTGGTTGCAA
TCACTGTGCTTATTTATCCAAACTCCATTCAGGAATACATACGGCAACTGCCTCCTAATTT
TCCCTACAGAGATGATGTCATGTCAGTGAATCCTACCTGTTTGGTCCTTATTATTCTTCTG
TTTATTAGCATTATCTTGACTTTTAAGGGTTACTTGATTAGCTGTGTTTGGAACTGCTACC
GATACATCAATGGTAGGAACTCCTCTGATGTCCTGGTTTATGTTACCAGCAATGACACTAC
GGTGCTGCTACCCCCGTATGATGATGCCACTGTGAATGGTGCTGCCAAGGAGCCACCGCCA
CCTTACGTGTCTGCCTAAGCCTTCAAGTGGGCGGAGCTGAGGGCAGCAGCTTGACTTTGCA
GACATCTGAGCAATAGTTCTGTTATTTCACTTTTGCCATGAGCCTCTCTGAGCTTGTTTGT
TGCTGAAATGCTACTTTTTAAAATTTAGATGTTAGATTGAAAACTGTAGTTTTCAACATAT
GCTTTGCTAGAACACTGTGATAGATTAACTGTAGAATTCTTCCTGTACGATTGGGGATATA
ATGGGCTTCACTAACCTTCCCTAGGCATTGAAACTTCCCCCAAATCTGATGGACCTAGAAG
TCTGCTTTTGTACCTGCTGGGCCCCAAAGTTGGGCATTTTTCTCTCTGTTCCCTCTCTTTT
GAAAATGTAAAATAAAACCAAAAATAGACAACTTTTTCTTCAGCCATTCCAGCATAGAGAA
CAAAACCTTATGGAAACAGGAATGTCAATTGTGTAATCATTGTTCTAATTAGGTAAATAGA
AGTCCTTATGTATGTGTTACAAGAATTTCCCCCACAACATCCTTTATGACTGAAGTTCAAT
GACAGTTTGTGTTTGGGTGGTAAAGGATTTTCTCCATGGCCTGAATTAAGACCATTAGAAA
GCACCAGGCCGTGGGAGCAGTGACCATCTGCTGACTGTTCTTGTGGATCTTGTGTCCAGGG
ACATGGGGTGACATGCCTCGTATGTGTTAGAGGGTGGAATGGATGTGTTTGGCGCTGCATG
GGATCTGGTGCCCCTCTTCTCCTGGATTCACATCCCCACCCAGGGCCCGCTTTTACTAAGT
GTTCTGCCCTAGATTGGTTCAAGGAGGTCATCCAACTGACTTTATCAAGTGGAATTGGGAT
ATATTTGATATACTTCTGCCTAACAACATGGAAAAGGGTTTTCTTTTCCCTGCAAGCTACA
TCCTACTGCTTTGAACTTCCAAGTATGTCTAGTCACCTTTTAAAATGTAAACATTTTCAGA
AAAATGAGGATTGCCTTCCTTGTATGCGCTTTTTACCTTGACTACCTGAATTGCAAGGGAT

TGTAAATTAAGTTGTTTGCAATTAAAACAAGGTTTGCCCACAAAAAAAAAA

FIGURA 47

ATTCGCCCTTCCGAAAGATGGTGAACTATGCCTGGGCAGGGCGGGCTCCAGGCGAGGCGGT
CGACGCTCCTGAAAACTTGCGCGCGCGCTCGCGCCACTGCGCCCGGAGCGATGAAGATGGT
CGCGCCCTGGACTCGGTTCTACTCCAACAGCTGCTGCTTGTGCTGCCATGTCCGCACCGGC
ACCATCCTGCTCGGCGTCTGGTATCTGATCATCAATGCTGTGGTACTGTTGATTTTATTGA
GTGCCCTGGCTGATCCGGATCAGTATAACTTTTCAAGTTCTGAACTGGGAGGTGACTTTGA
GTTCATGGATGATGCCAACATGTGCATTGCCATTGCGATTTCTCTTCTCATGATCCTGATA
TGTGCTATGGCTACTTACGGAGCGTACAAGCAACGCGCAGCCTGGATCATCCCATTCTTCT
GTTACCAGATCTTTGACTTTGCCCTGAACATGTTGGTTGCAATCACTGTGCTTATTTATCC
AAACTCCATTCAGGAATACATACGGCAACTGCCTCCTAATTTTCCCTACAGAGATGATGTC
ATGTCAGTGAATCCTACCTGTTTGGTCCTTATTATTCTTCTGTTTATTAGCATTATCTTGA
CTTTTAAGGGTTACTTGATTAGCTGTGTTTGGAACTGCTACCGATACATCAATGGTAGGAA
CTCCTCTGATGTCCTGGTTTATGTTACCAGCAATGACACTACGGTGCTGCTACCCCCGTAT
GATGATGCCACTGTGAATGGTGCTGCCAAGGAGCCACCGCCACCTTACGTGTCTGCCTAAG
CCTTCAAGTGGGCGGAGCTGAGGGCAGCAGCTTGACTTTGCAGACATCTGAGAAGGGC

FIGURA 48

AGCGCCATGGGGCGAGCCGGGGGCGGGGGCCCGGGCCGGGGGCCGCCGCCACTGCTGCTGT
TTCTGGGGGCCGCGCTGGTCCTGGCCTCTGGGGCCGTGCCGGCGCGTGAGGCGGGCAGCGC
GGTTGAGGCCGAAGAGCTGGTGAAGGGCAGCCCGGCGTGGGAGCCGCCTGCCAACGACACG
CGGGAAGAAGCCGGCCCACCAGCGGCTGGGGAAGATGAGGCGTCGTGGACGGCGCCCGGTG
GCGAGCTGGCCGGGCCAGAAGAGGTGCTGCAGGAGTCGGCTGCGGTGACCGGCACCGCCTG
GCTGGAAGCTGACAGCCCAGGCCTGGGAGGAGTGACCGCAGAGGCGGGCAGCGGCGATGCC
CAGGCCCTTCCAGCTACGCTCCAGGCTCCCCACGAGGTCCTCGGGCAGTCAATCATGCCCC
CTGCCATTCCTGAGGCTACAGAGGCCAGCGGGCCACCCTCCCCCACCCCCGGCGACAAGCT
GAGCCCAGCTTCTGAACTCCCCAAGGAGAGCCCCTTGGAGGTTTGGCTGAACCTGGGGGGC
AGCACACCCGACCCTCAAGTGCCAGAGCTGACTTACCCATTTCAGGGCACCCTGGAGCCCC
AACCGGCATCAGATATCATTGACATCGACTACTTCGAAGGACTGGATGGTGAGGGTCGTGG
CGCAGATCTGGGGAGCTTCCCAGGGTCACCAGGAACCTCAGAGAACCACCCTGATACTGAG
GGAGAGACCCCTTCCTGGAGCCTGCTTGACTTATACGATGATTTCACCCCTTCGATGAAT
CTGATTTCTACCCCACCACATCCTTTTATGATGACTTGGATGAAGAGGAGGAGGAAGAGGA
GGATGACAAAGATGCAGTAGGAGGTGGAGACCTAGAAGATGAAAATGAGCTTCTAGTGCCC
ACTGGGAAGCCTGGTCTGGGGCCCGGGACAGGCCAGCCCACCAGTCGGTGGCATGCTGTCC
CTCCACAGCACACTCTGGGGTCGGTCCCCGGCAGCAGCATCGCCCTCAGGCCCCGCCCAGG
AGAGCCAGGCAGGGACTTGGCCTCCAGTGAAAATGGCACTGAGTGCCGCAGTGGCTTTGTG
CGGCATAACGGCTCCTGCCGGTCAGTGTGCGACCTCTTCCCAAGTTACTGTCACAATGGCG
GCCAGTGCTACCTGGTGGAGAACATAGGGGCCTTCTGCAGGTGCAACACGCAGGACTACAT
CTGGCACAAGGGGATGCGCTGCGAGTCCATCATCACCGACTTCCAGGTGATGTGCGTGGCC
GTGGGCTCGGCTGCCCTCGTCCTGCTCCTGCTCTTCATGATGACGGTGTTCTTTGCCAAGA
AGCTCTACCTGCTCAAGACGGAGAATACCAAGCTGCGTAGGACCAACAAATTCCGGACCCC
ATCTGAGCTCCACAATGATAACTTCTCCCTCTCCACCATTGCCGAGGGCTCTCACCCAAAT
GATGATCCTAGTGCTCCCCACAAAATCCAGGAGGTTCTCAAGTCCTGCCTGAAAGAGGAGG
AGTCATTTAACATCCAGAACTCCATGTCGCCCAAACTTGAGGGTGGCAAAGGTGACCAGGC
TGACTTGGATGTGAACTGTCTTCAGAATAATTTAACCTAAAGCAGAGCAAGAAGAGAGGAA
GCGGGGTAGTGGGTGGGGGTAGGGAAGAAACATTATCTCCTCTTGTACAGAGTCTATTTCT
TGTAACCATTTGTTAAACTC

FIGURA 49

GGACAGGCGTGGCGGCCGGAGCCCCAGCATCCCTGCTTGAGGTCCAGGAGCGGAGCCCGCG
GCCACCGCCGCCTGATCAGCGCGACCCCGGCCCGCGCCCGCCCCGCCCGGCAAGATGCTGC
CCGTGTACCAGGAGGTGAAGCCCAACCCGCTGCAGGACGCGAACATCTGCTCACGCGTGTT
CTTCTGGTGGCTCAATCCCTTGTTTAAAATTGGCCATAAACGGAGATTAGAGGAAGATGAT
ATGTATTCAGTGCTGCCAGAAGACCGCTCACAGCACCTTGGAGAGGAGTTGCAAGGGTTCT
GGGATAAAGAAGTTTTAAGAGCTGAGAATGACGCACAGAAGCCTTCTTTAACAAGAGCAAT
CATAAAGTGTTACTGGAAATCTTATTTAGTTTTGGGAATTTTTACGTTAATTGAGGAAAGT
GCCAAAGTAATCCAGCCCATATTTTTGGGAAAAATTATTAATTATTTTGAAAATTATGATC
CCATGGATTCTGTGGCTTTGAACACAGCGTACGCCTATGCCACGGTGCTGACTTTTTGCAC
GCTCATTTTGGCTATACTGCATCACTTATATTTTTATCACGTTCAGTGTGCTGGGATGAGG
TTACGAGTAGCCATGTGCCATATGATTTATCGGAAGGCACTTCGTCTTAGTAACATGGCCA
TGGGGAAGACAACCACAGGCCAGATAGTCAATCTGCTGTCCAATGATGTGAACAAGTTTGA
TCAGGTGACAGTGTTCTTACACTTCCTGTGGGCAGGACCACTGCAGGCGATCGCAGTGACT
GCCCTACTCTGGATGGAGATAGGAATATCGTGCCTTGCTGGGATGGCAGTTCTAATCATTC
TCCTGCCCTTGCAAAGCTGTTTTGGGAAGTTGTTCTCATCACTGAGGAGTAAAACTGCAAC
TTTCACGGATGCCAGGATCAGGACCATGAATGAAGTTATAACTGGTATAAGGATAATAAAA
ATGTACGCCTGGGAAAAGTCATTTTCAAATCTTATTACCAATTTGAGAAAGAAGGAGATTT
CCAAGATTCTGAGAAGTTCCTGCCTCAGGGGGATGAATTTGGCTTCGTTTTTCAGTGCAAG
CAAAATCATCGTGTTTGTGACCTTCACCACCTACGTGCTCCTCGGCAGTGTGATCACAGCC
AGCCGCGTGTTCGTGGCAGTGACGCTGTATGGGGCTGTGCGGCTGACGGTTACCCTCTTCT
TCCCCTCAGCCATTGAGAGGGTGTCAGAGGCAATCGTCAGCATCCGAAGAATCCAGACCTT
TTTGCTACTTGATGAGATATCACAGCGCAACCGTCAGCTGCCGTCAGATGGTAAAAAGATG
GTGCATGTGCAGGATTTTACTGCTTTTTGGGATAAGGCATCAGAGACCCCAACTCTACAAG
GCCTTTCCTTTACTGTCAGACCTGGCGAATTGTTAGCTGTGGTCGGCCCCGTGGGAGCAGG
GAAGTCATCACTGTTAAGTGCCGTGCTCGGGGAATTGGCCCCAAGTCACGGGCTGGTCAGC
GTGCATGGAAGAATTGCCTATGTGTCTCAGCAGCCCTGGGTGTTCTCGGGAACTCTGAGGA
GTAATATTTTATTTGGGAAGAAATATGAAAAGGAACGATATGAAAAAGTCATAAAGGCTTG
TGCTCTGAAAAAGGATTTACAGCTGTTGGAGGATGGTGATCTGACTGTGATAGGAGATCGG
GGAACCACGCTGAGTGGAGGGCAGAAAGCACGGGTAAACCTTGCAAGAGCAGTGTATCAAG
ATGCTGACATCTATCTCCTGGACGATCCTCTCAGTGCAGTAGATGCGGAAGTTAGCAGACA
CTTGTTCGAACTGTGTATTTGTCAAATTTTGCATGAGAAGATCACAATTTTAGTGACTCAT
CAGTTGCAGTACCTCAAAGCTGCAAGTCAGATTCTGATATTGAAAGATGGTAAAATGGTGC
AGAAGGGGACTTACACTGAGTTCCTAAAATCTGGTATAGATTTTGGCTCCCTTTTAAAGAA
GGATAATGAGGAAAGTGAACAACCTCCAGTTCCAGGAACTCCCACACTAAGGAATCGTACC

TGGAGAGCCAAGATACAGAGAATGTCCCAGTTACACTATCAGAGGAGAACCGTTCTGAAGG
AAAAGTTGGTTTTCAGGCCTATAAGAATTACTTCAGAGCTGGTGCTCACTGGATTGTCTTC
ATTTTCCTTATTCTCCTAAACACTGCAGCTCAGGTTGCCTATGTGCTTCAAGATTGGTGGC
TTTCATACTGGGCAAACAAACAAAGTATGCTAAATGTCACTGTAAATGGAGGAGGAAATGT
AACCGAGAAGCTAGATCTTAACTGGTACTTAGGAATTTATTCAGGTTTAACTGTAGCTACC
GTTCTTTTTGGCATAGCAAGATCTCTATTGGTATTCTACGTCCTTGTTAACTCTTCACAAA
CTTTGCACAACAAAATGTTTGAGTCAATTCTGAAAGCTCCGGTATTATTCTTTGATAGAAA
TCCAATAGGAAGAATTTTAAATCGTTTCTCCAAAGACATTGGACACTTGGATGATTTGCTG
CCGCTGACGTTTTTTAGATTTCATCCAGACATTGCTACAAGTGGTTGGTGTGGTCTCTGTGG
CTGTGGCCGTGATTCCTTGGATCGCAATACCCTTGGTTCCCCTTGGAATCATTTTCATTTT
TCTTCGGCGATATTTTTTGGAAACGTCAAGAGATGTGAAGCGCCTGGAATCTACAACTCGG
AGTCCAGTGTTTTCCCACTTGTCATCTTCTCTCCAGGGGCTCTGGACCATCCGGGCATACA
AAGCAGAAGAGAGGTGTCAGGAACTGTTTGATGCACACCAGGATTTACATTCAGAGGCTTG
GTTCTTGTTTTTGACAACGTCCCGCTGGTTCGCCGTCCGTCTGGATGCCATCTGTGCCATG
TTTGTCATCATCGTTGCCTTTGGGTCCCTGATTCTGGCAAAAACTCTGGATGCCGGGCAGG
TTGGTTTGGCACTGTCCTATGCCCTCACGCTCATGGGGATGTTTCAGTGGTGTGTTCGACA
AAGTGCTGAAGTTGAGAATATGATGATCTCAGTAGAAAGGGTCATTGAATACACAGACCTT
GAAAAAGAAGCACCTTGGGAATATCAGAAACGCCCACCACCAGCCTGGCCCCATGAAGGAG
TGATAATCTTTGACAATGTGAACTTCATGTACAGTCCAGGTGGGCCTCTGGTACTGAAGCA
TCTGACAGCACTCATTAAATCACAAGAAAAGGTTGGCATTGTGGGAAGAACCGGAGCTGGA
AAAAGTTCCCTCATCTCAGCCCTTTTTAGATTGTCAGAACCCGAAGGTAAAATTTGGATTG
ATAAGATCTTGACAACTGAAATTGGACTTCACGATTTAAGGAAGAAAATGTCAATCATACC
TCAGGAACCTGTTTTTGTTCACTGGAACAATGAGGAAAAACCTGGATCCCTTTAAGGAGCAC
ACGGATGAGGAACTGTGGAATGCCTTACAAGAGGTACAACTTAAAGAAACCATTGAAGATC
TTCCTGGTAAAATGGATACTGAATTAGCAGAATCAGGATCCAATTTTAGTGTTGGACAAAG
ACAACTGGTGTGCCTTGCCAGGGCAATTCTCAGGAAAAATCAGATATTGATTATTGATGAA
GCGACGGCAAATGTGGATCCAAGAACTGATGAGTTAATACAAAAAAAAATCCGGGAGAAAT
TTGCCCACTGCACCGTGCTAACCATTGCACACAGATTGAACACCATTATTGACAGCGACAA
GATAATGGTTTTAGATTCAGGAAGACTGAAAGAATATGATGAGCCGTATGTTTTGCTGCAA
AATAAAGAGAGCCTATTTTACAAGATGGTGCAACAACTGGGCAAGGCAGAAGCCGCTGCCC
TCACTGAAACAGCAAAACAGGTATACTTCAAAAGAAATTATCCACATATTGGTCACACTGA
CCACATGGTTACAAACACTTCCAATGGACAGCCCTCGACCTTAACTATTTTCGAGACAGCA
CTGTGAATCCAACCAAAATGTCAAGTCCGTTCCGAAGGCATTTTCCACTAGTTTTTGGACT
ATGTAAACCACATTGTACTTTTTTTTACTTTGGCAACAAATATTTATACATACAAGATGCT
AGTTCATTTGAATATTTCTCCC

FIGURA 50

CTCCGTCCCAGGGGAGAAGGAGAGGCGTCTGCAGGGGGCAGAGACCGCAGCTACCTGCCGG
GTGCGCCCCCCACCCAGGAGCGCTCGCTTCGCCCCCTTTCCTCCCCCCGCCCCCACCTCCT
TATTGGTGCTAGTTTGCAGCGCCCAGCTCCTGCGCCTTCGCTTCGCGTTTGAATCTGGCTC
GCCCCTTCGTATTATGTCTGCACTCCGAAGGAAATTTGGGGACGATTATCAGGTAGTGACC
ACATCGTCCAGCGGCTCGGGCTTGCAGCCCCAGGGGCCAGGCCAGGACCCTCAGCAGCAGC
TTGTGCCCAAGAAGAAGCGGCAGCGGTTCGTGGACAAGAACGGCCGGTGCAATGTACAGCA
CGGCAACCTGGGCAGCGAGACAAGCCGCTACCTCTCGGACCTCTTCACCACGCTGGTGGAC
CTCAAGTGGCGCTGGAACCTCTTCATCTTCATTCTCACCTACACCGTGGCCTGGCTTTTCA
TGGCGTCCATGTGGTGGGTGATCGCCTACACTCGGGGCGACCTGAACAAAGCCCACGTCGG
TAACTACACGCCTTGCGTGGCCAATGTCTATAACTTCCCTTCTGCCTTCCTCTTCTTCATC
GAGACGGAGGCCACCATCGGCTATGGCTACCGATACATCACAGACAAGTGCCCCGAGGGCA
TCATCCTCTTCCTCTTCCAGTCCATCCTGGGCTCCATCGTGGACGCCTTCCTCATCGGCTG
CATGTTCATCAAGATGTCCCAGCCCAAGAAGCGCGCCGAGACCCTCATGTTCAGCGAGCAC
GCGGTGATCTCCATGAGGGACGGAAAACTCACGCTTATGTTCCGGGTGGGCAACCTGCGCA
ACAGCCACATGGTCTCCGCGCAGATTCGCTGCAAGCTGCTCAAATCTCGGCAGACACCTGA
GGGTGAGTTCCTTCCCCTTGACCAACTTGAACTGGATGTAGGTTTTAGTACAGGGGCAGAT
CAACTTTTTCTTGTGTCCCCCCTCACAATTTGCCACGTGATCGATGCCAAAAGCCCCTTTT
ATGACCTATCCCAGCGAAGCATGCAAACTGAACAGTTCGAGATTGTCGTCATCCTAGAAGG
CATTGTGGAAACAACTGGGATGACTTGTCAAGCTCGAACATCATATACTGAAGATGAAGTT
CTTTGGGGTCATCGTTTTTTTCCTGTAATTTCCTTAGAAGAGGGATTCTTTAAAGTTGATT
ACTCCCAGTTCCACGCAACATTTGAAGTCCCCACCCCACCTTACAGTGTGAAAGAGCAGGA
GGAAATGCTTCTCATGTCGTCCCCTTTAATAGCACCAGCCATAACTAACAGCAAAGAAAGA
CATAATTCTGTGGAATGCTTAGATGGACTAGATGATATTACTACAAAACTACCATCTAAGC
TGCAGAAAATTACTGGAAGAGAAGACTTTCCCAAAAAACTCTTGAGGATGAGTTCTACAAC
TTCAGAAAAAGCCTACAGCTTGGGAGACTTGCCCATGAAACTTCAACGAATAAGTTCAGTT
CCGGGCAACTCAGAAGAAAAACTGGTATCTAAAACCACCAAGATGTTATCTGATCCCATGA
GCCAGTCTGTGGCTGATTTGCCACCAAAGCTTCAAAAGATGGCTGGAGGAGCAGCTAGGAT
GGAAGGGAACCTTCCAGCCAAATTAAGAAAAATGAACTCTGATCGCTTCACATAACAAAGC
ACTCCCTTAGGCATTATTTAATGTTTGATTTAGTAATAGTCCAATATTTGGCGATGAGGTA
ATTCTCCCTAAGGAATCTGAAAGTATATTTTCCTCCCAGTTCTACAGCATATTTCGAGAAC
CCTTCCTTTCCCAAGTATTGCGAATGTGCAGAAAGCAACAGTTACGGAGGGAGGACATCAT
AAGGAAGTTATTAACGGGCATGTATTATCACATCAAGCATGCAATAATGTGCAAATTTTGC
ATTTAGTTTTATGGCATGATTTATATATGGCATATTTATATTGTATATTCTGGAAAAAAAA
ATATATATATATATTTAAAGGGGAGATACTCTCCCTGACATTTCTAACATATGTATTAAGC

TATGTATACACACATATACATATATATATACACATACATACACATACATACATACATACAT
ATATATCTGATAAAATTGTGATGTTTTGTTCAAAGTTGTAGTTCTTGTGCATGTTTACTTT
ATTAGAGTAGGAAGGCTACTGGCATTAATTATTAATACCAAATATTTTAGCCTTAAATTTT
TGTCATTTTAAAATCTGATTTAATGTTTTCTGCTGTTTAAGGTCTTGGGAGGCTTTCAATT
GTATTTTATATGAGAGAATCACACAAGTTTGTGCTATCTATGGCCCTGCAAAAATATAACC
ATTACATGTTTAAATTGTAAATTTTAGAGCATACCAGTACTCAGTATAGCATTGAACATTT
CTTATGATTTTTAAAAGTTGCTAGTACTGGGGAGAAATAATTGTTGATTAATTTGAGAATT
ATTCCTTTCCTAGACTAATTAAAATCTGGAAATCTGTTTTGTATATGATCTAATACAAAGA
TGAGCTCTGAACAAACACTGAATCATGTTAATAGACAGTAGCCAAGTTATATTGAATATAT
CAGAATCTGTGTGAAGTTACACAATTAATTGTCCCTGTTTCAAACTGAGTAAATTGGAAAC
ATTTTCTTTCTTTTTCTGGAAATTTTGTCCATTTTAAAAACCAATCATTTTAAGAAGACAT
GACAATGCAATGAAACAGATGATAAATATTTATGCTTAAAATAAAAAAAAAAAAA

FIGURA 51

GAGAGGGTCCTTCAGGGTCTGCTTATGCCCTTGTTCAAGAACACCAGTGTCAGCTCTCTGT
ACTCTGGTTGCAGACTGACCTTGCTCAGGCCTGAGAAGGATGGGGCAGCCACCAGAGTGGA
TGCTGTCTGCACCCATCGTCCTGACCCCAAAAGCCCTGGACTGGACAGAGAGCGGCTGTAC
TGGAAGCTGAGCCAGCTGACCCACGGCATCACTGAGCTGGGCCCCTACACCCTGGACAGGC
ACAGTCTCTATGTCAATGGTTTCACCCATCAGAGCTCTATGACGACCACCAGAACTCCTGA
TACCTCCACAATGCACCTGGCAACCTCGAGAACTCCAGCCTCCCTGTCTGGACCTACGACC
GCCAGCCCTCTCCTGGTGCTATTCACAATTAACTTCACCATCACTAACCTGCGGTATGAGG
AGAACATGCATCACCCTGGCTCTAGAAAGTTTAACACCACGGAGAGAGTCCTTCAGGGTCT
GCTCAGGCCTGTGTTCAAGAACACCAGTGTTGGCCCTCTGTACTCTGGCTGCAGACTGACC
TTGCTCAGGCCCAAGAAGGATGGGGCAGCCACCAAAGTGGATGCCATCTGCACCTACCGCC
CTGATCCCAAAAGCCCTGGACTGGACAGAGAGCAGCTATACTGGGAGCTGAGCCAGCTAAC
CCACAGCATCACTGAGCTGGGCCCCTACACCCTGGACAGGGACAGTCTCTATGTCAATGGT
TTCACACAGCGGAGCTCTGTGCCCACCACTAGCATTCCTGGGACCCCCACAGTGGACCTGG
GAACATCTGGGACTCCAGTTTCTAAACCTGGTCCCTCGGCTGCCAGCCCTCTCCTGGTGCT
ATTCACTCTCAACTTCACCATCACCAACCTGCGGTATGAGGAGAACATGCAGCACCCTGGC
TCCAGGAAGTTCAACACCACGGAGAGGGTCCTTCAGGGCCTGCTCAGGTCCCTGTTCAAGA
GCACCAGTGTTGGCCCTCTGTACTCTGGCTGCAGACTGACTTTGCTCAGGCCTGAAAAGGA
TGGGACAGCCACTGGAGTGGATGCCATCTGCACCCACCACCCTGACCCCAAAAGCCCTAGG
CTGGACAGAGAGCAGCTGTATTGGGAGCTGAGCCAGCTGACCCACAATATCACTGAGCTGG
GCCACTATGCCCTGGACAACGACAGCCTCTTTGTCAATGGTTTCACTCATCGGAGCTCTGT
GTCCACCACCAGCACTCCTGGGACCCCCACAGTGTATCTGGGAGCATCTAAGACTCCAGCC
TCGATATTTGGCCCTTCAGCTGCCAGCCATCTCCTGATACTATTCACCCTCAACTTCACCA
TCACTAACCTGCGGTATGAGGAGAACATGTGGCCTGGCTCCAGGAAGTTCAACACTACAGA
GAGGGTCCTTCAGGGCCTGCTAAGGCCCTTGTTCAAGAACACCAGTGTTGGCCCTCTGTAC
TCTGGCTCCAGGCTGACCTTGCTCAGGCCAGAGAAAGATGGGGAAGCCACCGGAGTGGATG
CCATCTGCACCCACCGCCCTGACCCCACAGGCCCTGGGCTGGACAGAGAGCAGCTGTATTT
GGAGCTGAGCCAGCTGACCCACAGCATCACTGAGCTGGGCCCCTACACACTGGACAGGGAC
AGTCTCTATGTCAATGGTTTCACCCATCGGAGCTCTGTACCCACCACCAGCACCGGGGTGG
TCAGCGAGGAGCCATTCACACTGAACTTCACCATCAACAACCTGCGCTACATGGCGGACAT
GGGCCAACCCGGCTCCCTCAAGTTCAACATCACAGACAACGTCATGAAGCACCTGCTCAGT
CCTTTGTTCCAGAGGAGCAGCCTGGGTGCACGGTACACAGGCTGCAGGGTCATCGCACTAA
GGTCTGTGAAGAACGGTGCTGAGACACGGGTGGACCTCCTCTGCACCTACCTGCAGCCCCT
CAGCGGCCCAGGTCTGCCTATCAAGCAGGTGTTCCATGAGCTGAGCCAGCAGACCCATGGC
ATCACCCGGCTGGGCCCCTACTCTCTGGACAAAGACAGCCTCTACCTTAACGGTTACAATG
AACCTGGTCTAGATGAGCCTCCTACAACTCCCAAGCCAGCCACCACATTCCTGCCTCCTCT

GTCAGAAGCCACAACAGCCATGGGGTACCACCTGAAGACCCTCACACTCAACTTCACCATC
TCCAATCTCCAGTATTCACCAGATATGGGCAAGGGCTCAGCTACATTCAACTCCACCGAGG
GGGTCCTTCAGCACCTGCTCAGACCCTTGTTCCAGAAGAGCAGCATGGGCCCCTTCTACTT
GGGTTGCCAACTGATCTCCCTCAGGCCTGAGAAGGATGGGGCAGCCACTGGTGTGGACACC
ACCTGCACCTACCACCCTGACCCTGTGGGCCCCGGGCTGGACATACAGCAGCTTTACTGGG
AGCTGAGTCAGCTGACCCATGGTGTCACCCAACTGGGCTTCTATGTCCTGGACAGGGATAG
CCTCTTCATCAATGGCTATGCACCCCAGAATTTATCAATCCGGGGCGAGTACCAGATAAAT
TTCCACATTGTCAACTGGAACCTCAGTAATCCAGACCCCACATCCTCAGAGTACATCACCC
TGCTGAGGGACATCCAGGACAAGGTCACCACACTCTACAAAGGCAGTCAACTACATGACAC
ATTCCGCTTCTGCCTGGTCACCAACTTGACGATGGACTCCGTGTTGGTCACTGTCAAGGCA
TTGTTCTCCTCCAATTTGGACCCCAGCCTGGTGGAGCAAGTCTTTCTAGATAAGACCCTGA
ATGCCTCATTCCATTGGCTGGGCTCCACCTACCAGTTGGTGGACATCCATGTGACAGAAAT
GGAGTCATCAGTTTATCAACCAACAAGCAGCTCCAGCACCCAGCACTTCTACCCGAATTTC
ACCATCACCAACCTACCATATTCCCAGGACAAAGCCCAGCCAGGCACCACCAATTACCAGA
GGAACAAAAGGAATATTGAGGATGCGCTCAACCAACTCTTCCGAAACAGCAGCATCAAGAG
TTATTTTTCTGACTGTCAAGTTTCAACATTCAGGTCTGTCCCCAACAGGCACCACACCGGG
GTGGACTCCCTGTGTAACTTCTCGCCACTGGCTCGGAGAGTAGACAGAGTTGCCATCTATG
AGGAATTTCTGCGGATGACCCGGAATGGTACCCAGCTGCAGAACTTCACCCTGGACAGGAG
CAGTGTCCTTGTGGATGGGTATTCTCCCAACAGAAATGAGCCCTTAACTGGGAATTCTGAC
CTTCCCTTCTGGGCTGTCATCTTCATCGGCTTGGCAGGACTCCTGGGACTCATCACATGCC
TGATCTGCGGTGTCCTGGTGACCACCCGCCGGCGGAAGAAGGAAGGAGAATACAACGTCCA
GCAACAGTGCCCAGGCTACTACCAGTCACACCTAGACCTGGAGGATCTGCAATGACTGGAA
CTTGCCGGTGCCTGGGGTGCCTTTCCCCCAGCCAGGGTCCAAAGAAGCTTGGCTGGGGCAG
AAATAAACCATATTGGTCG

FIGURA 52

GCCCGGGCAGGTGCTGATAGCACAGTTCTGTCCAGAGAAGGAAGGCGGAATAAACTTATTC
ATTCCCAGGAACTCTTGGGGTAGGTGTGTGTTTTTCACATCTTAAAGGCTCACAGACCCTG
CGCTGGACAAATGTTCCATTCCTGAAGGACCTCTCCAGAATCCGGATTGCTGAATCTTCCC
TGTTGCCTAGAAGGGCTCCAAACCACCTCTTGACAATGGGAAACTGGGTGGTTAACCACTG
GTTTTCAGTTTTGTTTCTGGTTGTTTGGTTAGGGCTGAATGTTTTCCTGTTTGTGGATGCC
TTCCTGAAATATGAGAAGGCCGACAAATACTACTACACAAGAAAAATCCTTGGGTCAACAT
TGGCCTGTGCCCGAGCGTCTGCTCTCTGCTTGAATTTTAACAGCACGCTGATCCTGCTTCC
TGTGTGTCGCAATCTGCTGTCCTTCCTGAGGGGCACCTGCTCATTTTGCAGCCGCACACTG
AGAAAGCAATTGGATCACAACCTCACCTTCCACAAGCTGGTGGCCTATATGATCTGCCTAC
ATACAGCTATTCACATCATTGCACACCTGTTTAACTTTGACTGCTATAGCAGAAGCCGACA
GGCCACAGATGGCTCCCTTGCCTCCATTCTCTCCAGCCTATCTCATGATGAGAAAAAGGGG
GGTTCTTGGCTAAATCCCATCCAGTCCCGAAACACGACAGTGGAGTATGTGACATTCACCA
GCGTTGCTGGTCTCACTGGAGTGATCATGACAATAGCCTTGATTCTCATGGTAACTTCAGC
TACTGAGTTCATCCGGAGGAGTTATTTTGAAGTCTTCTGGTATACTCACCACCTTTTTATC
TTCTATATCCTTGGCTTAGGGATTCACGGCATTGGTGGAATTGTCCGGGGTCAAACAGAGG
AGAGCATGAATGAGAGTCATCCTCGCAAGTGTGCAGAGTCTTTTGAGATGTGGGATGATCG
TGACTCCCACTGTAGGCGCCCTAAGTTTGAAGGGCATCCCCCTGAGTCTTGGAAGTGGATC
CTTGCACCGGTCATTCTTTATATCTGTGAAAGGATCCTCCGGTTTTACCGCTCCCAGCAGA
AGGTTGTGATTACCAAGGTTGTTATGCACCCATCCAAAGTTTTGGAATTGCAGATGAACAA
GCGTGGCTTCAGCATGGAAGTGGGGCAGTATATCTTTGTTAATTGCCCCTCAATCTCTCTC
CTGGAATGGCATCCTTTTACTTTGACCTCTGCTCCAGAGGAAGATTTCTTCTCCATTCATA
TCCGAGCAGCAGGGGACTGGACAGAAAATCTCATAAGGGCTTTCGAACAACAATATTCACC
AATTCCCAGGATTGAAGTGGATGGTCCCTTTGGCACAGCCAGTGAGGATGTTTTCCAGTAT
GAAGTGGCTGTGCTGGTTGGAGCAGGAATTGGGGTCACCCCCTTTGCTTCTATCTTGAAAT
CCATCTGGTACAAATTCCAGTGTGCAGACCACAACCTCAAAACAAAAAAGATCTATTTCTA
CTGGATCTGCAGGGAGACAGGTGCCTTTTCCTGGTTCAACAACCTGTTGACTTCCCTGGAA
CAGGAGATGGAGGAATTAGGCAAAGTGGGTTTTCTAAACTACCGTCTCTTCCTCACCGGAT
GGGACAGCAATATTGTTGGTCATGCAGCATTAAACTTTGACAAGGCCACTGACATCGTGAC
AGGTCTGAAACAGAAAACCTCCTTTGGGAGACCAATGTGGGACAATGAGTTTTCTACAATA
GCTACCTCCCACCCCAAGTCTGTAGTGGGAGTTTTCTTATGTGGCCCTCGGACTTTGGCAA
AGAGCCTGCGCAAATGCTGTCACCGATATTCCAGTCTGGATCCTAGAAAGGTTCAATTCTA
CTTCAACAAAGAAAATTTTTGAGTTATAGGAATAAGGACGGTAATCTGCATTTTGTCTCTT
TGTATCTTCAGTAATTGAGTTATAGGAATAAGGACGGTAATCTGCATTTTGTCTCTTTGTA
TCTTCAGTAATTTACTTGGTCTCGTCAGGTTTGAGCAGTCACTTTAGGATAAGAATGTGCC

TCAGCAACTTAAGAACTTCTGAAGTTCTTAAAGTTCTGAAGTTCTTAAAGCCCATGGATCC
TTTCTCAGAAAAATAACTGTAAATCTTTCTGGACAGCCATGACTGTAGCAAGGCTTGATAG
CAGAGGTTTGGTGGTTCAGAGTTATACAACTAATCCCAGGTGATTTTATCAATTCCAGTGT
TACCATCTCCTGAGTTTTGGTTTGTAATCTTTTGTCCCTCCCACCCCCACAGAAGATTTCT
AAGTAGGGTGACTTTTTAAATAAAAATTTATTGAATAATTAATGATAAAACATAATAATAA
ACATAAATAATAAACAAAATTACCGAGAACCCCATCCCCATATAACACCAACAGTGTACAT
GTTTACTGTCACTTTTGATATGGTCTTATCCAGTGTGAACAGCAATTTATTATTTTTGCTC
ATCAAAAAATAAAGGATTTTCTTCTTCACTTGATGAATGAGTCTTTTGTTACTTCTTTTTG
GGCTCTCCATTTGGGAAAGAAAATCTAGAAAGTATGCTCATTGTAATTGAAATTGTATTAT
AATTGCAAGTCAGCTATACTGTCATCTTGTATTATTCTGTTAAAAGCGATTGTGAAAACTT
GAACTCCATGTTAGGATGACTTATTTAGTCAAGAGGAGCTCTTTATCACTATTAAGGGAGT
TTATTAATGTCTAGTTTAAAAAATAACAATTTATAATTTTCATAGATTTTTGTGGACTCCT
ACCTATATCGCTAAGGAATCATTCAGCTAATAACATAAGCAAAATGCCTTCATTCTCTATC
AGAATGCTTTTTATTTCTGCCATCTACCCAACCTCGGCCTACAAGGACCACACTGGCTCTG
GGATGTTAGTCGTCACAGTGAATCCCTAGAAGGAGAGAAGCCCAGGGCTCTTAAGAGAGAA
ACCTCACACTTCATTGTGCCCAGACAATTCAACTCATTATAGATTAGAAGAATGGTTGTAC
TGAATTAGCATTATATTATCTATTTTTTCTACCATTAAGTGACCACGAATGAACAAATGGT
TGTTTCTAATAGTTTAGGAACCTCTACGGCTTGAGGGATAAAACTGTTTTTGGTTTTTTTT
TGTCACTAGTCTGAGCACAATAAAGGTAGGTCATGAACTCATGGTGGCTAATCATGTTTCA
TTTATTACAACAGTAAAACAATACTGATGGTTGTTTACATAGGGAGGGATACCACTGCTGT
TGCCAATTTGTTCACCCGGATTTCAGCTAATTCAGGTTTAATAAGGCTTTGGGAAAATGCA
CAATGAAACCTTGAAGACATGATTCTTATTAGACAACTACTGACCAAATGACAACTTCATC
TTCAATGACCTTGATTTGGAATCTTGAAACTTACAACGAGTTCCTAATGAGAAAGGCATCA
CTCTTGGTCTCCATGTCTAAATATATTGACACACTCACTTCAAGATATGATGGCTGCCAAA
ATCAATATCCCAAGGCTTGAATATTTATGGGAGTAGGCATAAATTCGTGATTATATTCATG
GTAATGAATATAATAAATCCTATGCATAGCCAACATCCAAGTTGAACTGGGGATGAGCAGG
ATACACAGAAAGCTCCTCACCTGCTTTGAGCAAGTTTTTCTTGCAAAGTTGTAATATGAAA
TACCTTAATCAACAAATTTCTTATAAATGTTTATATGAGTAGAAAATCCCTCAATTTCATT
GTCACTCAAAAAGTTGATTATCT

FIGURA 53

CCCTTGGGCTCCAAACCACCTCTTGACAATGGGAAACTGGGTGGTTAACCACTGGTTTTCA
GTTTTGTTTCTGGTTGTTTGGTTAGGGCTGAATGTTTTCCTGTTTGTGGATGCCTTCCTGA
AATATGAGAAGGCCGACAAATACTACTACACAAGAAAAATCCTTGGGTCAACATTGGCCTG
TGCCCGAGCGTCTGCTCTCTGCTTGAATTTTAACAGCACGCTGATCCTGCTTCCTGTGTGT
CGCAATCTGCTGTCCTTCCTGAGGGGCACCTGCTCATTTTGCAGCCGCACACTGAGAAAGC
AATTGGATCACAACCTCACCTTCCACAAGCTGGTGGCCTATATGATCTGCCTACATACAGC
TATTCACATCATTGCACACCTGTTTAACTTTGACTGCTATAGCAGAAGCCGACAGGCCACA
GATGGCTCCCTTGCCTCCATTCTCTCCAGCCTATCTCATGATGAGAAAAAGGGGGGTTCTT
GGCTAAATCCCATCCAGTCCCGAAACACGACAGTGGAGTATGTGACATTCACCAGCATTGC
TGGTCTCACTGGAGTGATCATGACAATAGCCTTGATTCTCATGGTAACTTCAGCTACTGAG
TTCATCCGGAGGAGTTATTTTGAAGTCTTCTGGTATACTCACCACCTTTTTATCTTCTATA
TCCTTGGCTTAGGGATTCACGGCATTGGTGGAATTGTCCGGGGTCAAACAGAGGAGAGCAT
GAATGAGAGTCATCCTCGCAAGTGTGCAGAGTCTTTTGAGATGTGGGATGATCGTGACTCC
CACTGTAGGCGCCCTAAGTTTGAAGGGCATCCCCCTGAGTCTTGGAAGTGGATCCTTGCAC
CGGTCATTCTTTATATCTGTGAAAGGATCCTCCGGTTTTACCGCTCCCAGCAGAAGGTTGT
GATTACCAAGGTTGTTATGCACCCATCCAAAGTTTTGGAATTGCAGATGAACAAGCGTGGC
TTCAGCATGGAAGTGGGGCAGTATATCTTTGTTAATTGCCCCTCAATCTCTCTCCTGGAAT
GGCATCCTTTTACTTTGACCTCTGCTCCAGAGGAAGATTTCTTCTCCATTCATATCCGAGC
AGCAGGGGACTGGACAGAAAATCTCATAAGGGCTTTCGAACAACAATATTCACCAATTCCC
AGGATTGAAGTGGATGGTCCCTTTGGCACAGCCAGTGAGGATGTTTTCCAGTATGAAGTGG
CTGTGCTGGTTGGAGCAGGAATTGGGGTCACCCCCTTTGCTTCTATCTTGAAATCCATCTG
GTACAAATTCCAGTGTGCAGACCACAACCTCAAAACAAAAAAGATCTATTTCTACTGGATC
TGCAGGGAGACAGGTGCCTTTTCCTGGTTCAACAACCTGTTGACTTCCCTGGAACAGGAGA
TGGAGGAATTAGGCAAAGTGGGTTTTCTAAACTACCGTCTCTTCCTCACCGGATGGGACAG
CAATATTGTTGGTCATGCAGCATTAAACTTTGACAAGGCCACTGACATCGTGACAGGTCTG
AAACAGAAAACCTCCTTTGGGAGACCAATGTGGGACAATGAGTTTTCTACAATAGCTACCT
CCCACCCCAAGTCTGTAGTGGGAGTTTTCTTATGTGGCCCTCGGACTTTGGCAAAGAGCCT
GCGCAAATGCTGTCACCGATATTCCAGTCTGGATCCTAGAAAGGTTCAATTCTACTTCAAC
AAAGAAAATTTTTGAGTTATAGGAATAAGGACGGTAATCTGCATTTTGTCTCTTTGTATCT
TCAGTAATTTACTTGGTCTCGTCAGGTTTGAGCAGTCACTTTAGGAAG

FIGURA 54

GGGCTCCAAACCACCTCTTGACAATGGGAAACTGGGTGGTTAACCACTGGTTTTCAGTTTT
GTTTCTGGTTGTTTGGTTAGGGCTGAATGTTTTCCTGTTTGTGGATGCCTTCCTGAAATAT
GAGAAGGCCGACAAATACTACTACACAAGAAAAATCCTTGGGTCTTGGAAGTGGATCCTTG
CACCGGTCATTCTTTATATCTGTGAAAGGATCCTCCGGTTTTACCGCTCCCAGCAGAAGGT
TGTGATTACCAAGGTTGTTATGCACCCATCCAAAGTTTTGGAATTGCAGATGAACAAGCGT
GGCTTCAGCATGGAAGTGGGGCAGTATATCTTTGTTAATTGCCCCTCAATCTCTCTCCTGG
AATGGCATCCTTTTACTTTGACCTCTGCTCCAGAGGAAGATTTCTTCTCCATTCATATCCG
AGCAGCAGGGGACTGGACAGAAAATCTCATAAGGGCTTTCGAACAACAATATTCACCAATT
CCCAGGATTGAAGTGGATGGTCCCTTTGGCACAGCCAGTGAGGATGTTTTCCAGTATGAAG
TGGCTGTGCTGGTTGGAGCAGGAATTGGGGTCACCCCCTTTGCTTCTATCTTGAAATCCAT
CTGGTACAAATTCCAGTGTGCAGACCACAACCTCAAAACAAAAAAGGTTGGTCATGCAGCA
TTAAACTTTGACAAGGCCACTGACATCGTGACAGGTCTGAAACAGAAAACCTCCTTTGGGA
GACCAATGTGGGACAATGAGTTTTCTACAATAGCTACCTCCCACCCCAAGTCTGTAGTGGG
AGTTTTCTTATGTGGCCCTCGGACTTTGGCAAAGAGCCTGCGCAAATGCTGTCACCGATAT
TCCAGTCTGGATCCTAGAAAGGTTCAATTCTACTTCAACAAAGAAAATTTTTGAGTTATAG
GAATAAGGACGGTAATCTGCATTTTGTCTCTTTGTATCTTCAGTAATTTACTTGGTCTCGT
CAGGTTTGAGCAGTCACTTTAGGA

FIGURA 55

GGGCTCCAAACCACCTCTTGACAATGGGAAACTGGGTGGTTAACCACTGGTTTTCAGTTTT
GTTTCTGGTTGTTTGGTTAGGGCTGAATGTTTTCCTGTTTGTGGATGCCTTCCTGAAATAT
GAGAAGGCCGACAAATACTACTACACAAGAAAAATCCTTGGGTCAACATTGGCCTGTGCCC
GAGCGTCTGCTCTCTGCTTGAATTTTAACAGCACGCTGATCCTGCTTCCTGTGTGTCGCAA
TCTGCTGTCCTTCCTGAGGGGCACCTGCTCATTTTGCAGCCGCACACTGAGAAAGCAATTG
GATCACAACCTCACCTTCCACAAGCTGGTGGCCTATATGATCTGCCTACATACAGCTATTC
ACATCATTGCACACCTGTTTAACTTTGACTGCTATAGCAGAAGCCGACAGGCCACAGATGG
CTCCCTTGCCTCCATTCTCTCCAGCCTATCTCATGATGAGAAAAAGGGGGGTTCTTGGCTA
AATCCCATCCAGTCCCGAAACACGACAGTGGAGTATGTGACATTCACCAGCATTGCTGGTC
TCACTGGAGTGATCATGACAATAGCCTTGATTCTCATGGTAACTTCAGCTACTGAGTTCAT
CCGGAGGAGTTATTTTGAAGTCTTCTGGTATACTCACCACCTTTTTATCTTCTATATCCTT
GGCTTAGGGATTCACGGCATTGGTGGAATTGTCCGGGGTCAAACAGAGGAGAGCATGAATG
AGAGTCATCCTCGCAAGTGTGCAGAGTCTTTTGAGATGTGGGATGATCGTGACTCCCACTG
TAGGCGCCCTAAGTTTGAAGGGCATCCCCCTGAGTCTTGGAAGTGGATCCTTGCACCGGTC
ATTCTTTATATCTGTGAAAGGATCCTCCGGTTTTACCGCTCCCAGCAGAAGGTTGTGATTA
CCAAGGTTGTTATGCACCCATCCAAAGTTTTGGAATTGCAGATGAACAAGCGTGGCTTCAG
CATGGAAGTGGGGCAGTATATCTTTGTTAATTGCCCCTCAATCTCTCTCCTGGAATGGCAT
CCTTTTACTTTGACCTCTGCTCCAGAGGAAGATTTCTTCTCCATTCATATCCGAGCAGCAG
GGGACTGGACAGAAAATCTCATAAGGGCTTTCGAACAACAATATTCACCAATTCCCAGGAT
TGAAGTGGATGGTCCCTTTGGCACAGCCAGTGAGGATGTTTTCCAGTATGAAGTGGCTGTG
CTGGTTGGAGCAGGAATTGGGGTCACCCCCTTTGCTTCTATCTTGAAATCCATCTGGTACA
AATTCCAGTGTGCAGACCACAACCTCAAAACAAAAAAGGTTGGTCATGCAGCATTAAACTT
TGACAAGGCCACTGACATCGTGACAGGTCTGAAACAGAAAACCTCCTTTGGGAGACCAATG
TGGGACAATGAGTTTTCTACAATAGCTACCTCCCACCCCAAGTCTGTAGTGGGAGTTTTCT
TATGTGGCCCTCGGACTTTGGCAAAGAGCCTGCGCAAATGCTGTCACCGATATTCCAGTCT
GGATCCTAGAAAGGTTCAATTCTACTTCAACAAAGAAAATTTTTGAGTTATAGGAATAAGG
ACGGTAATCTGCATTTTGTCTCTTTGTATCTTCAGTAATTTACTTGGTCTCGTCAGGTTTG
AGCAGTCACTTTAGGAAG

FIGURA 56

GGACTGCTTTGTAACTGCTAAGATTGCAGACAGAAATAGCACACAACCACTGTGAGCTGTA
TGCGATTCAGAAACCAAGACCAAATTTTGCTCACTTTCATTAATCAGTTGCTCAGATAGAA
GGAAATGACATCTGGTTCTGTCTTCTTCTACATCTTAATTTTTGGAAAATATTTTTCTCAT
GGGGGTGGACAGGATGTCAAGTGCTCCCTTGGCTATTTCCCCTGTGGGAACATCACAAAGT
GCTTGCCTCAGCTCCTGCACTGTAACGGTGTGGACGACTGCGGGAATCAGGCCGATGAGGA
CAACTGTGGAGACAACAATGGATGGTCCATGCAATTTGACAAATATTTTGCCAGTTACTAC
AAAATGACTTCCCAATATCCTTTTGAGGCAGAAACACCTGAATGTTTGGTCGGTTCTGTGC
CAGTGCAATGTCTTTGCCAAGGTCTGGAGCTTGACTGTGATGAAACCAATTTACGAGCTGT
TCCATCGGTTTCTTCAAATGTGACTGCAATGTCACTTCAGTGGAACTTAATAAGAAAGCTT
CCTCCTGATTGCTTCAAGAATTATCATGATCTTCAGAAGCTGTACCTGCAAAACAATAAGA
TTACATCCATCTCCATCTATGCTTTCAGAGGACTGAATAGCCTTACTAAACTGTATCTCAG
TCATAACAGAATAACCTTCCTGAAGCCGGGTGTTTTTGAAGATCTTCACAGACTAGAATGG
CTGATAATTGAAGATAATCACCTCAGTCGAATTTCCCCACCAACATTTTATGGACTAAATT
CTCTTATTCTCTTAGTCCTGATGAATAACGTCCTCACCCGTTTACCTGATAAACCTCTCTG
TCAACACATGCCAAGACTACATTGGCTGGACCTTGAAGGCAACCATATCCATAATTTAAGA
AATTTGACTTTTATTTCCTGCAGTAATTTAACTGTTTTAGTGATGAGGAAAAACAAAATTA
ATCACTTAAATGAAAATACTTTTGCACCTCTCCAGAAACTGGATGAATTGGATTTAGGAAG
TAATAAGATTGAAAATCTTCCACCGCTTATATTCAAGGACCTGAAGGAGCTGTCACAATTG
AATCTTTCCTATAATCCAATCCAGAAAATTCAAGCAAACCAATTTGATTATCTTGTCAAAC
TCAAGTCTCTCAGCCTAGAAGGGATTGAAATTTCAAATATCCAACAAAGGATGTTTAGACC
TCTTATGAATCTCTCTCACATATATTTTAAGAAATTCCAGTACTGTGGGTATGCACCACAT
GTTCGCAGCTGTAAACCAAACACTGATGGAATTTCATCTCTAGAGAATCTCTTGGCAAGCA
TTATTCAGAGAGTATTTGTCTGGGTTGTATCTGCAGTTACCTGCTTTGGAAACATTTTTGT
CATTTGCATGCGACCTTATATCAGGTCTGAGAACAAGCTGTATGCCATGTCAATCATTTCT
CTCTGCTGTGCCGACTGCTTAATGGGAATATATTTATTCGTGATCGGAGGCTTTGACCTAA
AGTTTCGTGGAGAATACAATAAGCATGCGCAGCTGTGGATGGAGAGTACTCATTGTCAGCT
TGTAGGATCTTTGGCCATTCTGTCCACAGAAGTATCAGTTTTACTGTTAACATTTCTGACA
TTGGAAAAATACATCTGCATTGTCTATCCTTTTAGATGTGTGAGACCTGGAAAATGCAGAA
CAATTACAGTTCTGATTCTCATTTGGATTACTGGTTTTATAGTGGCTTTCATTCCATTGAG
CAATAAGGAATTTTTCAAAAACTACTATGGCACCAATGGAGTATGCTTCCCTCTTCATTCA
GAAGATACAGAAAGTATTGGAGCCCAGATTTATTCAGTGGCAATTTTTCTTGGTATTAATT
TGGCCGCATTTATCATCATAGTTTTTTCCTATGGAAGCATGTTTTATAGTGTTCATCAAAG
TGCCATAACAGCAACTGAAATACGGAATCAAGTTAAAAAAGAGATGATCCTTGCCAAACGT
TTTTTCTTTATAGTATTTACTGATGCATTATGCTGGATACCCATTTTTGTAGTGAAATTTC
TTTCACTGCTTTCAGGTAGAAATACCAGGGTACCATAACCTCTTGGGTAGTGATTTTTATTCT

GCCCATTAACAGTGCTTTGAACCCAATTCTCTATACTCTGACCACAAGACCATTTAAAGAA
ATGATTCATCGGTTTTGGTATAACTACAGACAAAGAAAATCTATGGACAGCAAAGGTCAGA
AAACATATGCTCCATCATTCATCTGGGTGGAAATGTGGCCACTGCAGGAGATGCCACCTGA
GTTAATGAAGCCGGACCTTTTCACATACCCCTGTGAAATGTCACTGATTTCTCAATCAACG
AGACTCAATTCCTATTCATGA

FIGURA 57

TAGAGATCCCTCGACCTCGACCCACGCGTCCGAGGAAAGAAAAAAAGAGGAATGGAAAGAG
ACAGAGAAAGGAAATGGGAGTGGAAGGAGGGAGGACTGCTTTGTAACTGCTAAGATTGCAG
ACAGAAATAGCACACAACCACTGTGAGCTGTATGCGATTCAGAAACCAAGACCAAATTTTG
CTCACTTTCATTAATCAGTTGCTCAGATAGAAGGAAATGACATCTGGTTCTGTCTTCTTCT
ACATCTTAATTTTTGGAAAATATTTTTCTCATGGGGGTGGACAGGATGTCAAGTGCTCCCT
TGGCTATTTCCCCTGTGGGAACATCACAAAGTGCTTGCCTCAGCTCCTGCACTGTAACGGT
GTGGACGACTGCGGGAATCAGGCCGATGAGGACAACTGTGGAGACAACAATGGATGGTCTC
TGCAATTTGACAAATATTTTGCCAGTTACTACAAAATGACTTCCCAATATCCTTTTGAGGC
AGAAACACCTGAATGTTTGGTCGGTTCTGTGCCAGTGCAATGTCTTTGCCAAGGTCTGGAG
CTTGACTGTGATGAAACCAATTTACGAGCTGTTCCATCGGTTTCTTCAAATGTGACTGCAA
TGTCACTTCAGTGGAACTTAATAAGAAAGCTTCCTCCTGATTGCTTCAAGAATTATCATGA
TCTTCAGAAGCTGTACCTGCAAAACAATAAGATTACATCCATCTCCATCTATGCTTTCAGA
GGACTGAATAGCCTTACTAAACTGTATCTCAGTCATAACAGAATAACCTTCCTGAAGCCGG
GTGTTTTTGAAGATCTTCACAGACTAGAATGGCTGATAATTGAAGATAATCACCTCAGTCG
AATTTCCCCACCAACATTTTATGGACTAAATTCTCTTATTCTCTTAGTCCTGATGAATAAC
GTCCTCACCCGTTTACCTGATAAACCTCTCTGTCAACACATGCCAAGACTACATTGGCTGG
ACCTTGAAGGCAACCATATCCATAATTTAAGAAATTTGACTTTTATTTCCTGCAGTAATTT
AACTGTTTTAGTGATGAGGAAAAACAAAATTAATCACTTAAATGAAAATACTTTTGCACCT
CTCCAGAAACTGGATGAATTGGATTTAGGAAGTAATAAGATTGAAAATCTTCCACCGCTTA
TATTCAAGGACCTGAAGGAGCTGTCACAATTGAATCTTTCCTATAATCCAATCCAGAAAAT
TCAAGCAAACCAATTTGATTATCTTGTCAAACTCAAGTCTCTCAGCCTAGAAGGGATTGAA
ATTTCAAATATCCAACAAAGGATGTTTAGACCTCTTATGAATCTCTCTCACATATATTTTA
AGAAATTCCAGTACTGTGGGTATGCACCACATGTTCGCAGCTGTAAACCAAACACTGATGG
AATTTCATCTCTAGAGAATCTCTTGGCAAGCATTATTCAGAGAGTATTTGTCTGGGTTGTA
TCTGCAGTTACCTGCTTTGGAAACATTTTTGTCATTTGCATGCGACCTTATATCAGGTCTG
AGAACAAGCTGTATGCCATGTCAATCATTTCTCTCTGCTGTGCCGACTGCTTAATGGGAAT
ATATTTATTCGTGATCGGAGGCTTTGACCTAAAGTTTCGTGGAGAATACAATAAGCATGCG
CAGCTGTGGATGGAGAGTACTCATTGTCAGCTTGTAGGATCTTTGGCCATTCTGTCCACAG
AAGTATCAGTTTTACTGTTAACATTTCTGACATTGGAAAAATACATCTGCATTGTCTATCC
TTTTAGATGTGTGAGACCTGGAAAATGCAGAACAATTACAGTTCTGATTCTCATTTGGATT
ACTGGTTTTATAGTGGCTTTCATTCCATTGAGCAATAAGGAATTTTTCAAAAACTACTATG
GCACCAATGGAGTATGCTTCCCTCTTCATTCAGAAGATACAGAAAGTATTGGAGCCCAGAT
TTATTCAGTGGCAATTTTTCTTGGTATTAATTTGGCCGCATTTATCATCATAGTTTTTTCC
TATGGAAGCATGTTTTATAGTGTTCATCAAAGTGCCATAACAGCAACTGAAATACGGAATC
AAGTTTAAAAAAGGAGATGATCCTTTGCCAAACGTTTTTTTCTTTTATAGTATTTACTGATGCATT

ATGCTGGATACCCATTTTTGTAGTGAAATTTCTTTCACTGCTTCAGGTAGAAATACCAGGT
ACCATAACCTCTTGGGTAGTGATTTTTATTCTGCCCATTAACAGTGCTTTGAACCCAATTC
TCTATACTCTGACCACAAGACCATTTAAAGAAATGATTCATCGGTTTTGGTATAACTACAG
ACAAAGAAAATCTATGGACAGCAAAGGTCAGAAAACATATGCTCCATCATTCATCTGGGTG
GAAATGTGGCCACTGCAGGAGATGCCACCTGAGTTAATGAAGCCGGACCTTTTCACATACC
CCTGTGAAATGTCACTGATTTCTCAATCAACGAGACTCAATTCCTATTCATGACTGACTCT
GAAATTCATTTCTTCGCAGAGAATACTGTGGGGGTGCTTCATGAGGGATTTACTGGTATGA
AATGAATACCACAAAATTAATTTATAATAATAGCTAAGATAAATATTTTACAAGGACATGA
GGAAAAATAAAAATGACTAATGCTCTTACAAAGGGAAGTAATTATATCAATAATGTATATA
TATTAGTAGACATTTTGCATAAGAAATTAAGAGAAATCTACTTCAGTAACATTCATTCATT
TTTCTAACATGCATTTATTGAGTACCCACTACTATGTGCATAGCATTGCAATATAGTCCTG
GAAGTAGACAGTGCAGAACCTTTCAATCTGTAGATGGTGTTTAATGACAAAAGACTATACA
AAGTCCATCTGCAGTTCCTAGTTTAAAGTAGAGCTTTACCTGTCATGTGCATCAGCAAGAA
TCATAGGCACTTTTAAATAAAGGTTTAAAGTTTTGGAAT

FIGURA 58

GATCCATTGTGCTGGCGCGGATTCTTTATCACTGGAAGGAGGAGGAGGGGGAGAGTCGCTC
CCGCCGGGCGAGCATGGGGCGCCTGGCCTCGAGGCCGCTGCTGCTGGCGCTCCTGTCGTTG
GCTCTTTGCCGAGGGCGTGTGGTGAGAGTCCCCACAGCGACCCTGGTTCGAGTGGTGGGCA
CTGAGCTGGTCATCCCCTGCAACGTCAGTGACTATGATGGCCCCAGCGAGCAAAACTTTGA
CTGGAGCTTCTCATCTTTGGGGAGCAGCTTTGTGGAGCTTGCAAGCACCTGGGAGGTGGGG
TTCCCAGCCCAGCTGTACCAGGAGCGGCTGCAGAGGGGCGAGATCCTGTTAAGGCGGACTG
CCAACGACGCCGTGGAGCTCCACATAAAGAACGTCCAGCCTTCAGACCAAGGCCACTACAA
ATGTTCAACCCCCAGCACAGATGCCACTGTCCAGGGAAACTATGAGGACACAGTGCAGGTT
AAAGTGCTGGCCGACTCCCTGCACGTGGGCCCCAGCGCGCGGCCCCCGCCGAGCCTGAGCC
TGCGGGAGGGGGAGCCCTTCGAGCTGCGCTGCACCGCCGCCTCCGCCTCGCCGCTGCACAC
GCACCTGGCGCTGCTGTGGGAGGTGCACCGCGGCCCGGCCAGGCGGAGCGTCCTCGCCCTG
ACCCACGAGGGCAGGTTCCACCCGGGCCTGGGGTACGAGCAGCGCTACCACAGTGGGGACG
TGCGCCTCGACACCGTGGGCAGCGACGCCTACCGCCTCTCAGTGTCCCGGGCTCTGTCTGC
CGACCAGGGCTCCTACAGGTGTATCGTCAGCGAGTGGATCGCCGAGCAGGGCAACTGGCAG
GAAATCCAAGAAAAGGCCGTGGAAGTTGCCACCGTGGTGATCCAGCCGACAGTTCTGCGAG
CAGCTGTGCCCAAGAATGTGTCTGTGGCTGAAGGAAAGGAACTGGACCTGACCTGTAACAT
CACAACAGACCGAGCCGATGACGTCCGGCCCGAGGTGACGTGGTCCTTCAGCAGGATGCCT
GACAGCACCCTACCTGGCTCCCGCGTGTTGGCGCGGCTTGACCGTGATTCCCTGGTGCACA
GCTCGCCTCATGTTGCTTTGAGTCATGTGGATGCACGCTCCTACCATTTACTGGTTCGGGA
TGTTAGCAAAGAAAACTCTGGCTACTATTACTGCCACGTGTCCCTGTGGGCACCCGGACAC
AACAGGAGCTGGCACAAAGTGGCAGAGGCCGTGTCTTCCCCAGCTGGTGTGGGTGTGACCT
GGCTAGAACCAGACTACCAGGTGTACCTGAATGCTTCCAAGGTCCCCGGGTTTGCGGATGA
CCCCACAGAGCTGGCATGCCGGGTGGTGGACACGAAGAGTGGGGAGGCGAATGTCCGATTC
ACGGTTTCGTGGTACTACAGGATGAACCGGCGCAGCGACAATGTGGTGACCAGCGAGCTGC
TTGCAGTCATGGACGGGGACTGGACGCTAAAATATGGAGAGAGGAGCAAGCAGCGGGCCCA
GGATGGAGACTTTATTTTTTCTAAGGAACATACAGACACGTTCAATTTCCGGATCCAAAGG
ACTACAGAGGAAGACAGAGGCAATTATTACTGTGTTGTGTCTGCCTGGACCAAACAGCGGA
ACAACAGCTGGGTGAAAAGCAAGGATGTCTTCTCCAAGCCTGTTAACATATTTTGGGCATT
AGAAGATTCCGTGCTTGTGGTGAAGGCGAGGCAGCCAAAGCCTTTCTTTGCTGCCGGAAAT
ACATTTGAGATGACTTGCAAAGTATCTTCCAAGAATATTAAGTCGCCACGCTACTCTGTTC
TCATCATGGCTGAGAAGCCTGTCGGCGACCTCTCCAGTCCCAATGAAACGAAGTACATCAT
CTCTCTGGACCAGGATTCTGTGGTGAAGCTGGAGAATTGGACAGATGCATCACGGGTGGAT
GGCGTTGTTTTAGAAAAAGTGCAGGAGGATGAGTTCCGCTATCGAATGTACCAGACTCAGG
TCTCAGACGCAGGGCTGTACCGCTGCATGGTGACAGCCTGGTCTCCTGTCAGGGGCAGCCT
TTGGCGGAGAAGCAGCAACGAGTCTTCTCCAATCCTATTGAGATAGACTTCCAAACCTCAGGT

CCTATATTTAATGCTTCTGTGCATTCAGACACACCATCAGTAATTCGGGGAGATCTGATCA
AATTGTTCTGTATCATCACTGTCGAGGGAGCAGCACTGGATCCAGATGACATGGCCTTTGA
TGTGTCCTGGTTTGCGGTGCACTCTTTTGGCCTGGACAAGGCTCCTGTGCTCCTGTCTTCC
CTGGATCGGAAGGGCATCGTGACCACCTCCCGGAGGGACTGGAAGAGCGACCTCAGCCTGG
AGCGCGTGAGTGTGCTGGAATTCTTGCTGCAAGTGCATGGCTCCGAGGACCAGGACTTTGG
CAACTACTACTGTTCCGTGACTCCATGGGTGAAGTCACCAACAGGTTCCTGGCAGAAGGAG
GCAGAGATCCACTCCAAGCCCGTTTTTATAACTGTGAAGATGGATGTGCTGAACGCCTTCA
AGTATCCCTTGCTGATCGGCATCGGTCTGTCCACGGTCATCGGGCTCCTGTCCTGTCTCAT
CGGGTACTGCAGCTCCCACTGGTGTTGTAAGAAGGAGGTTCAGGAGACACGGCGCGAGCGC
CGCAGGCTCATGTCGATGGAGATGGACTAGGCTGGCCCGGGAGGGGAGTGACAGAGGGACG
TTCTAGGAGCAATTGGGGCAAGAAGAGGACAGTGATATTTTAAAACAAAGTGTGTTACACT
AAAAACCAGTCCTCTCTAATCTCAGGTGGGACTTGGCGCTCTCTCTTTTCTGCATGTCAAG
TTCTGAGCGCGGACATGTTTACCAGCACACGGCTCTTCTTCCCACGGCACTTTCTGATGTA
ACAATCGAGTGTGTGTTTTCCCAACTGCCAGTGATAAAGAATCCGCGCCAGCACAATGGAT
CTCGAGGGATCTTCCATACCTACCAGTTCTGCGCCTGCAGGTC.

FIGURA 59

CCCACGCGTCCGCGGTAAATCACTTCTTGGAGGTGCCCTGCACGCCGGTCCTGGGAGCAGG
CGGCCTCCCGGGGGTGCGGGAGCCCCACTCCTCCGTGGTGTGTTCCATTTGCTTCCCACAT
CTGGAGGAGCTGACGTGCCAGCCTCCCCCAGCACCACCCAGGGACGGGAGGCATGAGCCGG
TCAAGGCACCTGGGCAAAATCCGGAAGCGTCTGGAAGATGTCAAGAGCCAGTGGGTCCGGC
CAGCCAGGGCTGACTTTAGTGACAACGAGAGTGCCCGGCTGGCCACGGACGCCCTCTTGGA
TGGGGGTTCTGAAGCCTACTGGCGGGTGCTCAGCCAGGAAGGCGAGGTGGACTTCTTGTCC
TCGGTGGAGGCCCAGTACATCCAGGCCCAGGCCAGGGAGCCCCCGTGTCCCCCAGACACCC
TGGGAGGGGCGGAAGCAGGCCCTAAGGGACTGGACTCCAGCTCCCTACAGTCCGGCACCTA
CTTCCCTGTGGCCTCAGAGGGCAGCGAGCCGGCCCTACTGCACAGCTGGGCCTCAGCTGAG
AAGCCCTACCTGAAGGAAAAATCCAGCGCCACTGTGTACTTCCAGACCGTCAAGCACAACA
ACATCAGAGACCTCGTCCGCCGCTGCATCACCCGGACTAGCCAGGTCCTGGTCATCCTGAT
GGATGTGTTCACGGATGTGGAGATCTTCTGTGACATTCTAGAGGCAGCCAACAAGCGTGGG
GTGTTCGTTTGTGTGCTCCTGGACCAGGGAGGTGTGAAGCTCTTCCAGGAGATGTGTGACA
AAGTCCAGATCTCTGACAGTCACCTCAAGAACATTTCCATCCGGAGTGTGGAAGGAGAGAT
ATACTGTGCCAAGTCAGGCAGGAAATTCGCTGGCCAAATCCGGGAGAAGTTCATCATCTCG
GACTGGAGATTTGTCCTGTCTGGATCTTACAGCTTCACCTGGCTCTGCGGACACGTGCACC
GGAACATCCTCTCCAAGTTCACAGGCCAGGCGGTGGAGCTGTTTGACGAGGAGTTCCGCCA
CCTCTACGCCTCCTCCAAGCCTGTGATGGGCCTGAAGTCCCCGCGGCTGGTCGCCCCCGTC
CCGCCCGGAGCAGCCCCGGCCAATGGCCGCCTTAGCAGCAGCAGTGGCTCCGCCAGTGACC
GCACGTCCTCCAACCCCTTCAGCGGCCGCTCGGCAGGCAGCCACCCCGGTACCCGAAGTGT
GTCCGCGTCTTCAGGGCCCTGTAGCCCCGCGGCCCCACACCCGCCTCCACCGCCCCGGTTC
CAGCCCCACCAAGGCCCTTGGGGAGCCCCGAGTCCCCAGGCCCACCTCTCCCCGCGGCCCC
ACGACGGCCCGCCCGCCGCTGTCTACAGCAACCTGGGGGCCTACAGGCCCACGCGGCTGCA
GCTGGAGCAGCTGGGCCTGGTGCCGAGGCTGACTCCAACCTGGAGGCCCTTCCTGCAGGCC
TCCCCTCACTTCTGAAGGTCCCATCCCCTGCTGCCCTCCGCAGGCCCAGGGCTGGGCACTC
CCTGAGACCCAAAGACCCACCTCAACGACGAGTGGCGTTGAGCCACTTCCCTTTGAAAAGA
CACTCAAAATCACTGCCATGGTTCAATGTTGCCAGGCCCCAGGCCATCCACTTGCCGGCCC
CCACCAGTTCTTGGGTTCCCCGCTCTAGTTTGACCTGTGCAGCACATTCCAGAAGGTTCCA
GGGGAGGTTGTGGGGCAGCTAGAGGACAAAATCATGAAAACAGAGTCCCTGTCTTCCAGAG
ATCATCCGGGGCTTTAATATTAATGGCCCCCAAAACTCCGTAAGAAGCAGGAAATGCAGCC
CAAGTTTTACAAATGGGTAAACAGAGGCACTGAGAGATAGATGGTAGTTTGGTACTTCTGG
TTCCCAGTGCCCAGGAATGGTCCACTCCCAAGAAATTCAGGAAAGAAAGACTGAGGAGAAG
GTGTGGGAACATTCTGGATGTTTCGGGAGAGTTGGGGAAACTCCTCCTCTTAGGAAAGGCT
AATACTAGGGTATCCTTGGGCCCAATGAATTAGGGGTGAGGCCCCAGAACCCGTTATCTAT
GAGTTGTATGGGGGAGCCATCTTGAAGCTGTAGCCCACCAGGGATGCAGCTAGCTGAGGAGTT

TGGGGTGTTGGGTTGGACAAGGCAGGTTAGTAGACTCAGATTCTTGCTTCAAAGAGCCTTG
GGCTGGCCTGGAGGTCCCTGGAGTCTAGACTGGACCTAGGAGCTTGAGTTGTCAGGGGCCA
GGACTGGCCCCACTGCAGTGCCCAGGCCAGTCTTGAGCAGCAGGGAGGGCTCAGCTGTCCC
CAGATCCAGGTGCCTCTGACCAGCCTGGTCACCTCCTGAGGAATAAATGCTGAACCTCACA
AGCCCCATCATTCATTTCTTCTCAATTCACAGTGCCCCTCTTTGTTTCTGGGGTGGAACTA
GGTCCTGAGGGCACAGCCTAGCTGAGTGCAAAGAAATATAGGATGCTTAGAAAGCATACAG
GAGGGGCCAGGCGTGGTGGCTCATGCCTGTAATCCCAGAACTTTGGGATGCCAAGGTGGTT
GGATTACCTGAGATCAGGTGGATTACCTGGTCTCGAGACCAGCCTGACCAATATGGTGAAA
CCCCGTCTCTACTAAAAATACAAAAATTAGGCTGAGACAGGAGAATTGCTTGAACCCAGGA
AGCAGAGGTTGCAATGAGCTGAGATTGCATCACTGCACTCCAGCATGGGCAACAAAGCAAG
ACTCCGTCACAGAAAAAAAAAAAAAA

FIGURA 60

ATGGCTCTGCCGCAGCCGCCGCCGCCCGCCGGGGCGCGTGACCCGGTGACCCCGGAGCACT
GGGCATCCGGACCCCCCTCGCAGGCGCAGCCGCTATTAAGGCAGGAGGCTAAAGAGGAGGA
GGAAGGGGAGGAGACAGGCGTCCAGGGCGCCTGGGGAACCGGCACGGCGGAGCAGCGGCGG
CGGGGCTGGGGGGAGGCCGCCGAGTCGGCCGCGGCCGAGGAGGGGCAGGCGGAGGTCGGGG
GCGCCGCGGCCGCGGGGTCAGGTTCCCCGGCGGGAGGCGCAGGTGGCGGCCTGGGGAGCTG
GCGGCCGCTCCTCGCCTGGCTGCAGCGGCGGCAGCCCCAGTGCTGCCCCTGTGCGGCGCCC
CTTTCCCGCTCCGCCGCGCACTGTTGTCATGGAGGAACCAAGATGGCGGCTCTGGCCTACA
ACCTGGGCAAGCGGGAGATCAACCACTACTTCAGCGTGAGGAGCGCCAAGGTGCTGGCGCT
GGTGGCCGTGCTGCTGCTCGCAGCGTGCCACCTCGCCTCCCGCCGCTACCGAGGCAATGAT
TCGTGTGAATACCTTCTCTCAAGTGGCAGATTTCTTGGAGAGAAAGTTTGGCAACCTCACA
GTTGTATGATGCATAAATACAAAATCAGTGAAGCAAAGAACTGCCTTGTAGATAAACATAT
TGCATTTATTGGAGATTCCAGAATTCGTCAATTGTTTTATTCTTTTGTAAAAATAATTAAT
CCCCAATTCAAAGAAGAAGGAAATAAGCATGAAAACATTCCTTTTGAAGACAAGACTGCAT
CAGTTAAAGTGGATTTTCTGTGGCATCCTGAAGTTAATGGTTCTATGAAACAGTGTATCAA
AGTGTGGACTGAGGATTCCATTGCAAAGCCACATGTGATTGTAGCAGGAGCTGCCACATGG
TCCATCAAGATTCACAATGGTAGCAGTGAAGCGCTTTCTCAATATAAAATGAACATCACCT
CCATAGCACCACTTTTAGAAAAATTGGCAAAGACTAGTGATGTTTATTGGGTCTTACAAGA
TCCTGTTTATGAAGATCTATTAAGTGAAAATAGGAAGATGATCACTAATGAGAAGATAGAT
GCTTACAATGAAGCTGCAGTCAGTATTTTGAATAGTAGCACCAGAAATTCTAAATCAAATG
TTAAGATGTTCAGTGTTTCCAAATTAATTGCTCAAGAAACCATCATGGAATCTTTGGATGG
CTTACATCTTCCTGAATCGAGCAGAGAAACTACTGCAATGATTCTTATGAATGTGTATTGC
AATAAGATTTTGAAGCCTGTAGATGGGTCCTGTTGTCAACCTCGGCCTCCTGTTACTCTCA
TACAGAAGCTAGCTGCTTGTTTTTTCACTTTATCTATTATCGGATATTTAATTTTTTACAT
AATTCATCGTAATGCTCATCGGAAGAATAAGCCGTGTACTGATTTGGAAAGTGGAGAGGAA
AAGAAAAATATTATCAATACCCCTGTGTCTTCATTAGAAATACTTTTACAATCTTTCTGCA
AACTTGGCCTGATTATGGCATATTTCTATATGTGTGACCGTGCAAATCTGTTCATGAAGGA
AAACAAATTTTATACACATTCATCTTTCTTTATTGCAATTATCTACATTTTGGTTTTGGGA
GTATTTTATAATGAAAATACTAAAGAGACTAAAGTATTAAATAGAGAACAAACAGACGAAT
GGAAAGGCTGGATGCAACTTGTGATTTTGATTTATCACATTTCTGGAGCAAGTACATTTTT
GCCTGTATACATGCACATTCGAGTTCTGGTTGCTGCATATTTATTTCAGACAGGGTATGGG
CATTTCTCATACTTTTGGATAAAAGGAGATTTTGGAATCTATAGAGTATGTCAGGTTTTAT
TTCGTCTCAATTTCCTGGTAGTGGTGTTATGTATAGTAATGGATCGACCTTATCAATTCTA
TTACTTTGTCCCCTTGGTCACTGTATGGTTCATGGTCATATATGTTACTTTAGCACTATGG
CCACAAATAATCCAAAAAAAAGCAAACGGAAATTGTTTCTGGCATTTTGGCTTACTGTTGA
AACTAGGCTTTTTGCTGTTATTCATATGTTTTTTGGCATATTCTCAGGGTGCATTTGAGAA

GATCTTTTCTCTTTGGCCATTGTCCAAGTGTTTTGAACTGAAAGGGAATGTATATGAATGG
TGGTTCAGATGGAGGTTAGACCGTTATGTAGTTTTCCACGGAATGCTGTTTGCTTTTATTT
ATCTGGCTTTGCAGAAGCGTCAAATACTTTCTGAAGGAAAGGGTGAACCTCTTTTTTCAAA
CAAAATTTCAAATTTTCTGTTGTTTATTTCAGTAGTTTCTTTCTTGACCTATTCCATCTGG
GCTAGCAGTTGTAAAAACAAAGCAGAGTGCAATGAACTCCATCCGTCTGTTTCTGTGGTAC
AGATTTTAGCCTTCATCCTAATAAGAAACATCCCTGGATATGCCCGTTCAGTTTACAGTTC
ATTTTTTGCTTGGTTTGGAAAAATTTCATTAGAGCTATTTATTTGCCAGTATCACATATGG
CTGGCAGCGGACACAAGGGGTATCTTGGTACTGATACCTGGAAACCCTATGCTCAACATCA
TTGTCAGCACTTTCATATTTGTTTGTGTGGCACATGAAATTTCTCAGATCACTAATGATCT
TGCACAGATTATTATTCCTAAAGATAACTCATCTCTCTTGAAAAGGTTGGCATGTATAGCT
GCATTTTTTTGTGGACTCCTCATCTTATCATCCATTCAAGATAAATCAAAACATTAGGTTC
CAAAAATTCTAAAAAACCTAAACTCTTCAGGCTACCTTTGTGTGTCTCTAGAAGAGAAAAG
CATCTATCTGGAGATATAAATGTGTWTGTAAATATAAACGTTTGTGGCAAGAGGACAGTTC
TGTGACATCTGTTGAACATATGTGGTTGTATATATTGGAAATGTACATATCCAATATGAAA
TACTAAAACAAACAAACAAACAAAAAACCAGAATGCATTGTATAGGATTGCATGTGAAGTC
TTTTCTACTGAATCTATATTTCCATTTGTAAGTGATTTTAAGTTAACATATGAAGGCAGGG
AAATGATTACCTTTCCAGTAAAAAGTATAGATAATTTAATTAACTTAGTGACACCA:CCAA
GTGTTTTGATATAACTAAATTTGTGGTAATAAGACTGTCTGCACCTGTATTCATTGTGGAA
CTTCCTCTTTCATTGGAAACTTTCTTACTCAAGAATGACGGCAGTATTGTTTTCTTATATG
TGCAATGAAGTGGAATGATAAACAGTATGCCTTTAATTTATATGTGTTCTTGTTCTGATGT
TGTTTCCTGAAATGATTTTTCTTCCTAACTGTGGTTTTCGGGTATGCAAGCCTAAATCTTT
GTACACTTTGTCTCACAGAATAGTTCTGAGGCTCCATGACAGGGTTTTGTCATTGTTGATG
TTATTGTTGCTTCGTTTTATAAAAAAGCCAAAATTTTTTTTCCAATCCAAACGTTCACCTG
TTTCCTTTCCTCAAGCTATACCAGTGTAATACCAGTTACCCTGTGGATCCATTTAATATGT
TATCCCCACTAATTAATTTTCGTATATTATTTCCAATATTTGGAAAGCTCTTTATAGCCAT
TTGGTATTTCCTATTACCCACCTCCTATTTTAAATATTTATCAGTCTAAACTTGTGCAGTG
TAGTAAACATGCAAGTTGTTACGATTGAGCTGTATTACCATAAGTAGAATTTTAAGTAAAC
TGGTGAATTTGGGCAATAAATGTTTTTGCTTTTTGTTTGATTTTTTTTTACAAGCTAACTG
TTAGAGGTATACATTTATTTATCTGTTGTACAGATTTGATTATGATTTTAATGTTTGAAAG
ATTGCACTTGTTTGCTTTTACTATATGTGGGGTAAAATATATTTTCTGTTCACAGTATATG
AAAATATGGAGTAATTTAAACAGTAAATAAACATTCTGTGGATGCTTATTTTTGTATTGGC
AAAGTATCAATTAAACTATATGTGTTCTTTTTCAA

FIGURA 61A

GAAAGACATACACACTTCATGTAATGCTACCTGCAAGTCTCCCTAGAAAAGCAGTTTTTGT
AGGTGAAAACAATGAAGCCAGGTAATATTGCAAGGAGGCTGTAATTTTAGCAGACCTACCA
ACAACACTGATGTAGGAAGCTCATTATTTTAATTTCTGGAGCCTTTTAATTTTTTCTTTAG
AAAGTGTATAAATAATTGCAGTGCTGCTTTGCTTCCAAAACTGGGCAGTGAGTTCAACAAC
AACGACAACAACAGCCGCAGCTCATCCTGGCCGTCATGGAGTTTCTTGAAAGAACGTATCT
TGTGAATGATAAAGCTGCCAAGATGTATGCTTTCACACTAGAAAGAAGGAGCTGCAAATGA
ACACTTCATAGCAATGTGGAACTCCAACAGAAACCGGTGAATAAAGATCAGTGTCCCAGAG
AGAGACCAGAGGAGCTGGAGTCAGGAGGCATGTACCACTGCCACAGTGGCTCCAAGCCCAC
AGAAAAGGGGGCGAATGAGTACGCCTATGCCAAGTGGAAACTCTGTTCTGCTTCAGCAATA
TGCTTCATTTTCATGATTGCAGAGGTCGTGGGTGGGCACATTGCTGGGAGTCTTGCTGTTG
TCACAGATGCTGCCCACCTCTTAATTGACCTGACCAGTTTCCTGCTCAGTCTCTTCTCCCT
GTGGTTGTCATCGAAGCCTCCCTCTAAGCGGCTGACATTTGGATGGCACCGAGCAGAGATC
CTTGGTGCCCTGCTCTCCATCCTGTGCATCTGGGTGGTGACTGGCGTGCTAGTGTACCTGG
CATGTGAGCGCCTGCTGTATCCTGATTACCAGATCCAGGCGACTGTGATGATCATCGTTTC
CAGCTGCGCAGTGGCGGCCAACATTGTACTAACTGTGGTTTTGCACCAGAGATGCCTTGGC
CACAATCACAAGGAAGTACAAGCCAATGCCAGCGTCAGAGCTGCTTTTGTGCATGCCCTTG
GAGATCTATTTCAGAGTATCAGTGTGCTAATTAGTGCACTTATTATCTACTTTAAGCCAGA
GTATAAAATAGCCGACCCAATCTGCACATTCATCTTTTCCATCCTGGTCTTGGCCAGCACC
ATCACTATCTTAAAGGACTTCTCCATCTTACTCATGGAAGGTGTGCCAAAGAGCCTGAATT
ACAGTGGTGTGAAAGAGCTTATTTTAGCAGTCGACGGGGTGCTGTCTGTGCACAGCCTGCA
CATCTGGTCTCTAACAATGAATCAAGTAATTCTCTCAGCTCATGTTGCTACAGCAGCCAGC
CGGGACAGCCAAGTGGTTCGGAGAGAAATTGCTAAAGCCCTTAGCAAAAGCTTTACGATGC
ACTCACTCACCATTCAGATGGAATCTCCAGTTGACCAGGACCCCGACTGCCTTTTCTGTGA
AGACCCCTGTGACTAGCTCAGTCACACCGTCAGTTTCCCAAATTTGACAGGCCACCTTCAA
ACATGCTGCTATGCAGTTTCTGCATCATAGAAAATAAGGAACCAAAGGAAGAAATTCATGT
CATGGTGCAATGCACATTTTATCTATTTATTTAGTTCCATTCACCATGAAGGAAGAGGCAC
TGAGATCCATCAATCAATTGGATTATATACTGATCAGTAGCTGTGTTCAATTGCAGGAATG
TGTATATAGATTATTCCTGAGTGGAGCCGAAGTAACAGCTGTTTGTAACTATCGGCAATAC
CAAATTCATCTCCCTTCCAATAATGCATCTTGAGAACACATAGGTAAATTTGAACTCAGGA
AAGTCTTACTAGAAATCAGTGGAAGGGACAAATAGTCACAAAATTTTACCAAAACATTAGA
AACAAAAAATAAGGAGAGCCAAGTCAGGAATAAAAGTGACTCTGTATGCTAACGCCACATT
AGAACTTGGTTCTCTCACCAAGCTGTAATGTGATTTTTTTTCTACTCTGAATTGGAAATA
TGTATGAATATACAGAGAAGTGCTTACAACTAATTTTTATTTACTTGTCACATTTTGGCAA
TAAATCCCTCTTATTTCTAAATTCTAACTTGTTTATTTCAAAACTTTATATAATCACTGTT

TTGTAGAGACCCAGAAGTCTTCAAGAACAGCCGACAAAAACATTCGAGTTGACCCCACCAA
GTTGTTGCCACAGATAATTTAGATATTTACCTGCAAGAAGGAATAAAGCAGATGCAACCAA
TTCATTCAGTCCACGAGCATGATGTGAGCACTGCTTTGTGCTAGACATTGGGCTTAGCATT
GAAACTATAAAGAGGAATCAGACGCAGCAAGTGCTTCTGTGTTCTGGTAGCAACTCAACAC
TATCTGTGGAGAGTAAACTGAAGATGTGCAGGCCAACATTCTGGAAATCCTATGTCACTGG
GTTTGGTTTGGAAACCTGGACTTCTGCATTTTTAAAAGTTACCCAGAGATGCTTCTAAAGA
TGAGCCATAGTCTAGAAGATTGTCAACCACAGGAGTTCATTGAGTGGGACAGCTAGAYACA
TACATTGGCAGCTACAATAGTATCATGAATTGCAATGATGTAGTGGGGTATAAAAGGAAAG
CGATGGATATTGCCGGATGGGCATGGCCAGTGATGTTTCACGTCATTGAGGTGACAGCTCT
GCTGGACTTTGAATTACATATGGAGGCTCTCCAGGAAGACGAAGAAGAGAAGGACATTCTA
GGCAAAAAGAAGACTAGGCACAAGGCACACTTATGTTTGTCTGTTAGCTTTTAGTTGAAAA
AGCAAAATACATGATGCAAAGAAACCTCTCCACGCTGTGATTTTTAAAACTACATACTTTT
TGCAACTTTATGGTTATGAGTATTGTAGAGAACAGGAGATAGGTCTTAGATGATTTTTATG
TTGTTGTCAGACTCTAGCAAGGTACTAGAAACCTAGCAGGCATTAATAATTGTTGAGGCAA
TGACTCTGAGGCTATATCTGGGCCTTGTCATTATTTATCATTTATATTTGTATTTTTTTCT
GAAATTTGAGGGCCAAGAAAACATTGACTTTGACTGAGGAGGTCACATCTGTGCCATCTCT
GCAAATCAATCAGCACCACTGAAATAACTACTTAGCATTCTGCTGAGCTTTCCCTGCTCAG
TAGAGACAAATATACTCATCCCCCACCTCAGTGAGCTTGTTTAGGCAACCAGGATTAGAGC
TGCTCAGGTTCCCAACGTCTCCTGCCACATCGGGTTCTCAAAATGGAAAGAATGGTTTATG
CCAAATCACTTTTCCTGTCTGAAGGACCACTGAATGGTTTTGTTTTTCCATATTTTGCATA
GGACGCCCTAAAGACTAGGTGACTTGGCAAACACACAAGTGTTAGTATAATTCTTTGCTTC
TGCTTCTTTTTGAAAATCATGTTTAGATTTGATTTTAAGTCAGAAATTCACTGAATGTCAG
GTAATCATTATGGAGGGAGATTTGTGTGTCAACCAAAGTAATTGTCCCATGGCCCCAGGGT
ATTTCTGTTGTTTCCCTGAAATTCTGCTTTTTTAGTCAGCTAGATTGAAAACTCTGAACAG
TAGATGTTTATATGGCAAAATGCAAGACAATCTATAAGGGAGATTTTAAGGATTTTGAGAT
GAAAAAACAGATGCTACTCAGGGGCTTTATGGACCATCCATCAATTCTGAAGTTCTGACTC
TCCCATTACCCTTTCCCTGGTGTGGTCAGAACTCCAGGTCACTGGAAGTTAGTGGAATCAT
GTAGTTGAATTCTTTACTTCAAGACATTGTATTCTCTCCAGCTATCAAAACATTAATGATC
TTTTATGTCTTTTTTTTGTTATTGTTATACTTTAAGTTCTGGGGTACATGTGCGGAACATG
TAGGTTTGTTACATAGGTATACATGTGCCATGGTGGTTTGCTGCACTCATCAACCTGTCAT
CTACATTCTTTTATGTCTGTCTTTCAAAGCAACACTCTGTTCTTCTGAGTAGTGAAATCAG
GTCAACTTTACCACCAGCCTCCATTTTTAATATGCTTCACCATCATCCAGCACCTACTTAA
GATTTATCTAGGGCTCTGTGGTGATGTTAGGACCCATAAAAGAAATTTATGCCTTCCATAT
GTTTGGTTACAGATGGGAAATGGGAATGTTGAAGGACATGAAAGAAAGGATGTTTACACAT
TAAGCATCAGTTCTGAAGCTAGATTGTCTGAGTTTGAATCTTAGCTCTTCCCTTTATTAGC
TCTGTGACCTCGAGCTAGTTACTTAAATGCTCTGATCCTCTATTTCCTGATCAGTGAAACC

ATAGCATGTAGTGTTCAGTACATGTTAAATGTTGTTTTTTATTATGTACAAACATGTGTGG
GCACAGAATTTTAAATCATCTCAACTTTTGAGAAATTTTGAGTTATCAACACCGTTCCCAC

FIGURA 61B

AAGACAGTGGCAAAATTATTGGTGAGAATTAAACAGCTGTTTCTCAGAGGAAGCAATGGAG
GCTTGCTGGGATAAAGGCATTTACTGAGAGGCTGTTACCTAGTGAGAGTGATGAATTAATT
AAAATAGTCGAATCCCTTTCTGACTGTCTCTGAAAGCTTCCGCTTTTATCTTTGAAGAGCA
GAATTGTCACCCCAAGGACATTTATTAATAAAAAGAACAACTGTCCAGTGCAATGAAGGCA
AAGTCATAGGTCTCCCAAGTCTTACCCCATTCCTGTGAAATATCAAGTTCTTGGCTTTTCT
CTGTCATGTAGCCTCAACTTTCTCCGACCGGGTGCATTTCTTTCTCTGGTTTCTAAATTGC
CAGTGGCAAATTTGGATCACTTACTTAATATCTGTTAAATTTTGTGACCCAACAAAGTCTT
TTAGCACTGTGGTGTCAAAAAGAAAAACACCTCCCAGGCATATACATTTTATAGATTCCTG
GAGAATGTTGCTCTCCAGCTCCATCCCCACCCAATGAAATATGATCCAGAGAGTCTTGCAA
AGAGACAAGCCTCATTTTCCACAATTAGCTCTAAAGTGCCTCCAGGAAATGATTTTCTCAG
CTCATCTCTCTGTATTCCCTGTTTTGGATCACAGGGCAATCTGTTTAAATGACTAATTACA
GAAATCATTAAAGGCACCAAGCAAATGTCATCTCTGAATACACACATCCCAAGCTTTACAA
ATCCTGCCTGGCTTGACAGTGATGAGGCCACTTAACAGTCCAGCGCAGGCGGATGTTAAAA
AAAATAAAAAGGTGACCATCTGCGGTTTAGTTTTTTAACTTTCTGATTTCACACTTAACGT
CTGTCATTCTGTTACTGGGCACCTGTTTAAATTCTATTTTAAAATGTTAATGAGTGTTGTT
TAAAATAAAATCAGGAAAGAGAGAAAAAAAAAAAAAAAAAAA

FIGURA 62A

GCGGCGGGAGAGCTGCTGGCTCGCCCGGATCCCGGGAGCTGCCTGGAGGCGGGCCCGGCCC
GGGGAAGGTGAGCGGCTGCGGGACCCAGCCCCTCGCCGGGAGCGGGCACCATGGTGCTGTC
GGTGCCTGTGATCGCGCTGGGCGCCACGCTGGGCACAGCCACCAGCATCCTCGCGTTGTGC
GGGGTCACCTGCCTGTGTCGGCACATGCACCCCAAGAAGGGGCTGCTGCCGCGGGACCAGG
ACCCCGACCTGGAGAAGGCGAAGCCCAGCTTGCTCGGGTCTGCACAACAGTTCAATGTTAA
AAAGTCCACGGAACCTGTTCAGCCCCGTGCCCTCCTCAAGTTCCCAGACATCTATGGACCC
AGGCCAGCTGTGACGGCTCCAGAGGTCATCAACTATGCAGACTATTCACTGAGGTCTACGG
AGGAGCCCACTGCACCTGCCAGCCCCCAACCCCCGAATGACAGTCGCCTCAAGAGGCAGGT
CACAGAGGAGCTGTTCATCCTCCCTCAGAATGGTGTGGTGGAGGATGTCTGTGTCATGGAG
ACCTGGAACCCAGAGAAGGCTGCCAGTTGGAACCAGGCCCCCAAACTCCACTACTGCCTGG
ACTATGACTGTCAGAAGGCAGAATTGTTTGTGACTCGCCTGGAAGCTGTGACCAGCAACCA
CGACGGAGGCTGTGACTGCTACGTCCAAGGGAGTGTGGCCAATAGGACCGGCTCTGTGGAG
GCTCAGACAGCCCTAAAGAAGCGGCAGCTGCACACCACCTGGGAGGAGGGCCTGGTGCTCC
CCCTGGCGGAGGAGGAGCTCCCCACAGCCACCCTGACGCTGACCTTGAGGACCTGCGACCG
CTTCTCCCGTCACAGCGTGGCCGGGGAGCTCCGCCTGGGCCTGGACGGGACATCTGTGCCT
CTAGGGGCTGCCCAGTGGGGCGAGCTGAAGACTTCAGCGAAGGAGCCATCTGCAGGAGCTG
GAGAGGTCCTACTATCCATCAGCTACCTCCCGGCTGCCAACCGCCTCCTGGTGGTGCTGAT
TAAAGCCAAGAACCTCCACTCTAACCAGTCCAAGGAGCTCCTGGGGAAGGATGTCTCTGTC
AAGGTGACCTTGAAGCACCAGGCTCGGAAGCTGAAGAAGAAGCAGACTAAACGAGCTAAGC
ACAAGATCAACCCCGTGTGGAACGAGATGATCATGTTTGAGCTGCCTGACGACCTGCTGCA
GGCCTCCAGTGTGGAGCTGGAAGTGCTGGGCCAGGACGATTCAGGGCAGAGCTGTGCGCTT
GGCCACTGCAGCCTGGGCCTGCACACCTCGGGCTCTGAGCGCAGCCACTGGGAGGAGATGC
TCAAAAACCCTCGCCGGCAGATTGCCATGTGGCACCAGCTGCACCTGTAACCAGCTGCCCA
GCTGCCTCCCTTCTTGGACAGCCCTGACCCGTCCTCTGCAACCTCCTTTCTGTGCCCCTTT
CCTCATTCTGACACCCAGAAGACAGTGACAGATGTGTTTGCAAGGCTGGGATGGCTCTCTC
ATCATACTCTTGTTTCTTAGAAATAAGCAAGACAGAGCAGGAAATGGAATATGCGGGTCAC
ACTGAGGAATGCATTTTGCTCATCTGTGTTATTGAAGGAGGTGCTTATTAAATACAGTTCC
TATGCCTGTTTTATAGGTGGGGTTAGGCCAGATGCAGAGAAAGCTAAATGTGGGAATCATG
GATGCAAAGAAGAATTTGGCTTTTTGAAAAACAAGCATTTCAAAAATGATGAAGGAAGTGA
AAGTATCCTGGATCAACTCCTAGAGTTAGAGATTGCCCAGGTGGAAAGAAACCTTAGCCAG
CGTTCAATCAAGCTCACCATGCAGGGCAGTCACCCGGCAGTTCTCAAACTTTAGCATGTGA
AGAGTCACCAGCAGATTCCTGGGCTCGCCTGGAGACATTCCTAGTCGGTATTCCTGGTCGA
AGCCCAGGAGCCTTCCTTTTTAACAAGCTGATGTAGAGGGTGGAGCACTGTATGTGGAGAA
ATTCCTTCTACAATATTCCACACAGGTTTTTGGCCACAGTCCTTGATGGAGTCCCAAAACC
[illegible]

CTTTTTAATTTAATTATTTTAATTCTTAATACTTTAATTTGTGCATTTCATAAGCCCCCTG
CTCTTGGACTGAATTTTGTGCTTTTTATTGAAGAATTTTATTGTTTTTATCTTAAAATCAG
TTTCTATTATCCTTGGGGAGACCATCCCTAACAAAGTACAGGTGGGATCTCCTGTGAGTCA
TTGGCTGGGTTCTGATTGCTAGATGTCACACCCACCAGCATCACCAAAGTGACTCTGAGAT
AGACCGGTCCCTTCTCAGCGTTCCAGTCACTTCAGGAGGAATTTAGTTATTGACTTAGTCT
ATGACATCTGGCTACATGTAGGTAGAGAAGAAAGACAATTTTAAAAAGGAAATCAGGTCTT
TTGCAACTGTGCCTCCCTCTGTCTGTTTTCACTTGAATGGGTAAATAACCAGCAGCTAGGT
TTTGAATTCCTACCTTGTTATTCTAAACAGATGTCCACATTGTTAATTAAATCTAAATTAT
GAGCCTTGCTGAGTGGATACGGTACTTACACCTGAACCAGGATTCCTGGGTTCTGTTGTTG
ACATTGCCCTTCAGCACCTGTTTGGCCAGCTGTATAAGATAGGACTAATGACTAGGAAGCC
TACCCCAATGAATGATATACTAGATGAAATAGTGTTCAAAACCTGTAGGCACTCTCTGGCT
AAAAACAAACTCTGAGGCCACCAGCAGATCATCTTTAAGCTAAGTTACTATTTTTCACCTT
TTTTTTTAGACGGAGTTTTGCTCTTTGTTGCCCAGGCTGGAGTGCAGTGGCACGATCTCGG
CTCACTGCAACCTCCGCCTCCCAAGTTCAAGCGATTCTCCTGTCTCAGCCTCCTGGGTAGC
TGGGATTACAGGTGCCCACCAACATGCCTGGCTAATTTTTGTACTTTTAGTAGAGATGGGG
TTTCACCATGTTGGCCAGGCTGGTCTTCAACTCCAGATCTCAGGTGATCTACCCTCCTCGG
CCTCCCAAAGTACTGGGATTACAGGCCTGAGCCACCGCGCCCGGCCTATTTTTCACTTTAA
TTTGGCAGCTGAGAATGCCCAAAAAGTGCCAGAAGCATCGTGGCATTTCCAGAACCATGGA
TTCTGCCTTTGGACCCCTCTCTATTAATATTAAAACTCTGGGCCTTCAGATGTCACCCTAA
TCCACTGCCCTAAGACAGAATTTCTGGACAAGATGGGTAAGGGCTTCATTCCTTCAACAAG
TCAAGTCATACTTGGCCTCTCCCTGAGAATCTGAGCAGGAGCCTTATAACCTGTGGTCATT
ATTTTTTCTTTCTGTACAGAAATAGAAAAGCATTAGAAATAACTTCTAACCATCCTCTGAA
AAAACAGAAAAAATATCGAATCCCTCTTTCATGAGAAGTCTTTTGGATAATTGGAAACCTT
CATCACTGAGGTTGGCCAGCCCCTGCCAAGTGTTGTGTAGGCAAAGCACTTGTTAGTGGCT
TCCTATGAAATGTTTTAGAGATCTCTTCACCATACTGGTTTCTTCTCTTTGGTTGGTGTGG
GTAAAAGAAAACAAAACATTTCCTATAAGCTGAAAGCTGACCAGCATTCTCTTCTTGGTAA
CATCTACTACTCCAACCTAGAAAATTTGGATTCTAGACCAAAAATCAGGAAACATGGCTCC
TTATAAATCTGTGCAGCTGCCTTATAGTACCATCAAAGGAATTTCAGGTGGGCTGGGCGGG
GCCCCGATCCCAGAATTATCAACTCCACCCATCATCATTTGGTCATGAAGCATCCTTTCAT
TCTTCTTCTTCTTTTTTTTGGGGGGGGCGGGGCGGGGGAGGGATCTCAAAGTTTTAGTCTT
CCAGAATCCAAATTAAAGGTTGCCCCTGATGGGGGCCAGGTTCCGCCACAGAACATGTTAG
ATGTCAGCCTTGACCTCACTTAGCAGGGATTACAGAAATGAGATACATTTTGAAGGAGAGT
TGTCTGTTATGTTCACTGTATTCTAAGTGCCTGGGATAAAGCTGTCTCATGGGTGCTCCAT
ATATATTCATATATATTTGTTGAGTGAATTAATGAATTAAGAGTGGCTGGCAGAGTAGGCA
GAAAAAGACACTGCAAATGGCATAAAAATTAAAGTCCTAGCTGAGTTCTCAATGGTAAAGG
CATCAGATGTCTTAGCAGTCAAGCTAGAAATTCATGACAATGAGTATTACTATTTGCCTAA

CTTTTCCACTAAAACTCCATCTTAGTGAATTTTAAATTATCCAGAGATGTCAAACTGCCAA
ATAAAAATATTTCAGTAGTCTTTGCATCAGCTTACCTTGTACCAGAAACATTTCCAATTTA

FIGURA 62B

CTATCAAATTATAGTAACTGAGCCTGTGTGAAGTATCTCATCATTTTCGAAAGGAACACCT
TGTGTGATGCCAGTGAGCATTTCTAAAAAGGGTGTGAGGTAGAGGTAAAAATAAGGTGAGA
GACCATTTCAGAATGCACTGTTGCTCAAAAAGGTGATCTGGTTCTTTCTTCAGAGATTTCT
ACGGGGATAGAAAATCGGGAGTCTGCCCTCATTAATCTGTGACTCCACCTCTTGCATCAAA
TCAATATCTATTTGTTGAGCACTTATTGATTAAGACCTTGCATATGTCTGTCCATTTTGAT
TTGAGATACAACTTTTTGTGTGGGTTGAATGACAAATCACTCCAAACAAAACTGGGCACAG
AGAATCAGCTAGGAGACCAGTTATTCAGGGTCCATTTCTCTTGGATGTAAAGGAGTCCTGG
GTAAAATGTGGCTGTAACCTAAACCAACTAGTCCTTGTGATTTGTTTCTGCCCTCTGTGTT
TCCTGTTGTCAAATGCTAAGTGTGTGTTTTGCAGTCATGAACTAAAGCACAAAAAGATGCA
TGAGACATTGTAGTCATATGTCTGGTGTGACACTTTGGAGCAAAAACCTTGCAGTGGTAAA
TAAAAAATTTCCAACAGGGAAAAAAAAAAAAAAAAAAAAAAAAA

911

FIGURA 63A

AGCGGCAGCTCCGCTGAAAACTGCATTCAGCCAGTCCTCCGGACTTCTGGAGCGGGGACAG
GGCGCAGGGCATCAGCAGCCACCAGCAGGACCTGGGAAATAGGGATTCTTCTGCCTCCACT
TCAGGTTTTAGCAGCTTGGTGCTAAATTGCTGTCTCAAAATGCAGAGGATCTAATTTGCAG
AGGAAAACAGCCAAAGAAGGAAGAGGAGGAAAAGGAAAAAAAAAAGGGGTATATTGTGGATG
CTCTACTTTTCTTGGAAATGCAAAAGATTATGCATATTTCTGTCCTCCTTTCTCCTGTTTT
ATGGGGACTGATTTTTGGTGTCTCTTCTAACAGCATACAGATAGGGGGGCTATTTCCTAGG
GGCGCCGATCAAGAATACAGTGCATTTCGAGTAGGGATGGTTCAGTTTTCCACTTCGGAGT
TCAGACTGACACCCCACATCGACAATTTGGAGGTGGCAAACAGCTTCGCAGTCACTAATGC
TTTCTGCTCCCAGTTTTCGAGAGGAGTCTATGCTATTTTTGGATTTTATGACAAGAAGTCT
GTAAATACCATCACATCATTTTGCGGAACACTCCACGTCTCCTTCATCACTCCCAGCTTCC
CAACAGATGGCACACATCCATTTGTCATTCAGATGAGACCCGACCTCAAAGGAGCTCTCCT
TAGCTTGATTGAATACTATCAATGGGACAAGTTTGCATACCTCTATGACAGTGACAGAGGC
TTATCAACACTGCAAGCTGTGCTGGATTCTGCTGCTGAAAAGAAATGGCAAGTGACTGCTA
TCAATGTGGGAAACATTAACAATGACAAGAAAGATGAGATGTACCGATCACTTTTTCAAGA
TCTGGAGTTAAAAAAGGAACGGCGTGTAATTCTGGACTGTGAAAGGGATAAAGTAAACGAC
ATTGTAGACCAGGTTATTACCATTGGAAAACACGTTAAAGGGTACCACTACATCATTGCAA
ATCTGGAATTTACTGATGGAGACCTATTAAAAATCCAGTTTGGAGGTGCAAATGTCTCTGG
ATTTCAGATAGTGGACTATGATGATTCGTTGGTATCTAAATTTATAGAAAGATGGTCAACA
CTGGAAGAAAAAGAATACCCTGGAGCTCACACAACAACAATTAAGTATACTTCTGCTCTGA
CCTATGATGCCGTTCAAGTGATGACTGAAGCCTTCCGCAACCTAAGGAAGCAAAGAATTGA
AATCTCCCGAAGGGGGAATGCAGGAGACTGTCTGGCAAACCCAGCAGTGCCCTGGGGACAA
GGTGTAGAAATAGAAAGGGCCCTCAAACAGGTTCAGGTTGAAGGTCTCTCAGGAAATATAA
AGTTTGACCAGAATGGAAAAAGAATAAACTATACAATTAACATCATGGAGCTCAAAACTAA
TGGGCCCCGGAAGATTGGCTACTGGAGTGAAGTGGACAAAATGGTTGTTACCCTTACTGAG
CTCCCTTCTGGAAATGACACCTCTGGGCTTGAGAATAAGACTGTTGTTGTCACCACAATTT
TGGAATCTCCGTATGTTATGATGAAGAAAAATCATGAAATGCTTGAAGGCAATGAGCGCTA
TGAGGGCTACTGTGTTGACCTGGCTGCAGAAATCGCCAAACATTGTGGGTTCAAGTACAAG
TTGACAATTGTTGGTGATGGCAAGTATGGGGCCAGGGATGCAGACACGAAAATTTGGAATG
GGATGGTTGGAGAACTTGTATATGGGAAAGCTGATATTGCAATTGCTCCATTAACTATTAC
CCTTGTGAGAGAAGAGGTGATTGACTTCTCAAAGCCCTTCATGAGCCTCGGGATATCTATC
ATGATCAAGAAGCCTCAGAAGTCCAAACCAGGAGTGTTTTCCTTTCTTGATCCTTTAGCCT
ATGAGATCTGGATGTGCATTGTTTTTGCCTACATTGGGGTCAGTGTAGTTTTATTCCTGGT
CAGCAGATTTAGCCCCTACGAGTGGCACACTGAGGAGTTTGAAGATGGAAGAGAAACACAA
AGTAGTGAATCAACTAATGAATTTGGGATTTTTAATAGTCTCTGGTTTTCCTTGGGTGCCT
[illegible]TCCCTCTCTGGGCGCATTGTTGGAGGTGT

GTGGTGGTTCTTTACCCTGATCATAATCTCCTCCTACACGGCTAACTTAGCTGCCTTCCTG
ACTGTAGAGAGGATGGTGTCTCCCATCGAAAGTGCTGAGGATCTTTCTAAGCAAACAGAAA
TTGCTTATGGAACATTAGACTCTGGCTCCACTAAAGAGTTTTTCAGGAGATCTAAAATTGC
AGTGTTTGATAAAATGTGGACCTACATGCGGAGTGCGGAGCCCTCTGTGTTTGTGAGGACT
ACGGCCGAAGGGGTGGCTAGAGTGCGGAAGTCCAAAGGGAAATATGCCTACTTGTTGGAGT
CCACGATGAACGAGTACATTGAGCAAAGGAAGCCTTGCGACACCATGAAAGTTGGTGGAAA
CCTGGATTCCAAAGGCTATGGCATCGCAACACCTAAAGGATCCTCATTAGGAACCCCAGTA
AATCTTGCAGTATTGAAACTCAGTGAGCAAGGCGTCTTAGACAAGCTGAAAAACAAATGGT
GGTACGATAAAGGTGAATGTGGAGCCAAGGACTCTGGAAGTAAGGAAAAGACCAGTGCCCT
CAGTCTGAGCAACGTTGCTGGAGTATTCTACATCCTTGTCGGGGGCCTTGGTTTGGCAATG
CTGGTGGCTTTGATTGAGTTCTGTTACAAGTCAAGGGCCGAGGCGAAACGAATGAAGGTGG
CAAAGAATGCACAGAATATTAACCCATCTTCCTCGCAGAATTCACAGAATTTTGCAACTTA
TAAGGAAGGTTACAACGTATATGGCATCGAAAGTGTTAAAATTTAGGGGATGACCTTGAAT
GATGCCATGAGGAACAAGGCAAGGCTGTCAATTACAGGAAGTACTGGAGAAAATGGACGTG
TTATGACTCCAGAATTTCCCAAAGCAGTGCATGCTGTCCCTTACGTGAGTCCTGGCATGGG
AATGAATGTCAGTGTGACTGATCTCTCGTGATTGATAAGAACCTTTTGAGTGCCTTACACA
ATGGTTTTCTTGTGTGTTTATTGTCAAAGTGGTGAGAGGCATCCAGTATCTTGAAGACTTT
TCTTTCAGCCAAGAATTCTTAAATATGTGGAGTTCATCTTGAATTGTAAGGAATGATTAAT
TAAAACACAACATCTTTTTCTACTCGAGTTACAGACAAAGCGTGGTGGACATGCACAGCTA
ACATGGAAGTACTATAATTTACCTGAAGTCTTTGTACAGACAACAAACCTGTTTCTGCAGC
CACTATTGTTAGTCTCTTGATTCATAATGACTTAAGCACACTTGACATCAACTGCATCAAG
ATGTGACATGTTTTATAAAAAAAGGAAAAAAAACATTTAAAACTAAAAAATATTTTTAGGT
ATTTTCACAAACAAACTGGCTTTTAAATAAATTTGCTTCCATATTGGTTGAATAAGACAAA
AACAATTAAACTGAGTGGGAAGTGAATAAAAAAAGGCTTTAGGTATCGATTCCATATTTTT
CAAAGCCAAATATGTAAATGCTAAGGAAAGTAAACAAAGAGGAGATTCCAATCTTGTAATT
TAATATTGTTATTAAAACTTTAATGTATCCTATTCTTTAACATTTGGTGTTAATATAAAAT
TACTTGGCAATGCTTGACATTTGAAATAAACATTTTTCTATTGTTTTATTGCAAGTGGTCC
AATTAATTTTGCTTAGCTACAGTTTGGTCATAAATCAAGTGAGTTTAAAGACACTACCAAG
TTGTTAGGTGCCCAGAGAAAATTTCTCCCTTTTAAAAAGGCCAGGTGATTTTTCAAATGTA
ATCTTGCCCCCAAAGTAATATCCGAATATCTTTTTGACATGTCTAAATATATATATATATA
AAGAAATATTTGTTAACACAAAAGCATTTGATCTATGTAGATAAATGCTAATAGATTTAAA
AAGCTAATATTAACAAATACCAGAATACGTGAAGTTCCATTTTTAAAGTGTTTGAGCTTAC
AGAAGAGAAACATTCATTTTAAATGAAGTAAAAAATGCCTTGAAAGTAATTCTTTAGATAG
TTGCCCATTGATTAAATTCCAAAAACTAAATATGTTTTTAGCTTTAAAATTATAAAAGCTG
TCATAAACTTTATATATTATGAATTTTAAAATATGTTTGAGTCTCCTGCAATATAGTTTCA
TCCCATTGACATCAATTAAAAATAACCCTAATATATTATTTTTATATTTATTCCTCAGGTG

AAGAGCCTATATTTATCTAGTTAATGAATTTAAAGGATCTATCTTTCCCTTCATAAAATAC
CTCTTATTTCCATTAAAGCCCCCCAAGTTTAATTAATTTAGGATTTTGAATGATTATTGAC

FIGURA 63B

ATCCAATAGTTATTTTTAATATTTGTATTCTTGTTATTTCTGGAAGAAAGCCTTTGTGTAG
CACTTGGTATTTTGCAAAGTGCTTTTAAAACATTCTTACTTACCGTATTTCATAGAAGGGA
AGGAAAAATGTAAGGTTTAACAGTAAGCACTTGCATTGAACATGGAGGCATGTGGTATCAT
GATATTCTTCACTAAATTTAGCTGTCCCTAATCACAGATCCTAAGGTAATATAATATAATT
TTAGTGCATTTCTCCTCATCAGGAATGCTGGAGGTGCATTTTAAGTTTTAATAATAAGTGC
TAGAATGACCAAATTGCAGACTAATTGTTTCCATATTGTACTTAAAATGAGTTTTTAAAAG
TGAAAAAGAAATGACTATATACAATCAATGCTATTTATTGTACCTCTGGGCCTACTCTTCT
AAAAATTGTAGCTTATCGATTTTTCTCTGTCAAGCTTGAACTAATGTAAATAATTGAAATA
ATGTAAAGTTATATTTTCATGTTTTTATAGATACAACATGACAAGAATACATAATGTAAGA
GTATTTCAACTATGGATAATGTTGATTGGATAATGCACATCTCAGTTACAAGCAGTACTCA
TAGTTTAATATCCATGTAACGGTGCATCAATATATTGCTATATAAATATGTCTGTGTGCAT
ATAAGTGAAAAGTGGTCAAACAAGAGTGATGACAGCTGTCTAAAGGTTTTTTTATTCATTT
TATATAAAAACTGTTATGGAAAGACCAAAATGTTTATGAACTATTCTTATGTAAATTTACA
ATTGTCCTTTACTGTACTTTTTTGTTTACAGTATAGTACCTTATTTTCTGCTGTGTTAAGT
GGGTGTCAAACTCCAAGAAGACATACACTTTCTATAACTTCTATTGAAGATATTGGAATTT
CCAATTTTTCATGTGTACTATGTCAGAAAATGCTTTCGATTTTATTTTTAAATCTAACATC
GGATGGCTTTTCCGGAGTGTTGTAAAAACTTCAATCATACATAAAACATGTTCTTACAAAA
GGCAAAGATCTT

FIGURA 64

GCCAACACTGGCCAAAAGGGATAGCTGTCCAAGGTCTCCCCCAGCACTGAGGAGCTCGCCT
GCTGCCCTCTTGCGCGCGGGAAGCAGCACCAAGTTCACGGCCAACGCCTTGGCACTAGGGT
CCAGAATGGCTACAACAGTCCCTGATGGTTGCCGCAATGGCCTGAAATCCAAGTACTACAG
ACTTTGTGATAAGGCTGAAGCTTGGGGCATCGTCCTAGAAACGGTGGCCACAGCCGGGGTT
GTGACCTCGGTGGCCTTCATGCTCACTCTCCCGATCCTCGTCTGCAAGGTGCAGGACTCCA
ACAGGCGAAAAATGCTGCCTACTCAGTTTCTCTTCCTCCTGGGTGTGTTGGGCATCTTTGG
CCTCACCTTCGCCTTCATCATCGGACTGGACGGGAGCACAGGGCCCACACGCTTCTTCCTC
TTTGGGATCCTCTTTTCCATCTGCTTCTCCTGCCTGCTGGCTCATGCTGTCAGTCTGACCA
AGCTCGTCCGGGGGAGGAAGCCCCTTTCCCTGTTGGTGATTCTGGGTCTGGCCGTGGGCTT
CAGCCTAGTCCAGGATGTTATCGCTATTGAATATATTGTCCTGACCATGAATAGGACCAAC
GTCAATGTCTTTTCTGAGCTTTCCGCTCCTCGTCGCAATGAAGACTTTGTCCTCCTGCTCA
CCTACGTCCTCTTCTTGATGGCGCTGACCTTCCTCATGTCCTCCTTCACCTTCTGTGGTTC
CTTCACGGGCTGGAAGAGACATGGGGCCCACATCTACCTCACGATGCTCCTCTCCATTGCC
ATCTGGGTGGCCTGGATCACCCTGCTCATGCTTCCTGACTTTGACCGCAGGTGGGATGACA
CCATCCTCAGCTCCGCCTTGGCTGCCAATGGCTGGGTGTTCCTGTTGGCTTATGTTAGTCC
CGAGTTTTGGCTGCTCACAAAGCAACGAAACCCCATGGATTATCCTGTTGAGGATGCTTTC
TGTAAACCTCAACTCGTGAAGAAGAGCTATGGTGTGGAGAACAGAGCCTACTCTCAAGAGG
AAATCACTCAAGGTTTTGAAGAGACAGGGGACACGCTCTATGCCCCCTATTCCACACATTT
TCAGCTGCAGAACCAGCCTCCCCAAAAGGAATTCTCCATCCCACGGGCCCACGCTTGGCCG
AGCCCTTACAAAGACTATGAAGTAAAGAAAGAGGGCAGCTAACTCTGTCCTGAAGAGTGGG
ACAAATGCAGCCGGGCGGCAGATCTAGCGGGAGCTCAAAGGGATGTGGGCGAAATCTTGAG
TCTTCTGAGAAAACTGTACAAGACACTACGGGAACAGTTTGCCTCCCTCCCAGCCTCAACC
ACAATTCTTCCATGCTGGGGCTGATGTGGGCTAGTAAGACTCCAGTTCTTAGAGGCGCTGT
AGTATTTTTTTTTTTTTGTCTCATCCTTTAGATACTTCTTTTAAGTGGGAGTCTCAG

FIGURA 65

TCATTCCATTAAACATATTTCTAAATAATAGTAAGTGGTACTAACAAAATAAATAATAATT
TAATAGCCTTAGAAATAAATGACTGTATACTTATACAGGTTGAAAAAAACTCGGTAGGAAT
AAGTTACCTTTTTGTTTACTAATGTTGGTTTCAAAAATACTCAGATTCATTTTAGTTGGCT
GACATCTGGAAGTAGTTAACAACTAACCAGTGGACTTCAACAATCATTTGCTCCCAGGCTT
CCCCCATCATCACCCTCACCACATATCCTGCTAATATCCAACAACAAACAAATATTTAATA
TTGAAATAGCCCATTGCCTGAGAATGAACACAAGCTAAAATACATGCAAGGGTACTTAATG
GAAGCCAAACCATGTTCTATACCTAAGGAGAAAACATGGACATGTAGAATGCTTTTATTCA
TGTATTCAAAATCAGAACAAATCAGTGTATATCACTAGAACATCAGATGGAGGATAACACA
AGAAGTGATACAGATCAGGGTTCACTTCTCTTACCCTCTCTCTGTTAGGACCACATTCCTA
TTTTAGCCAAATGTTTCTGGTACGGGCCATCTTTTCACCATAAATGGCATTATGTTTCAAA
TGGCTAAAAGCTTATTATGGGTATGCTCAAAGGAGTAAAACCCAATTCACAGAGATGTGGC
TTTTCTAAAGAACCAAATTAAGGGAATACTTAGTATCTACAACTAAAGTATTAGTCATTCT
GAGGATTATTTGTTGCTTTAAGTATTAGCTCCCTACTACATCCATAACACCCTACTTACTA
AATTTAATTACACACAACTTTTCAAGAATTCGTATTTTATTTGAAGGGAGGTACCTGTCTA
CTTTATCTACAATAAAAACAAAGGGTATTGGCTTTCTCTAATCCATGCAAACTACAAATTC
CATCGGGAGTCCTACATCACTAACAGTGGTATGAACAAAACTAAGAAAGTACTTCCTATCC
ACATGTGAATGTTTTAAAAAAGTTTTTGCCATAAAACCTAAGTGTAATTTAGCATACCACA
GTGCTCTGAAGATGGGTCATTGACGATGTACCATTTGTATATAGGTAATACACATGTAAAT
CACTAATTGTTAATTAAATATAAGAAGTAGGTTTTTATCTTTTTANNAAAAAAACAAAAAA
GGTGACTCCCTTTCTCATTCTGTTCTGTTGTATTGCGTCCAAAAGTTGTGTGTAATTTTTT
TAAGGTCCAAGAAATGTAAAGAAATTTGTTGGAATACCTGAATTTCTGTAAAAAAATAAAA
ATAAAAATAAGATTTTGTTACTT

FIGURA 66

GGGAGGCTGTGCCAGGCGAGCCGGAGGGGTGCTCCGCGCTCCCCCGCCCTCCTTCCGGGAG
CGAGGATGCAGACTCTGAAACTGGTGCTGCTGGGCTGAGGCGGAGGCAGGGGAGTTGCAGC
GCGCGAGGCTCCGTGAGTGTGTCTCCTGCGCGCTGAGAGGCGGGGGGAGGCGGAGGACCAG
GAGGAGGAGGAGGAGGAGGAGGAGGGGGAGAATGCCCGGAGCCGCCGCCGCTGCCGCCGCC
GCCGCCGCCGCGATGCTCCCGGCTCAGGAGGCTGCCAAGCTGTACCACACCAACTATGTGC
GGAACTCGCGGGCCATCGGCGTGCTGTGGGCCATCTTCACCATCTGCTTTGCCATCGTCAA
CGTGGTGTGCTTCATCCAGCCCTACTGGATAGGCGACGGCGTGGACACCCCGCAAGCCGGC
TATTTCGGGCTCTTCCACTACTGCATCGGCAACGGCTTCTCCCGGGAGCTGACCTGCAGGG
GCAGCTTCACGGACTTCTCCACGCTGCCCTCGGGCGCCTTCAAAGCCGCCTCCTTCTTTAT
CGGCCTCTCCATGATGCTCATCATTGCCTGCATCATTTGCTTTACCCTCTTCTTCTTCTGC
AACACGGCCACTGTGTACAAGATATGTGCCTGGATGCAGCTCACCTCCGCTGCCTGCCTTG
TGCTTGGCTGTATGATTTTCCCTGATGGCTGGGACTCAGATGAAGTAAAACGGATGTGTGG
AGAAAAGACAGACAAGTACACTCTTGGGGCTTGCTCAGTCCGCTGGGCATACATCCTGGCT
ATTATTGGAATTTTGGATGCCCTGATCCTCTCATTTCTAGCATTTGTGCTTGGTAATCGAC
AAGACAGCTTGATGGCAGAGGAACTGAAGGCAGAAAACAAAGTTCTGCTAAGCCAATATTC
TCTAGAATGAGCACAAAACAAATCGAATAACAGCTAAACAAATCGAATAACAGCTAAACGA
ATCGAATAACAGCTTTTGTACATCAACATCAAGAAGGAATACGCCTGAGAGAGATCAGAGT
ATATAGATGAATATGAACAAGAATGGAACATTCACTTGTCAACGCACTTTCTAAATCTAGA
TCAGCAGAGATGGGAGTGATTTTCTGGAAAGAGATGTGATCATGGATTAAACACCAGCTCA
TTGGAAACTCATTGGATGAGATCAGAAAACGTTCATGAAAAATCATATTCAGGAAATAAGG
AAGAGGAATATAAATGCTCTAGAGTTAACATGTAAAATATATACGTACTGAGGTTTGTAAA
CTGTCCTTTTTAAATCAAACTGAAAACAAAAAGCTTTAACCTTTCAACAGAATTTTTAAAA
AGGCAGTTAGTTCTAAATTATTCCTATCTCAATAGCCAAGAGGCTGATCAAGCGTCATTTA
TTGAGGAAGCATCTTAGAAAATGCCTCTGAATGTTTTCATAGGAGCCGTGACCTTTGGTTC
TTCATCTCTACCATTCATTTACTTCACTGTGTAATTAGTTACAACCACTCAGTTATTAAGA
GACGTAACGCTTCAAACTTTTTACCAAGTCTGTGTTCTGTTTAATCTGTCCATACAAGTTA
TTACTGAGAAAGTGTTTATGCCATATACTATTACTCCATCAAGCTGTATATTACAGGAAGT
ACATCTTTACATCATAGGTTCCCAAGCAACATAGATTTCCCTATCTTTCAGGAAACAGCAT
CAAGGAACTCTGAAAAATATAGAAAAAGTTCATTTTCACCTTGGAAGCTCACGTGTAATAT
TATAGGCTACTATCAAATAAACACTTTTTTTCTAATTCTCCCTAGTATATGCATAGGAATT
TAATATACTTTATAAATAAGTATCTAAAATGTCTCCTACTTTTTTCCTATTTCTTTGCCAT
ACATGTTATCAGAAATCCATGTCTTCTATTTCCCTTACTGATGGGCACTCATTTTTATTTT
TTTAAAAATCATTCCATTAAACATATTTCTAAATAATAGTAAGTGGTACTAACAAAATAAA
TAATAATTTAATAGCCTTAGAAATAAATGACTGTATACTTATACAGGTTGAAAAAAACTCG

AGTTGGCTGACATCTGGAAGTAGTTAACAACTAACCAGTGGACTTCAACAATCATTTGCTC
CCAGGCTTCCCCCATCATCACCCTCACCACATATCCTGCTAATATCCAACAACAAACAAAT
ATTTAATATTGAAATAGCCCATTGCCTGAGAATGAACACAAGCTAAAATACATGCAAGGGT
ACTTAATGGAAGCCAAACCATGTTCTATACCTAAGGAGAAAACATGGACATGTAGAATGCT
TTTATTCATGTATTCAAAATCAGAACAAATCAGTGTATATCACTAGAACATCAGATGGAGG
ATAACACAAGAAGTGATACAGATCAGGGTTCACTTCTCTTACCCTCTCTCTGTTAGGACCA
CATTCCTATTTTAGCCAAATGTTTCTGGTACGGGCCATCTTTTCACCATAAATGGCATTAT
GTTTCAAATGGCTAAAAGCATATTATGGGTATGCTCAAAGGAGTAAAACCCAATTCACAGA
GATGTGGCTTTTCTAAAGAACCAAATTAAGGGAATACTTAGTATCTACAACTAAAGTATTA
GTCATTCTGAGGATTATTTGTTGCTTTAAGTATTAGCTCCCTACTACATCCATAACACCCT
ACTTACTAAATTTAATTACACACAACTTTTCAAGAATTCGTATTTTATTTGAAGGGAGGTA
CCTGTCTACTTTATCTACAATAAAAACAAAGGGTATTGGCTTTCTCTAATCCATGCAAACT
ACAAATTCCATCGGGAGTCCTACATCACTAACAGTGGTATGAACAAAACTAAGAAAGTACT
TCCTATCCACATGTGAATGTTTTAAAAAAGTTTTTGCCATAAAACCTAAGTGTAATTTAGC
ATACCACAGTGCTCTGAAGATGGGTCATTGACGATGTACCATTTGTATATAGGTAATACAC
ATGTAAATCACTAATTGTTAATTAAATATAAGAAGTAGGTTTTTATCTTTTTAAAAAAAAA
ACAAAAAAGGTGACTCCCTTTCTCATTCTGTTCTGTTGTATTGCGTCCAAAAGTTGTGTGT
AATTTTTTTAAGGTCCAAGAAATGTAAAGAAATTTGTTGGAATACCTGAATTTCTGTAAAA
AAATAAAAATAAAAATAAGATTTTGTTACTTAAAA

FIGURA 67

GGGAGGCTGTGCCAGGCGAGCCGGAGGGGTGCTCCGCGCTCCCCCGCCCTCCTTCCGGGAG
CGAGGATGCAGACTCTGAAACTGGTGCTGCTGGGCTGAGGCGGAGGCAGGGGAGTTGCAGC
GCGCGAGGCTCCGTGAGTGTGTCTCCTGCGCGCTGAGAGGCGGGGGGAGGCGGAGGACCAG
GAGGAGGAGGAGGAGGAGGAGGAGGGGGAGAATGCCCGGAGCCGCCGCCGCTGCCGCCGCC
GCCGCCGCCGCGATGCTCCCGGCTCAGGAGGCTGCCAAGCTGTACCACACCAACTATGTGC
GGAACTCGCGGGCCATCGGCGTGCTGTGGGCCATCTTCACCATCTGCTTTGCCATCGTCAA
CGTGGTGTGCTTCATCCAGCCCTACTGGATAGGCGACGGCGTGGACACCCCGCAAGCCGGC
TATTTCGGGCTCTTCCACTACTGCATCGGCAACGGCTTCTCCCGGGAGCTGACCTGCAGGG
GCAGCTTCACGGACTTCTCCACGCTGCCCTCGGGCGCCTTCAAAGCCGCCTCCTTCTTTAT
CGGCCTCTCCATGATGCTCATCATTGCCTGCATCATTTGCTTTACCCTCTTCTTCTTCTGC
AACACGGCCACTGTGTACAAGATATGTGCCTGGATGCAGCTCACCTCCGCTGCCTGCCTTG
TGCTTGGCTGTATGATTTTCCCTGATGGCTGGGACTCAGATGAAGTAAAACGGATGTGTGG
AGAAAAGACAGACAAGTACACTCTTGGGGCTTGCTCAGTCCGCTGGGCATACATCCTGGCT
ATTATTGGAATTTTGGATGCCCTGATCCTCTCATTTCTAGCATTTGTGCTTGGTAATCGAC
AAGACAGCTTGATGGCAGAGGAACTGAAGGCAGAAAACAAAGTTCTGCTAAGCCAATATTC
TCTAGAATGAGCACAAAACAAATCGAATAACAGCTAAACAAATCGAATAACAGCTAAACGA
ATCGAATAACAGCTTTTGTACATCAACATCAAGAAGGAATACGCCTGAGAGAGATCAGAGT
ATATAGATGAATATGAACAAGAATGGAACATTCACTTGTCAACGCACTTTCTAAATCTAGA
TCAGCAGAGATGGGAGTGATTTTCTGGAAAGAGATGTGATCATGGATTAAACACCAGCTCA
TTGGAAACTCATTGGATGAGATCAGAAAACGTTCATGAAAAATCATATTCAGGAAATAAGG
AAGAGGAATATAAATGCTCTAGAGTTAACATGTAAAATATATACGTACTGAGGTTTGTAAA
CTGTCCTTTTTAAATCAAACTGAAAACAAAAAGCTTTAACCTTTCAACAGAATTTTTAAAA
AGGCAGTTAGTTCTAAATTATTCCTATCTCAATAGCCAAGAGGCTGATCAAGCGTCATTTA
TTGAGGAAGCATCTTAGAAAATGCCTCTGAATGTTTTCATAGGAGCCGTGACCTTTGGTTC
TTCATCTCTACCATTCATTTACTTCACTGTGTAATTAGTTACAACCACTCAGTTATTAAGA
GACGTAACGCTTCAAACTTTTTACCAAGTCTGTGTTCTGTTTAATCTGTCCATACAAGTTA
TTACTGAGAAAGTGTTTATGCCATATACTATTACTCCATCAAGCTGTATATTACAGGAAGT
ACATCTTTACATCATAGGTTCCCAAGCAACATAGATTTCCCTATCTTTCAGGAAACAGCAT
CAAGGAACTCTGAAAAATATAGAAAAAGTTCATTTTCACCTTGGAAGCTCACGTGTAATAT
TATAGGCTACTATCAAATAAACACTTTTTTTCTAATTCTCCCTAGTATATGCATAGGAATT
TAATATACTTTATAAATAAGTATCTAAAATGTCTCCTACTTTTTTCCTATTTCTTTGCCAT
ACATGTTATCAGAAATCCATGTCTTCTATTTCCCTTACTGATGGGCACTCATTTTTATTTT
TTTAAAAATCATTCCATTAAACATATTTCTAAATAATAGTAAGTGGTACTAACAAAATAAA
TAATAATTTAATAGCCTTAGAAATAAATGACTGTATACTTATACAGGTTGAAAAAAACTCG
GTAGGAATAAGTTACCTTTTTTGTTTACTAATGTTGGTTTCAAAAATACTCAGATTCATTTT

AGTTGGCTGACATCTGGAAGTAGTTAACAACTAACCAGTGGACTTCAACAATCATTTGCTC
CCAGGCTTCCCCCATCATCACCCTCACCACATATCCTGCTAATATCCAACAACAAACAAAT
ATTTAATATTGAAATAGCCCATTGCCTGAGAATGAACACAAGCTAAAATACATGCAAGGGT
ACTTAATGGAAGCCAAACCATGTTCTATACCTAAGGAGAAAACATGGACATGTAGAATGCT
TTTATTCATGTATTCAAAATCAGAACAAATCAGTGTATATCACTAGAACATCAGATGGAGG
ATAACACAAGAAGTGATACAGATCAGGGTTCACTTCTCTTACCCTCTCTCTGTTAGGACCA
CATTCCTATTTTAGCCAAATGTTTCTGGTACGGGCCATCTTTTCACCATAAATGGCATTAT
GTTTCAAATGGCTAAAAGCATATTATGGGTATGCTCAAAGGAGTAAAACCCAATTCACAGA
GATGTGGCTTTTCTAAAGAACCAAATTAAGGGAATACTTAGTATCTACAACTAAAGTATTA
GTCATTCTGAGGATTATTTGTTGCTTTAAGTATTAGCTCCCTACTACATCCATAACACCCT
ACTTACTAAATTTAATTACACACAACTTTTCAAGAATTCGTATTTTATTTGAAGGGAGGTA
CCTGTCTACTTTATCTACAATAAAAACAAAGGGTATTGGCTTTCTCTAATCCATGCAAACT
ACAAATTCCATCGGGAGTCCTACATCACTAACAGTGGTATGAACAAAACTAAGAAAGTACT
TCCTATCCACATGTGAATGTTTTAAAAAAGTTTTTGCCATAAAACCTAAGTGTAATTTAGC
ATACCACAGTGCTCTGAAGATGGGTCATTGACGATGTACCATTTGTATATAGGTAATACAC
ATGTAAATCACTAATTGTTAATTAAATATAAGAAGTAGGTTTTTATCTTTTTAAAAAAAAA
ACAAAAAAAGGTGACTCCCTTTCTCATTCTGTTCTGTTGTATTGCGTCCAAAAGTTGTGTGT
AATTTTTTTAAGGTCCAAGAAATGTAAAGAAATTTGTTGGAATACCTGAATTTCTGTAAAA
AAATAAAAATAAAAATAAGATTTTGTTACTTAAAA

FIGURA 68

GGGAGGCTGTGCCAGGCGAGCCGGAGGGGTGCTCCGCGCTCCCCCGCCCTCCTTCCGGGAG
CGAGGATGCAGACTCTGAAACTGGTGCTGCTGGGCTGAGGCGGAGGCAGGGGAGTTGCAGC
GCGCGAGGCTCCGTGAGTGTGTCTCCTGCGCGCTGAGAGGCGGGGGGAGGCGGAGGACCAG
GAGGAGGAGGAGGAGGAGGAGGAGGGGGAGAATGCCCGGAGCCGCCGCCGCTGCCGCCGCC
GCCGCCGCCGCGATGCTCCCGGCTCAGGAGGCTGCCAAGCTGTACCACACCAACTATGTGC
GGAACTCGCGGGCCATCGGCGTGCTGTGGGCCATCTTCACCATCTGCTTTGCCATCGTCAA
CGTGGTGTGCTTCATCCAGCCCTACTGGATAGGCGACGGCGTGGACACCCCGCAAGCCGGC
TATTTCGGGCTCTTCCACTACTGCATCGGCAACGGCTTCTCCCGGGAGCTGACCTGCAGGG
GCAGCTTCACGGACTTCTCCACGCTGCCCTCGGGCGCCTTCAAAGCCGCCTCCTTCTTTAT
CGGCCTCTCCATGATGCTCATCATTGCCTGCATCATTTGCTTTACCCTCTTCTTCTTCTGC
AACACGGCCACTGTGTACAAGATATGTGCCTGGATGCAGCTCACCTCCGCTGCCTGCCTTG
TGCTTGGCTGTATGATTTTCCCTGATGGCTGGGACTCAGATGAAGTAAAACGGATGTGTGG
AGAAAAGACAGACAAGTACACTCTTGGGGCTTGCTCAGTCCGCTGGGCATACATCCTGGCT
ATTATTGGAATTTTGGATGCCCTGATCCTCTCATTTCTAGCATTTGTGCTTGGTAATCGAC
AAGACAGCTTGATGGCAGAGGAACTGAAGGCAGAAAACAAAGTTCTGCTAAGCCAATATTC
TCTAGAATGAGCACAAAACAAATCGAATAACAGCTAAACAAATCGAATAACAGCTAAACGA
ATCGAATAACAGCTTTTGTACATCAACATCAAGAAGGAATACGCCTGAGAGAGATCAGAGT
ATATAGATGAAATATGAACAAGAATGGAACATTCACTTGTCAACGCACTTTCTAAATCTAGA
TCAGCAGAGATGGGAGTGATTTTCTGGAAAGAGATGTGATCATGGATTAAACACCAGCTCA
TTGGAAACTCATTGGATGAGATCAGAAAACGTTCATGAAAAATCATATTCAGGAAATAAGG
AAGAGGAATATAAATGCTCTAGAGTTAACATGTAAAATATATACGTACTGAGGTTTGTAAA
CTGTCCTTTTTAAATCAAACTGAAAACAAAAAGCTTTAACCTTTCAACAGAATTTTTAAAA
AGGCAGTTAGTTCTAAATTATTCCTATCTCAATAGCCAAGAGGCTGATCAAGCGTCATTTA
TTGAGGAAGCATCTTAGAAAATGCCTCTGAATGTTTTCATAGGAGCCGTGACCTTTGGTTC
TTCATCTCTACCATTCATTTACTTCACTGTGTAATTAGTTACAACCACTCAGTTATTAAGA
GACGTAACGCTTCAAACTTTTTACCAAGTCTGTGTTCTGTTTAATCTGTCCATACAAGTTA
TTACTGAGAAAGTGTTTATGCCATATACTATTACTCCATCAAGCTGTATATTACAGGAAGT
ACATCTTTACATCATAGGTTCCCAAGCAACATAGATTTCCCTATCTTTCAGGAAACAGCAT
CAAGGAACTCTGAAAAATATAGAAAAAGTTCATTTTCACCTTGGAAGCTCACGTGTAATAT
TATAGGCTACTATCAAATAAACACTTTTTTTCTAATTCTCCCTAGTATATGCATAGGAATT
TAATATACTTTATAAATAAGTATCTAAAATGTCTCCTACTTTTTTCCTATTTCTTTGCCAT
ACATGTTATCAGAAATCCATGTCTTCTATTTCCCTTACTGATGGGCACTCATTTTTATTTT
TTTAAAAATCATTCCATTAAACATATTTCTAAATAATAGTAAGTGGTACTAACAAAATAAA
TAATAATTTAATAGCCTTAGAAATAAATGACTGTATACTTATACAGGTTGAAAAAAACTCG
[illegible]TTCAAAAATACTCAGATTCATTTT

AGTTGGCTGACATCTGGAAGTAGTTAACAACTAACCAGTGGACTTCAACAATCATTTGCTC
CCAGGCTTCCCCCATCATCACCCTCACCACATATCCTGCTAATATCCAACAACAAACAAAT
ATTTAATATTGAAATAGCCCATTGCCTGAGAATGAACACAAGCTAAAATACATGCAAGGGT
ACTTAATGGAAGCCAAACCATGTTCTATACCTAAGGAGAAAACATGGACATGTAGAATGCT
TTTATTCATGTATTCAAAATCAGAACAAATCAGTGTATATCACTAGAACATCAGATGGAGG
ATAACACAAGAAGTGATACAGATCAGGGTTCACTTCTCTTACCCTCTCTCTGTTAGGACCA
CATTCCTATTTTAGCCAAATGTTTCTGGTACGGGCCATCTTTTCACCATAAATGGCATTAT
GTTTCAAATGGCTAAAAGCATATTATGGGTATGCTCAAAGGAGTAAAACCCAATTCACAGA
GATGTGGCTTTTCTAAAGAACCAAATTAAGGGAATACTTAGTATCTACAACTAAAGTATTA
GTCATTCTGAGGATTATTTGTTGCTTTAAGTATTAGCTCCCTACTACATCCATAACACCCT
ACTTACTAAATTTAATTACACACAACTTTTCAAGAATTCGTATTTTATTTGAAGGGAGGTA
CCTGTCTACTTTATCTACAATAAAAACAAAGGGTATTGGCTTTCTCTAATCCATGCAAACT
ACAAATTCCATCGGGAGTCCTACATCACTAACAGTGGTATGAACAAAACTAAGAAAGTACT
TCCTATCCACATGTGAATGTTTTAAAAAAGTTTTTGCCATAAAACCTAAGTGTAATTTAGC
ATACCACAGTGCTCTGAAGATGGGTCATTGACGATGTACCATTTGTATATAGGTAATACAC
ATGTAAATCACTAATTGTTAATTAAATATAAGAAGTAGGTTTTTATCTTTTTAAAAAAAAA
ACAAAAAAGGTGACTCCCTTTCTCATTCTGTTCTGTTGTATTGCGTCCAAAAGTTGTGTGT
AATTTTTTTAAGGTCCAAGAAATGTAAAGAAATTTGTTGGAATACCTGAATTTCTGTAAAA
AAATAAAAATAAAAATAAGATTTTGTTACTTAAAA

FIGURA 69

CAGGCCTCTGAGGCTCCCTTGCCGAGGGCCCCGAGCTGCAGGGACAGTGAGCAGTGAGTCC
CTTGGGCATCCCGCTCCTGGGCAGGTCACCAATAGGTCCCCGCAGTTCCCAATGGAACTGT
TCCAGTCCTCCCCGAGGCCTCCACTTCAACCTGTCTGTGTCTGCCCAGGCCTGGAGTTGTG
TGACCCTCCCCACCGCCTGGCCTTCTCCATGGGGGCTGGCCTTTTCTCGGTGGTGGGCACC
CTGCTGCTGCCCGGCCTGGCTGCGCTTGTGCAGGACTGGCGTCTTCTGCAGGGGCTGGGTG
CCCTGATGAGTGGACTCTTGCTGCTCTTTTGGGGGTTCCCGGCCCTGTTCCCCGAGTCTCC
CTGCTGGCTGCTGGCCACAGGTCAGGTAGCTCGAGCCAGGAAGATCCTGTGGCGCTTTGCA
GAAGCCAGTGGCGTGGGCCCCGGGGACAGTTCCTTGGAGGAGAACTCCCTGGCTACAGAGC
TGACCATGCTGTCTGCACGGAGCCCCCAGCCCCGGTACCACTCCCCACTGGGGCTTCTGCG
TACCCGAGTCACCTGGAGAAACGGGCTTATCTTGGGCTTCAGCTCGCTGGTTGGTGGAGGC
ATCAGAGCTAGCTTCCGCCGCAGCCTGGCACCTCAGGTGCCGACCTTCTACCTGCCCTACT
TCCTGGAGGCCGGCCTGGAGGCGGCAGCCTTGGTCTTCCTGCTCCTGACGGCAGATTGCTG
TGGACGCCGCCCCGTGCTGCTGCTGGGCACCATGGTCACAGGCCTGGCATCCCTGCTGCTC
CTCGCTGGGGCCCAGTATCTGCCAGGCTGGACTGTGCTGTTCCTCTCTGTCCTGGGGCTCC
TGGCCTCCCGGGCTGTGTCCGCACTCAGCAGCCTCTTCGCGGCCGAGGTCTTCCCCACGGT
GATCAGGGGGGCCGGGCTGGGCCTGGTGCTGGGGGCCGGGTTCCTGGGCCAGGCAGCCGGC
CCCCTGGACACCCTGCACGGCCGGCAGGGCTTCTTCCTGCAACAAGTCGTCTTCGCCTCCC
TTGCTGTCCTTGCCCTGCTGTGTGTCCTGCTGCTGCCTGAGAGCCGAAGCCGGGGGCTGCC
CCAGTCACTGCAGGACGCCGACCGCCTGCGCCGCTCCCCACTCCTGCGGGGCCGCCCCCGC
CAGGACCACCTGCCTCTGCTGCCGCCCTCCAACTCCTACTGGGCCGGCCACACCCCCGAGC
AGCACTAGTCCTGCCTGGTGGCCCTGGGAGCCAGGATGGGACCAAAGTCAAGGCCTGGGGC
ATGGCTGAGTACCCCAGACGTCTGGTCCAGGGCAGACACATTCCTCTCAGAAGCCCGTGTC
TCAGTGCAGGTGGAGCCGTGGGGACAGCGTGAAGGTGTCTCCAGCCAGGCCCCAGGCACTG
GGAGGCCCTGGGTCTCCCCCCAGCCACACCCAGTAGGTGTGGAGGATAAAGGCTTCTGTGG
AACTTGGCGTCCTCCCTGGTTCTTCCCTCAGGCCACACCTGGGGCCTGGAAGGGTCTCACC
TGTGCCTGCTCCCCACCTCTGGGCTCACCAGCAAGGAGCCAGGGTGTCAGCGGCTGCTCCA
TAGCCCCTGTCCAGGTAGTGGCAGCCCCCGGCCGCGTCCTGCCCCAGCCAAGCAGGGTGGT
ATGGGGGCCTGGCCCTCAGGTCCCTGCCAGGCCAGGCCTCACTGGCCCCATCCCCCTCATG
GTCAGTGTGGCAGTGGGCGCCCGTCCTGCATGAAAGAGACTAGCACG

FIGURA 70

AGCGGGGCGGGTCCCCGAAGCCCAGTACCAGACCTCCrGTGCGAAGAGTAACCTGCTCTTC
GGACTGATAACTGGATGGAATGGCCTTTAGAAAGGGAGCATCTTCAGAATCTGCAGCTGGG
ATCAGATTGTTCTTTTGTTGAATCAACTCAGCTCTGCCGAGCCAGCATCTTGCTAAGGGTT
ACAAAATAAATGCCATTTGAACAGTGCCATCTGTCATGGAAAAACGTGAGACGTTTGTACA
ATGCCGTGTCTAAGGAGCTGGTTGGAGAGTTTTTGCAATTTGTTCAACTTGATAAAGAGGC
CTCTGATCCTTTCAGCCTAAATGAATTACTAGATGAATTATCAAGGAAACAGAAAGAAGAA
TTATGGCAAAGGCTGAAGAATTTATTGACAGATGTGTTGTTAGAAAGCCCAGTGGATGGGT
GGCAGGTAGTGGAAGCCCAGGGTGAAGACAATATGGAAACCGAACATGGCTCAAAAATGAG
AAAAAGCATAGAAATAATTTATGCAATTACATCTGTGATTCTTGCTTCTGTGTCTGTAATA
AATGAAAGTGAGAACTACGAAGCCCTACTGGAATGTGTTATTATATTAAATGGTATTTTAT
ATGCATTACCTGAGTCTGAACGAAAACTACAGAGTTCTATTCAGGATTTGTGTGTTACCTG
GTGGGAGAAAGGCCTGCCTGCCAAGGAAGACACAGGAAAGACAGCCTTTGTCATGTTACTA
AGGAGGAGTCTGGAGACTAAGACAGGTGCAGACGTATGTCGGCTTTGGCGTATCCATCAAG
CTTTATATTGCTTTGATTATGATTTGGAGGAAAGTGGAGAAATTAAAGATATGTTACTTGA
GTGCTTCATAAATATTAATTATATCAAGAAAGAAGAGGGAAGAAGATTTCTTAGTTGTCTC
TTCAACTGGAATATCAACTTCATCAAAATGATCCACGGGACCATTAAAAACCAGTTACAGG
GATTACAAAAGTCTTTGATGGTATACATTGCAGAAATTTATTTCAGAGCTTGGAAAAAGGC
TTCAGGGAAAATACTGGAGGCGATTGAAAATGATTGCATCCAGGACTTCATGTTCCACGGG
ATACACCTTCCGAGGAGGTCTCCAGTGCATTCCAAAGTGCGGGAGGTTCTGAGTTACTTTC
ACCATCAAAAGAAAGTTCGGCAGGGAGTGGAAGAGATGCTTTATAGATTATATAAGCCCAT
CCTTTGGAGAGGATTAAAGGCCAGAAACTCTGAAGTTCGATCAAATGCTGCATTGTTGTTT
GTTGAAGCATTTCCTATTAGGGATCCAAACCTTCATGCTATTGAAATGGATAGTRAAATCC
AGAAACAGTTTGAAGAGCTCTATAGCCTTTTAGAAGATCCTTACCCGATGGTCCGTTCCAC
AGGGATCCTTGGTGTTTGTAAAATAACTTCTAAGTACTGGGAAATGATGCCCCCGACCATT
CTTATTGACCTCCTGAAGAAGGTGACTGGGGAACTGGCATTTGACACGAGCTCAGCTGATG
TTCGTTGTTCTGTCTTTAAGTGTCTGCCAATGATTTTGGACAACAAACTGAGCCACCCATT
GTTAGAGCAGCTCCTTCCAGCTCTCAGATACAGTCTCCACGACAATTCGGAGAAAGTGAGG
GTAGCTTTTGTGGACATGCTGTTGAAGATCAAAGCTGTGAGGGCTGCTAAGTTTTGGAAAA
TATGTCCCATGGAGCACATTCTGGTTCGTCTGGAAACTGATTCTCGACCTGTGTCTCGGCG
CCTGGTGAGCCTCATCTTTAATTCTTTCCTGCCTGTGAATCAGCCGGAGGAGGTCTGGTGC
GAGCGCTGTGTCACCCTGGTGCAGATGAACCACGCCGCTGCCAGGAGGTTCTATCAGTACG
CCCACGAACACACCGCCTGCACCAACATAGCAAAGCTGATTCACGTTATTCGTCATTGCTT
AAATGCCTGTATCCAGAGGGCAGTGAGAGAGCCTCCAGAGGACGAGGAGGAAGAGGACGGA
AGGGAGAAGGAGAATGTGACTGTTCTGGACAAAACACTGTCAGTAAACGATGTTGCATGCA
TGGCAGGTTTACTAGAAATCATTGTGATTCTCTGGAAAAGTATTGACAGAAGTATGGAAAA

TAATAAAGAGGCCAAACTTTACACGATTAACAAGTTTGCCTCTGTGCTTCCAGAGTATCTG
AAAGTATTTAAGGATGATCGCTGCAAGATCCCTTTATTCATGCTAATGTCCTTTATGCCGG
CCTCTGCTGTCCCCCCATTCAGCTGTGGTGTGATTTCCACGCTGAGAAGCCGGGAGGAGGG
CGCTGTGGACAAGAGCTACTGCACTTTGTTGGATTGCCTCTGCTCCTGGGGGCAGGTGGGG
CACATTCTGGAGCTTGTTGACAACTGGCTGCCCACAGAGCATGCCCAGGCCAAGAGCAACA
CAGCTTCTAAAGGTAGGGTGCAGATCCATGACACACGCCCAGTCAAACCTGAATTGGCATT
GGTCTACATTGAGTATCTGCTGACTCATCCAAAGAACCGCGAGTGCTTGCTCTCTGCTCCT
CGGAAGAAACTTAACCATCTTTTGAAAGCCCTTGAAACGTCAAAGGCAGATCTGGAGTCAC
TTCTGCAGACACCGGGTGGGAAGCCTCGTGGCTTCAGTGAAGCAGCTGCCCCGCGAGCCTT
TGGTCTCCACTGTCGCCTGAGCATCCATCTTCAGCACAAGTTCTGCTCAGAAGGAAAGGTG
TATTTGTCCATGTTGGAAGACACTGGCTTTTGGTTAGAAAGCAAAATTTTATCTTTTATTC
AAGATCAAGAAGAAGACTACCTGAAGCTTCATAGGGTCATTTATCAGCAAATTATCCAGAC
CTACCTGACTGTGTGTAAAGATGTTGTTATGGTAGGCCTTGGTGACCATCAGTTTCAGATG
CAACTCTTACAGCGGAGTCTTGGAATCATGCAAACAGTGAAGGGATTTTTTTATGTTTCAT
TACTTCTTGACATTCTGAAAGAGATAACTGGAAGTTCCTTGATTCAGAAAACAGATTCAGA
TGAAGAAGTTGCAATGCTGTTGGACACAGTCCAGAAAGTATTTCAGAAAATGTTGGAATGT
ATTGCACGGAGCTTCAGGAAGCAGCCGGAAGAAGGCCTGCGGCTGCTTTATTCTGTTCAGA
GGCCTCTTCATGAGTTCATTACTGCTGTTCAGTCTCGGCACACAGACACCCCTGTGCACCG
GGGTGTACTTTCTACTCTGATCGCTGGGCCTGTGGTTGAGATAAGTCACCAGCTACGGAAG
GTTTCTGACGTAGAAGAGCTTACCCCTCCAGAGCATCTTTCTGATCTTCCACCATTTTCAA
GGTGTTTAATAGGAATAATAATAAAGTCTTCGAATGTGGTCAGGTCATTTTTGGATGAATT
AAAGGCATGTGTGGCTTCTAATGATATTGAAGGCATTGTGTGCCTCACGGCTGCTGTGCAT
ATTATCCTGGTTATTAATGCAGGTAAACATAAAAGCTCAAAAGTGAGGGAGGTTGCAGCCA
CTGTTCACAGAAAACTAAAGACATTCATGGAAATTACTTTGGAAGAGGATAGCATTGAAAG
ATTTCTCTATGAATCATCATCAAGAACTCTGGGAGAACTTTTGAATTCATAACCAAGCCAA
CATCTCCAGACATGTAAAAATAGGGAAAAGTGATTCAAATTGAAATGCCTGTGTATTTTCC
TATTGTTTTTAATGTTAATAACCCATATAATAGGGAAAGGGTGGGATTTTTTTGTGGGAAT
GTGGGAAGGTGGGGGTTATGGAGGAGATAACTCAAAACTTCTTCAATTTTGCCTAGTGCCT
GCGTAAATAATATATTTAATATAAAGGACTCCAGGTATGAATGGTGTAGAAATCCATGATT
CCAAGAAAAAACACTTTTCTAGCAAACCTGGTTGTTTTTAAAATGACTTTTATATATGTAA
TATTGCTTGGAAACTATGAGTAATAAAGCAATGACAACAAMAAAAAAAAAAAAAAA

FIGURA 71

GCCCTTCAGAATCTGCAGCTGGGATCAGATTGTTCTTTTGTTGAATCAACTCAGCTCTGCC
GAGCCAGCATCTTGCTAAGGGTTACAAAATAAATGCCATTTGAACAGTGCCATCTGTCATG
GAAAAACGTGAGACGTTTGTACAAGCCGTGTCTAAGGAGCTGGTTGGAGAGTTTTTGCAAT
TTGTTCAACTTGATAAAGAGGCCTCTGATCCTTTCAGCCTAAATGAATTACTAGATGAATT
ATCAAGGAAACAGAAAGAAGAATTATGGCAAAGGCTGAAGAATTTATTGACAGATGTGTTG
TTAGAAAGCCCAGTGGATGGGTGGCAGGTAGTGGAAGCCCAGGGTGAAGACAATATGGAAA
CCGAACATGGCTCAAAAATGAGAAAAAGCATAGAAATAATTTATGCAATTACATCTGTGAT
TCTTGCTTCTGTGTCTGTAATAAATGAAAGTGAGAACTACGAAGCCCTACTGGAATGTGTT
ATTATATTAAATGGTATTTTATATGCATTACCTGAGTCTGAACGAAAACTACAGAGTTCTA
TTCAGGATTTGTGTGTTACCTGGTGGGAGAAAGGCCTGCCTGCCAAGGAAGACACAGGAAA
GACAGCCTTTGTCATGTTACTAAGGAGGAGTCTGGAGACTAAGACAGGTGCAGACGTATGT
CGGCTTTGGCGTATCCATCAAGCTTTATATTGCTTTGATTATGATTTGGAGGAAAGTGGAG
AAATTAAAGATATGTTACTTGAGTGCTTCATAAATATTAATTATATCAAGAAAGAAGAGGG
AAGAAGATTTCTTAGTTGTCTCTTCAACTGGAATATCAACTTCATCAAAATGATCCACGGG
ACCATTAAAAACCAGTTACAGGGATTACAAAAGTCTTTGATGGTATACATTGCAGAAATTT
ATTTCAGAGCTTGGAAAAAGGCTTCAGGGAAAATACTGGAGGCGATTGAAAATGATTGCAT
CCAGGACTTCATGTTCCACGGGATACACCTTCCGAGGAGGTCTCCAGTGCATTCCAAAGTG
CGGGAGGTTCTGAGTTACTTTCACCATCAAAAGAAAGTTCGGCAGGGAGTGGAAGAGATGC
TTTATAGATTATATAAGCCCATCCTTTGGAGAGGATTAAAGGCCAGAAACTCTGAAGTTCG
ATCAAATGCTGCATTGTTGTTTGTTGAAGCATTTCCTATTAGGGATCCAAACCTTCATGCT
ATTGAAATGGATAGTGAAATCCAGAAACAGTTTGAAGAGCTCTATAGCCTTTTAGAAGATC
CTTACCCGATGGTCCGTTCCACAGGGATCCTTGGTGTTTGTAAAATAACTTCTAAGTACTG
GGAAATGATGCCCCCGACCATTCTTATTGACCTCCTGAAGAAGGTGACTGGGGAACTGGCA
TTTGACACGAGCTCAGCTGATGTTCGTTGTTCTGTCTTTAAGTGTCTGCCAATGATTTTGG
ACAACAAACTGAGCCACCCATTGTTAGAGCAGCTCCTTCCAGCTCTCAGATACAGTCTCCA
CGACAATTCGGAGAAAGTGAGGGTAGCTTTTGTGGACATGCTGTTGAAGATCAAAGCTGTG
AGGGCTGCTAAGTTTTTGGAAAATATGTCCCATGGAGCACATTCTGGTTCGTCTGGAAACTG
ATTCTCGACCTGTGTCTCGGCGCCTGGTGAGCCTCATCTTTAATTCTTTCCTGCCTGTGAA
TCAGCCGGAGGAGGTCTGGTGCGAGCGCTGTGTCACCCTGGTGCAGATGAACCACGCCGCT
GCCAGGAGGTTCTATCAGTACGCCCACGAACACACCGCCTGCACCAACATAGCAAAGCTGA
TTCACGTTATTCGTCATTGCTTAAATGCCTGTATCCAGAGGGCAGTGAGAGAGCCTCCAGA
GGACGAGGAGGAAGAGGACGGAAGGGAGAAGGAGAATGTGACTGTTCTGGACAAAACACTG
TCAGTAAACGATGTTGCATGCATGGCAGGTTTACTAGAAATCATTGTGATTCTCTGGAAAA
GTATTGACAGAAGTATGGAAAATAATAAAGAGGCCAAACTTTACACGATTAACAAGTTTGC
CTCTTGTGCTTTCCAGAGTATCTGAAAGTATTTAAGGATGATCGCTGCAAGATCCCTTTATTC

ATGCTAATGTCCTTTATGCCGGCCTCTGCTGTCCCCCCATTCAGCTGTGGTGTGATTTCCA
CGCTGAGAAGCCGGGAGGAGGGCGCTGTGGACAAGAGCTACTGCACTTTGTTGGATTGCCT
CTGCTCCTGGGGGCAGGTGGGGCACATTCTGGAGCTTGTTGACAACTGGCTGCCCACAGAG
CATGCCCAGGCCAAGAGCAACACAGCTTCTAAAGGTAGGGTGCAGATCCATGACACACGCC
CAGTCAAACCTGAATTGGCATTGGTCTACATTGAGTATCTGCTGACTCATCCAAAGAACCG
CGAGTGCTTGCTCTCTGCTCCTCGGAAGAAACTTAACCATCTTTTGAAAGCCCTTGAAACG
TCAAAGGCAGATCTGGAGTCACTTCTGCAGACACCGGGTGGGAAGCCTCGTGGCTTCAGTG
AAGCAGCTGCCCCGCGAGCCTTTGGTCTCCACTGTCGCCTGAGCATCCATCTTCAGCACAA
GTTCTGCTCAGAAGGAAAGGTGTATTTGTCCATGTTGGAAGACACTGGCTTTTGGTTAGAA
AGCAAAATTTTATCTTTTATTCAAGATCAAGAAGAAGACTACCTGAAGCTTCATAGGGTCA
TTTATCAGCAAATTATCCAGACCTACCTGACTGTGTGTAAAGATGTTGTTATGGTAGGCCT
TGGTGACCATCAGTTTCAGATGCAACTCTTACAGCGGAGTCTTGGAATCATGCAAACAGTG
AAGGGATTTTTTTATGTTTCATTACTTCTTGACATTCTGAAAGAGATAACTGGAAGTTCCT
TGATTCAGAAAACAGATTCAGATGAAGAAGTTGCAATGCTGTTGGACACAGTCCAGAAAGT
ATTTCAGAAAATGTTGGAATGTATTGCACGGAGCTTCAGGAAGCAGCCGGAAGAAGGCCTG
CGGCTGCTTTATTCTGTTCAGAGGCCTCTTCATGAGTTCATTACTGCTGTTCAGTCTCGGC
ACACAGACACCCCTGTGCACCGGGGTGTACTTTCTACTCTGATCGCTGGGCCTGTGGTTGA
GATAAGTCACCAGCTACGGAAGGTTTCTGACGTAGAAGAGCTTACCCCTCCAGAGCATCTT
TCTGATCTTCCACCATTTTCAAGGTGTTTAATAGGAATAATAATAAAGTCTTCGAATGTGG
TCAGGTCATTTTTGGATGAATTAAAGGCATGTGTGGCTTCTAATGATATTGAAGGCATTGT
GTGCCTCACGGCTGCTGTGCATATTATCCTGGTTATTAATGCAGGTAAACATAAAAGCTCA
AAAGTGAGGGAGGTTGCAGCCACTGTTCACAGAAAACTAAAGACATTCATGGAAATTACTT
TGGAAGAGGATAGCATTGAAAGATTTCTCTATGAATCATCATCAAGAACTCTGGGAGAACT
TTTGAATTCATAACCAAGCCAACATCTCCAGACATGAAGGGC

FIGURA 72

TCCAACCAGATTTTCACCGCTATGCCTCCTTTCACTCTGGGAGTCTTCCAGAGGTCTTGCA
CTCGGGAGAGCATGCTCAGGTTTCCCCAGCTCTACAAAATCACCCAGAATGCCAAAGACTT
CAACGTGACCAGTGGTCATGCCACTGCCTGTTGATTTGTTGAAAATATTGTTTACACGTAT
GTTCTTGTTACTGATTGTCAGAAAGCTGGTTTTGAGACTGCAGCTTGGACTAAATTCAGTC
ATCTGGCTGTCTGGGGAAGCATGCTGACCAGTCTGGTGTTCTTTGGCATCTACTCAGCCAT
CTGGTCCACCATTCTCATTGCCCCAAATATGAGAGGACAGAAGAATGGTACCGGTACTGCC
AATGGAGATGGAGGAAGGAGACAGAAAGAAACAGAGCCCAGACCCTAGGGACCACCAGCAT
TTGCAGAATGGATAAACAGCCTTCTTCCTAACAAAGGAAGCACAGCAACTGTGATCCTGAG
CTGTGCACACTTCTGGTTGGGATTATTTCTGGTTTCTACTTCCTGTTTGAAGATGTGGCAT
GGAGAGTGAACAAGCTGCTGCCCACCACCTGGCATCACAGCCCCAGAACTCAGCTATTTCC
ATGGGACCACAGCATCTCATCTCTGGGCTGAGCCAGAAAGACCCCTACTGAAGTCCAGAGG
CACTTTTCTGAAAGGCTCTGCTTTGACCTGAAGTATTTTATCTATCCTCAGTCTCAGGACA
CTGTTGATGGAATTAAGGCCAAGCACATCTGCAAAAAAGACATTGCTGGAGGAGGTGCAAA
GAGCTGGAAACCAAGTCTCCAGTCCTGGGAAAAGCAGTGGTATGGAAAAGCAATGGAAAGA
GCATTTTGAAAATGCCATTCCACTGTTTTCTGGCCTTTATGATTTCTGCTGAGAAATCCAC
TGTTAGTCTGATGGGGTCTCCTTCATAGCACCAATGACCTGAAGAGCCTTGTTGAAGGAAG
ACTCCATCTGATGACTCAGAGCAAGTATTTTTTAGTGTGTTATTGTTATTAGCAGAAAGAG
GGCCATAAAATACATGGGGCAAGCTGAATATATCTTAGGCAAAAGAAGAAAATATTCAAAT
TCTTATGTTATTTTATCTAATTATTTTATCTCTTTTTGTGTGTGACTTATAATGTGTGTAT
TGTATTAATAAAAGTATATAAACATGTAGTTT

FIGURA 73

CATTCCCACCCTTCCTTCTTTAATAAGCAGGAGCGAAAAAGACAAATTCCAAAGAGGATTG
TTCAGTTCAAGGGAATGAAGAATTCAGAATAATTTTGGTAAATGGATTCCAATATCGGGAA
TAAGAATAAGCTGAACAGTTGACCTGCTTTGAAGAAACATACTGTCCATTTGTCTAAAATA
ATCTATAACAACCAAACCAATCAAAATGAATTCAACATTATTTTCCCAGGTTGAAAATCAT
TCAGTCCACTCTAATTTCTCAGAGAAGAATGCCCAGCTTCTGGCTTTTGAAAATGATGATT
GTCATCTGCCCTTGGCCATGATATTTACCTTAGCTCTTGCTTATGGAGCTGTGATCATTCT
TGGTGTCTCTGGAAACCTGGCCTTGATCATAATCATCTTGAAACAAAAGGAGATGAGAAAT
GTTACCAACATCCTGATTGTGAACCTTTCCTTCTCAGACTTGCTTGTTGCCATCATGTGTC
TCCCCTTTACATTTGTCTACACATTAATGGACCACTGGGTCTTTGGTGAGGCGATGTGTAA
GTTGAATCCTTTTGTGCAATGTGTTTCAATCACTGTGTCCATTTTCTCTCTGGTTCTCATT
GCTGTGGAACGACATCAGCTGATAATCAACCCTCGAGGGTGGAGACCAAATAATAGACATG
CTTATGTAGGTATTGCTGTGATTTGGGTCCTTGCTGTGGCTTCTTCTTTGCCTTTCCTGAT
CTACCAAGTAATGACTGATGAGCCGTTCCAAAATGTAACACTTGATGCGTACAAAGACAAA
TACGTGTGCTTTGATCAATTTCCATCGGACTCTCATAGGTTGTCTTATACCACTCTCCTCT
TGGTGCTGCAGTATTTTGGTCCACTTTGTTTTATATTTATTTGCTACTTCAAGATATATAT
ACGCCTAAAAAGGAGAAACAACATGATGGACAAGATGAGAGACAATAAGTACAGGTCCAGT
GAAACCAAAAGAATCAATATCATGCTGCTCTCCATTGTGGTAGCATTTGCAGTCTGCTGGC
TCCCTCTTACCATCTTTAACACTGTGTTTGATTGGAATCATCAGATCATTGCTACCTGCAA
CCACAATCTGTTATTCCTGCTCTGCCACCTCACAGCAATGATATCCACTTGTGTCAACCCC
ATATTTTATGGGTTCCTGAACAAAAACTTCCAGAGAGACTTGCAGTTCTTCTTCAACTTTT
GTGATTTCCGGTCTCGGGATGATGATTATGAAACAATAGCCATGTCCACGATGCACACAGA
TGTTTCCAAAACTTCTTTGAAGCAAGCAAGCCCAGTCGCATTTAAAAAAATCAACAACAAT
GATGATAATGAAAAAATCTGAAACTACTTATAGCCTATGGTCCCGGATGACATCTGTTTAA
AAACAAGCACAACCTGCAACATACTTTGATTACCTGTTCTCCCAAGGAATGGGGTTGAAAT
CATTTGAAAATGACTAAGATTTTCTTGTCTTGCTTTTTACTGCTTTTGTTGTAGTTGTCAT
AATTACATTTGGAACAAAAGGTGTGGGCTTTGGGGTCTTCTGGAAATAGTTTTGACCAGAC
ATCTTTGAAGTGCTTTTTGTGAATTTATGCATATAATATAAAGACTTTTATACTGTACTTA
TTGGAATGAAATTTCTTTAAAGTATTACGATGCGCTGACTTCAGAAGTACCTGCCATCCAA
TACGGTCATTAGATTGGGTCATCTTGATTAGATTAGATTAGATTAGATTGTCAACAGATTG
GGCCATCCTTACTTTATGATAGGCATCATTTTAGTGTGTTACAATAGTAACAGTATGCAAA
AGCAGCATTCAGGAGCCGAAAGATAGTCTTGAAGTCATTCAGAAGTGGTTTGAGGTTTCTG
TTTTTTGGTGGTTTTTGTTTGTTTTTTTTTTTTTTCACCTTAAGGGAGGCTTTCATTTCCT
CCCGACTGATTGTCACTTAAATCAAAATTTAAAAATGAATAAAAAGACATACTTCTCAGCT
GCAAATATTATGGAGAATTGGGCACCCACAGGAATGAAGAGAGAAAGCAGCTCCCCAACTT

TTAGTATTGCTGCAAATAGCTAAATTATATTTATTTGAATTGATGGTCAAGAGATTTTCCA
TTTTTTTTACAGACTGTTCAGTGTTTGTCAAGCTTCTGGTCTAATATGTACTCGAAAGACT
TTCCGCTTACAATTTGTAGAAACACAAATATCGTTTTCCATACAGCAGTGCCTATATAGTG
ACTGATTTTAACTTTCAATGTCCATCTTTCAAAGGAAGTAACACCAAGGTACAATGTTAAA
GGAATATTCACTTTACCTAGCAGGGAAAAATACACAAAAACTGCAGATACTTCATATAGCC
CATTTTAACTTGTATAAACTGTGTGACTTGTGGCGTCTTATAAATAATGCACTGTAAAGAT
TACTGAATAGTTGTGTCATGTTAATGTGCCTAATTTCATGTATCTTGTAATCATGATTGAG
CCTCAGAATCATTTGGAGAAACTATATTTTAAAGAACAAGACATACTTCAATGTATTATAC
AGATAAAGTATTACATGTGTTTGATTTTAAAAGGGCGGACATTTTATTAAAATCAATATTG
TTTTTGCTTTTTCTGAGGAGTCTCTTTCAGTTTCATTTTTTCTCATCCCATGACTTCCCTC
CGATGGT

FIGURA 74

AGTATGAGGCGCCCACCATGGTGCGATGCGACCGCGGGCTGCAGATGCTGCTGACCACGGC
CGGAGCCTTCGCCGCCTTCTCGCTCATGGCCATCGCCATCGGCACCGACTACTGGCTGTAC
TCCAGCGCGCACATCTGCAACGGCACCAACCTGACCATGGACGACGGGCCCCCGCCCCGCC
GCGCCCGCGGCGACCTCACCCACTCTGGTCTGTGGCGGGTGTGCTGCATCGAAGGGATCTA
TAAAGGGCACTGCTTCCGGATCAATCACTTCCCAGAGGACAATGACTACGACCACGACAGC
TCGGAGTACCTCCTCCGCATCGTGCGAGCCTCCAGCGTCTTCCCCATCCTCAGCACCATCC
TGCTCCTGCTGGGTGGCCTGTGCATCGGTGCTGGCAGGATCTACAGCCGCAAGAACAACAT
CGTCCTCAGTGCCGGCATCCTCTTCGTGGCTGCAGGCCTCAGTAACATCATCGGTATCATC
GTCTACATTTCCAGCAACACAGGTGACCCGAGTGACAAGCGGGACGAAGACAAAAAGAACC
ATTACAACTACGGCTGGTCTTTTTACTTTGGAGCTCTGTCTTTCATTGTGGCTGAGACCGT
GGGCGTCCTGGCTGTAAACATTTACATTGAGAAAAATAAAGAGTTGAGGTTTAAGACCAAA
CGGGAATTCCTTAAGGCGTCTTCCTCTTCTCCTTATGCCAGGATGCCGAGCTACAGGTACC
GGCGACGGCGCTCGAGGTCCAGCTCAAGGTCCACCGAGGCCTCGCCCTCCAGGGACGTGTC
GCCCATGGGCCTGAAGATCACAGGGGCCATCCCCATGGGGGAGCTGTCCATGTACACGCTG
TCCAGGGAGCCCCTCAAGGTGACCACCGCAGCCAGCTACAGCCCCGACCAGGAGGCCAGCT
TCCTGCAGGTGCATGACTTTTTCCAGCAGGACCTGAAGGAAGGTTTCCACGTCAGCATGCT
GAACCGACGGACGACCCCTGTGTGA

FIGURA 75

CGGGTCTGATAGTCCCTACCTGTCAGGACTGGTGTTAGGATGAGATAATGTTTGTGAACTG
TAAACATATATAAACGTGTGCTACTGTGAGAACTGGAACAAAGAAGAGAGGGAGTGAGAGA
AATCAAGGGAGGGCTGGGGCTGGGAAAGAACGAAAAGGGAGTCGCGTATAGAGGAGAGGCG
ACAGTCGCGAGCCACACTTTGCAATGAAACTCTTTAGACTTTCTGCCGGGAGAGCGGCCCA
GACGCGCCAGGTCTGTAGCAGGAGGCCGCGAGGGCGGGTCCCCAGAAGCCTACAGGTGAGT
ATCGGTTCTCCCCTTCCCGGCTTTCGGTCCGGAGGAGGCGGGAGCAGCTTCCCTGTTCTGA
TCCTATCGCGGGCGGCGCAGGGCCGGCTTGGCCTTCCGTGGGACGGGGAGGGGGGCGGGAT
GTGTCACCCAAATACCAGTGGGGACGGTCGGTGGTGGAACCAGCCGGGCAGGTCGGGTAGA
GTATAAGAGCCGGAGGGAGCGGCCGGGGCGCAGACGCCTGCAGACCATCCCAGACGCCGGA
GCCCGAGCCCCGACGAGTCCCCGCGCCTCATCCGCCCGCGTCCGGTCCGCGTTCCTCCGCC
CCACCATGGCTCGGGGCCCCGGCCTCGCGCCGCCACCGCTGCGGCTGCCGCTGCTGCTGCT
GGTGCTGGCGGCGGTGACCGGCCACACGGCCGCGCAGGACAACTGCACGTGTCCCACCAAC
AAGATGACCGTGTGCAGCCCCGACGGCCCCGGCGGCCGCTGCCAGTGCCGCGCGCTGGGCT
CGGGCATGGCGGTCGACTGCTCCACGCTGACCTCCAAGTGTCTGCTGCTCAAGGCGCGCAT
GAGCGCCCCCAAGAACGCCCGCACGCTGGTGCGGCCGAGTGAGCACGCGCTCGTGGACAAC
GATGGCCTCTACGACCCCGACTGCGACCCCGAGGGCCGCTTCAAGGCGCGCCAGTGCAACC
AGACGTCGGTGTGCTGGTGCGTGAACTCGGTGGGCGTGCGCCGCACGGACAAGGGCGACCT
GAGCCTACGCTGCGATGAGCTGGTGCGCACCCACCACATCCTCATTGACCTGCGCCACCGC
CCCACCGCCGGCGCCTTCAACCACTCAGACCTGGACGCCGAGCTGAGGCGGCTCTTCCGCG
AGCGCTATCGGCTGCACCCCAAGTTCGTGGCGGCCGTGCACTACGAGCAGCCCACCATCCA
GATCGAGCTGCGGCAGAACACGTCTCAGAAGGCCGCCGGTGAAGTGGATATCGGCGATGCC
GCCTACTACTTCGAGAGGGACATCAAGGGCGAGTCTCTATTCCAGGGCCGCGGCGGCCTGG
ACTTGCGCGTGCGCGGAGAACCCCTGCAGGTGGAGCGCACGCTCATCTATTACCTGGACGA
GATTCCCCCGAAGTTCTCCATGAAGCGCCTCACCGCCGGCCTCATCGCCGTCATCGTGGTG
GTCGTGGTGGCCCTCGTCGCCGGCATGGCCGTCCTGGTGATCACCAACCGGAGAAAGTCGG
GGAAGTACAAGAAGGTGGAGATCAAGGAACTGGGGGAGTTGAGAAAGGAACCGAGCTTGTA
GGTACCCGGCGGGGCAGGGGATGGGGTGGGGTACCGGATTTCGGTATCGTCCCAGACCCAA
GTGAGTCACGCTTCCTGATTCCTCGGCGCAAAGGAGACGTTTATCCTTTCAAATTCCTGCC
TTCCCCCTCCCTTTTGCGCACACACCAGGTTTAATAGATCCTGGCCTCAGGGTCTCCTTTC
TTTCTCACTTCTGTCTTGAGGGAAGCATTTCTAAAATGTATCCCCTTTCGGTCCAACAACA
GGAAACCTGACTGGGGCAGTGAAGGAAGGGATGGCACAGCGTTATGTGTAAAAAACAAGTA
TCTGTATGACAACCCGGGATCGTTTGCAAGTAACTGAATCCATTGCGACATTGTGAAGGCT
TAAATGAGTTTAGATGGGAAATAGCGTTGTTATCGCCTTGGGTTTAAATTATTTGATGAGT
TCCACTTGTATCATGGCCTACCCGAGGAGAAGAGGAGTTTGTTAACTGGGCCTATGTAGTA
[illegible]TACTGACCACATATGCTTGTCACTGGGAAAGAAGCCTG

TTTCAGCTGCCTGAACGCAGTTTGGATGTCTTTGAGGACAGACATTGCCCGGAAACTCAGT
CTATTTATTCTTCAGCTTGCCCTTACTGCCACTGATATTGGTAATGTTCTTTTTTGTAAAA
TGTTTGTACATATGTTGTCTTTGATAATGTTGCTGTAATTTTTTAAAATAAAACACGAATT
TAATAAAATATGGGAAAGGCACAAACCAGAAGTCGGCATTTGTGAAAAGTCCCTCCAGATT
TCTATCACTTTGGTCTCTAATTTCCCAAGACTTGTATTTTTTTTTTATTTCAAATTATAAC
ACTTTTTTTTCCCCCAGAAGTGGGTGTTTCATGTTGCTACTCTGGTGTGTCCCAAGATATC
CTAACTGGCCAGTGTAAATGCTATTCTTTCTAAATAAGATTATTTGGAAACTTCCTTCAAA
CTGCAGGAGGGCGAGCTCTGAGGGCACGAGAAGCTAAAACTAGCTGCTTTTGATGAAAAAG
AGTGCCAGTCTTTGGTCATCTCTAAACAAGGCTTATCACCAATGGAGACAGAAAACTCTAG
TTCAAGAGCTGTACCTCCTTTGAATCCCAGCCCTACTCGAAATAAGTGGTACTATTTCCAT
TTAGCCTTTGAGCAAATCACTTAACTCAAAGGCGTTGTGGCTCTAAGATTAAACGACTTT

FIGURA 76

GCGGCCGCGGAGCAAGAAGGGCGCCGCGTCGTGCGGCCCGCGCAGCCCCCGGAGCCATGGG
CAAGTGCAGCGGGCGCTGCACGCTGGTCGCCTTCTGCTGCCTGCAGCTGGTGGCTGCGCTG
GAGCGGCAGATCTTTGACTTCCTGGGCTACCAGTGGGCTCCCATCCTAGCCAACTTCCTGC
ACATCATGGCAGTCATCCTGGGCATCTTTGGCACCGTGCAGTACCGCTCCCGGTACCTCAT
CCTGTATGCAGCCTGGCTGGTGCTCTGGGTTGGCTGGAATGCATTTATCATCTGCTTCTAC
TTGGAGGTTGGACAGCTGTCCCAGGACCGGGACTTCATCATGACCTTCAACACATCCCTGC
ACCGCTCCTGGTGGATGGAGAATGGGCCAGGCTGCCTGGTGACACCTGTTCTGAACTCCCG
CCTGGCTCTGGAGGACCACCATGTCATCTCTGTCACTGGCTGCCTGCTTGACTACCCCTAC
ATTGAAGCCCTCAGCAGCGCCCTGCAGATCTTCCTGGCACTGTTCGGCTTCGTGTTCGCCT
GCTACGTGAGCAAAGTGTTCCTGGAGGAGGAGGACAGCTTTGACTTCATCGGCGGCTTTGA
CTCCTACGGATACCAGGCGCCCCAGAAGACGTCGCATTTACAGCTGCAGCCTCTGTACACG
TCGGGGTAGCCTCTGCCCCGCGCCCACCCCGGCGCCTCGCCCTGGGCTGACCGCAGCTGCC
GCGAGCTCGGGCCAAGGCGCAGGCGTGTCCCCCTGGTGGCCCGCGCTCTCACTGCAGCCTG
TGCCCAACCCCGCGTCTGCATCTGGAGATGCGGACTTGGACGTGGACTTGGACTTGGACTT
GGATTTGAGCTTGGCTCTTCGCAGCCCGGACTTCGGAGGAGTGGGGCGGGGCGGGGGAGGG
GCACCACGGGTTTTTTGTTTTTTGTTTGTTTGTTTTTAATCTCAGCCTTGGCGTGAGCTGG
GGCCTTCCTCTCTTCTCCAGCCTCTCCCTTTCACTCTTCGCCCAGCATCCTGCCCCCCTGT
CCAAAAACAGCAGGACATCAGGCCCATCCCATCCCACCACACTCACTCACCAGCTCTGGGG
AAAGCTACTGTGAACTAGGAGCAGGATTCCTGGGTTCTAATCGCAGGTCCATCACTGACTG
TGACGTCTAGCAAAGCCCTTGCCCTCTCTGAGCCTCGGTTTCCGCACCTCAAGTAATTAAT
CCCTTAGCAAATGGACTCTTTTAGACTTCTCATTTAACTCAATTCCCTGAGCTAGACTGGG
ATTAAAATTCTCATTTTGCAGTACATTAAAACTGATGCCCAGAGATGTGATTTGCTTGAGG
CCACACAGCTAGATTTTTGGTGGAAGTGGGCCTTGAACACAGTGTACTTTCTGCAGTTTCT
GACTGTAAAACCCAGTGTCTGCTCTCTGAGTTCCATTTCCAAGCCCCCCTCCATCTTGGAC
CTATGTGGTCTCCACCATATTCACACACCACCACCACCACTTGCCAATGCCTCTCTTAAAG
CAATATACCCATTCGTTCTCTTATTGGGAACTGGATGGATGAAGCCCCAAATTCAGCCCCA
CCCACAGAGAAGCCTTCCTACACTCAGCCTCTGTCCACCCTTGGCAAATCTTTCAAGCTCT
CTCCTCCAGGAAAGTGGGGCCCCAACTCAGTCACTCCACCCCCTTCCAGGTCCCTGAGGCT
GGTTCTACTGTATCCCCATCACCTCCACAACTCCACTCACCCCTGACGGCTCCATCCACCT
CACCAGTTGGAAGGCTTGTGGTTTCAGAGAGGAGCAATGCTGGTCAGCGCTGCCCAGACTC
CAGTGTTTACAGATCACCAGCATTTACAACCAATCCAATGGCCAGAAGCCTCCTCTAACAA
GCCCAGAAGGAGTTCTGAAGGGGCAGATGGGGGTGTGAGTAGTCGGGGAGTCGGGATTGCC
AGCACCCTCACCCTTCCTTGGGGGCAAGTAGAGGTGAGAACACTTTCCCCACCTCCCTCCA
CAGACACTCCTGAGGACGCTGCATCCCACGCACTGCCTGGTGCGTCCATAGAGAGAGGATC
[illegible]

CATGGAGGAACTGGGGGGCACCCAGATGGTGCAGATGGTTTGCACACCTGAGCCTGTCTGT
GGTGACCATTCCGCTCCTCTCCCACTACCCTCCAATCTATCATTCCCTACTCTCTAAGGCC
AAAATATCCTGAGCAAGGCTGGCAACCCCACCCCACCATCCCAAATGCAAGCAGCCAGGCC
CAGGAGTTCCTCTGGCCCCCACAGGCATGGAGCTCCCAGCTGGTGGGTACAGCTTGAGAGG
GGGGCAGCTCCCTCAGGCTAAGCTACTGCCCTTCACTGGGCCAGCCCTGCCTCCAGCCCTC
ACCTCTCTCACCCCAACTGTCCCCCAAGCCCCTTTCTACTCAACGGGTGTAGCCACTGGTG
CTTTGAAGCCTTTTGTTTTTATAAGATGGTTTTTGCAAGGGGACCAGGTTCTCTTTTCACT
GGGACCTTGCAAGGAGGGGAGTGCTCTCCTGGTTTCTGTGCAGGCGGGTTGATTAAAGATG
GTGTTTTCTTCTCT

FIGURA 77

GGCGGGGAGAGGAGAGGAGAAGAGCCGCGGGGGGCCCAGCCCGGAGCCAGGATGCCCGC
GCCGCGCGCCCGGGAGCAGCCCCGCGTGCCCGGGGAGCGCCAGCCGCTGCTGCCTCGCGGT
GCGCGGGGCCCTCGACGGTGGCGGCGGGCGGCGGGCGCGGCCGTGCTGCTGGTGGAGATGC
TGGAGCGCGCCGCCTTCTTCGGCGTCACCGCCAACCTCGTGCTGTACCTCAACAGCACCAA
CTTCAACTGGACCGGCGAGCAGGCGACGCGCGCCGCGCTGGTATTCCTGGGCGCCTCCTAC
CTGCTGGCGCCCGTGGGCGGCTGGCTGGCCGACGTGTACCTGGGCCGCTACCGCGCGGTCG
CGCTCAGCCTGCTGCTCTACCTGGCCGCCTCGGGCCTGCTGCCCGCCACCGCCTTCCCCGA
CGGCCGCAGCTCCTTCTGCGGAGAGATGCCCGCGTCGCCGCTGGGACCTGCCTGCCCCTCG
GCCGGCTGCCCGCGCTCCTCGCCCAGCCCCTACTGCGCGCCCGTCCTCTACGCGGGCCTGC
TGCTACTCGGCCTGGCCGCCAGCTCCGTCCGGAGCAACCTCACCTCCTTCGGTGCCGACCA
GGTGATGGATCTCGGCCGCGACGCCACCCGCCGCTTCTTCAACTGGTTTTACTGGAGCATC
AACCTGGGTGCTGTGCTGTCGCTGCTGGTGGTGGCGTTTATTCAGCAGAACATCAGCTTCC
TGCTGGGCTACAGCATCCCTGTGGGCTGTGTGGGCCTGGCATTTTTCATCTTCCTCTTTGC
CACCCCCGTCTTCATCACCAAGCCCCCGATGGGCAGCCAAGTGTCCTCTATGCTTAAGCTC
GCTCTCCAAAACTGCTGCCCCCAGCTGTGGCAACGACACTCGGCCAGAGACCGTCAATGTG
CCCGCGTGCTGGCCGACGAGAGGTCTCCCCAGCCAGGGGCTTCCCCGCAAGAGGACATCGC
CAACTTCCAGGTGCTGGTGAAGATCTTGCCCGTCATGGTGACCCTGGTGCCCTACTGGATG
GTCTACTTCCAGATGCAGTCCACCTATGTCCTGCAGGGTCTTCACCTCCACATCCCAAACA
TTTTCCCAGCCAACCCGGCCAACATCTCTGTGGCCCTGAGAGCCCAGGGCAGCAGCTACAC
GATCCCGGAAGCCTGGCTCCTCCTGGCCAATGTTGTGGTGGTGCTGATTCTGGTCCCTCTG
AAGGACCGCTTGATCGACCCTTTACTGCTGCGGTGCAAGCTGCTTCCCTCTGCTCTGCAGA
AGATGGCGCTGGGGATGTTCTTTGGTTTTACCTCCGTCATTGTGGCAGGAGTCCTGGAGAT
GGAGCGCTTACACTACATCCACCACAACGAGACCGTGTCCCAGCAGATTGGGGAGGTCCTG
TACAACGCGGCACCACTGTCCATCTGGTGGCAGATCCCTCAGTACCTGCTCATTGGGATCA
GTGAGATCTTTGCCAGCATCCCAGGCCTGGAGTTTGCCTACTCAGAGGCCCCGCGCTCCAT
GCAGGGCGCCATCATGGGCATCTTCTTCTGCCTGTCGGGGGTGGGCTCACTGTTGGGCTCC
AGCCTAGTGGCACTGCTGTCCTTGCCCGGGGGCTGGCTGCACTGCCCCAAGGACTTTGGGA
ACATCAACAATTGCCGGATGGACCTCTACTTCTTCCTGCTGGCTGGCATTCAGGCCGTCAC
GGCTCTCCTATTTGTCTGGATCGCTGGACGCTATGAGAGGGCGTCCCAGGGCCCAGCCTCC
CACAGCCGTTTCAGCAGGGACAGGGGCTGAACAGGCCCTATTCCAGCCCCCTTGCTTCAC

FIGURA 78A

GGCCGCGGGCTCTCGCGGGGCGGCGACGCCGCGGGGAGGATGCTGCTTGCCGCGCCCGCGT
CCTCACCGTCCTCCCGGGCCGCCTGCTGGGGCTTTGTTGTGGCCCGGACGCCGCGGGCCAC
CCCCTGAAGTCGCCTGCCGCCGCCGCCGCCGCACCTAGCGGACGGGCGGGCGGGCGCGCGT
GTGCCCAGGAGTGCGCGCCTGTCGCGGTGGTGGGTGCAGGACTGGACCCACGGGCCCATTG
TGCGCCCGCCCGCGGCAGCCAGGACCATGTGGGTGAACCCGGAGGAGGTGTTGCTGGCCAA
CGCGCTGTGGATCACCGAGAGGGCCAACCCATACTTCATCCTGCAGCGGAGGAAGGGCCAC
GCCGGCGATGGAGGCGGCGGCGGCGGACTGGCGGGCCTGCTGGTGGGTACCCTTGATGTTG
TGTTGGACTCCAGCGCCCCGGGTCGCTCCTTACCGAATCTTGTACCAGACTCCAGACTCCCT
GGTCTACTGGACCATCGCCTGTGGTGGTTCCAGGAAAGAAATCACTGAACACTGGGAATGG
CTTGAGCAAAATCTCTTGCAGACACTCTCCATCTTTGAAAATGAGAATGATATCACCACAT
TTGTGAGAGGAAAAATACAGGGCATCATTGCAGAATACAACAAAATCAATGATGTAAAGGA
AGATGATGACACGGAGAAGTTTAAAGAAGCCATTGTGAAATTTCATAGGCTGTTTGGGATG
CCAGAGGAAGAGAAACTCGTCAACTATTACTCTTGCAGCTATTGGAAGGGGAAGGTCCCCC
GTCAGGGTTGGATGTACCTCAGCATTAACCACCTTTGCTTTTATTCTTTTCTTATGGGAAG
GGAAGCGAAACTGGTCATCCGGTGGGTAGACATCACTCAGCTTGAGAAGAATGCCACCCTG
CTTCTGCCTGATGTGATCAAAGTGAGCACACGGTCCAGTGAGCATTTCTTCTCTGTATTCC
TCAACATCAACGAGACCTTCAAGTTAATGGAGCAGCTTGCCAACATAGCCATGAGGCAACT
CTTAGACAATGAGGGATTTGAACAAGATCGATCCCTGCCCAAACTCAAAAGGAAATCTCCT
AAAAAAGTGTCTGCTCTAAAACGTGATCTTGATGCCAGGGCAAAGAGTGAGAGATACCGTG
CACTTTTCCGGCTGCCCAAAGATGAAAAATTAGATGGCCACACAGACTGCACTCTCTGGAC
TCCATTTAACAAAATGCACATTTTGGGGCAGATGTTTGTGTCCACAAATTACATCTGTTTT
ACCAGCAAGGAGGAGAACTTATGTAGCCTCATTATCCCGCTCCGTGAGGTGACAATTGTGG
AAAAGGCAGACAGCTCCAGTGTGCTCCCCAGTCCCTTATCCATCAGCACCCGAAACAGGAT
GACCTTCCTATTTGCCAACTTGAAAGATAGAGACTTTCTAGTGCAGAGGATCTCAGATTTC
CTGCAACAGACTACTTCCAAAATATATTCTGACAAGGAGTTTGCAGGAAGTTACAACAGTT
CAGATGATGAGGTGTACTCTCGACCCAGCAGCCTCGTCTCCTCCAGCCCCCAGAGAAGCAC
GAGCTCTGATGCTGATGGAGAGCGCCAGTTTAACCTAAATGGCAACAGCGTCCCCACAGCC
ACACAGACCCTGATGACCATGTATCGGCGGCGGTCTCCCGAGGAGTTCAACCCGAAATTGG
CCAAAGAGTTTCTGAAAGAGCAAGCCTGGAAGATTCACTTTGCTGAGTATGGGCAAGGGAT
CTGCATGTACCGCACAGAGAAAACGCGGGAGCTGGTGTTGAAGGGCATCCCGGAGAGCATG
CGTGGGGAGCTCTGGCTGCTGCTGTCAGGTGCCATCAATGAGAAGGCCACACATCCTGGGT
ACTATGAAGACCTAGTGGAGAAGTCCATGGGGAAGTATAATCTCGCCACGGAGGAGATTGA
GAGGGATTTACACCGCTCCCTTCCAGAACACCCAGCTTTTCAGAATGAAATGGGCATTGCT
GCACTAAGGAGAGTCTTAACAGCTTATGCTTTTCGAAATCCCAACATAGGGTATTGCCAGG
CCATGAATATTGTCACTTCAGTGCTGCTGCTTTATGCCAAAGAGGAGGAAGCTTTCTGGCT

GCTTGTGGCTTTGTGTGAGCGCATGCTCCCAGATTACTACAACACCAGAGTTGTGGGTGCA
CTGGTGGACCAAGGTGTCTTTGAGGAGCTAGCACGAGACTACGTCCCACAGCTGTACGACT
GCATGCAAGACCTGGGCGTGATTTCCACCATCTCCCTGTCTTGGTTCCTCACACTATTTCT
CAGTGTGATGCCTTTTGAGAGTGCAGTTGTGGTTGTTGACTGTTTCTTCTATGAAGGAATT
AAAGTGATATTCCAGTTGGCCCTAGCTGTGCTGGATGCAAATGTGGACAAACTGTTGAACT
GCAAGGATGATGGGGAGGCCATGACCGTTTTGGGAAGGTATTTAGACAGTGTGACCAATAA
AGACAGCACACTGCCTCCCATTCCTCACCTCCACTCCTTGCTCAGCGATGATGTGGAACCT
TACCCTGAGGTAGACATCTTTAGACTCATCAGAACTTCCTACGAGAAATTCGGAACTATCC
GGGCAGATTTGATTGAACAGATGAGATTCAAACAGAGACTGAAAGTGATCCAGACGCTGGA
GGATACTACGAAACGCAACGTGGTACGAACCATTGTGACAGAAACTTCCTTTACCATTGAT
GAGCTGGAAGAACTTTATGCTCTTTTCAAGGCAGAACATCTCACCAGCTGCTACTGGGGCG
GGAGCAGCAACGCGCTGGACCGGCATGACCCCAGCCTGCCCTACCTGGAACAGTATCGCAT
TGACTTCGAGCAGTTCAAGGGAATGTTTGCTCTTCTCTTTCCTTGGGCATGTGGAACTCAC
TCTGACGTTCTGGCCTCCCGCTTGTTCCAGTTATTAGATGAAAATGGAGACTCTTTGATTA
ACTTCCGGGAGTTTGTCTCTGGGCTAAGTGCTGCATGCCATGGGGACCTCACAGAGAAGCT
CAAACTCCTGTACAAAATGCACGTCTTGCCTGAGCCATCCTCTGATCAAGATGAACCAGAT
TCTGCTTTTGAAGCAACTCAGTACTTCTTTGAAGATATTACCCCAGAATGTACACATGTTG
TTGGATTGGATAGCAGAAGCAAACAGGGTGCAGATGATGGCTTTGTTACGGTGAGCCTAAA
GCCAGACAAAGGGAAGAGAGCAAATTCCCAAGAAAATCGTAATTATTTGAGACTGTGGACT
CCAGAAAATAAATCTAAGTCAAAGAATGCAAAGGATTTACCCAAATTAAATCAGGGGCAGT
TCATTGAACTGTGTAAGACAATGTATAACATGTTCAGCGAAGACCCCAATGAGCAGGAGCT
GTACCATGCCACGGCAGCAGTGACCAGCCTCCTGCTGGAGATTGGGGAGGTCGGCAAGTTG
TTCGTGGCCCAGCCTGCAAAGGAGGGCGGGAGCGGAGGCAGTGGGCCGTCCTGCCACCAGG
GCATCCCAGGCGTGCTCTTCCCCAAGAAAGGGCCAGGCCAGCCTTACGTGGTGGAGTCTGT
TGAGCCCCTGCCGGCCAGCCTGGCCCCCGACAGCGAGGAACACTCCCTTGGAGGACAAATG
GAGGACATCAAGCTGGAGGACTCCTCGCCCCGGGACAACGGGGCCTGCTCCTCCATGCTGA
TCTCTGACGACGACACCAAGGACGACAGCTCCATGTCCTCATACTCGGTGCTGAGTGCCGG
CTCCCACGAGGAGGACAAGCTGCACTGCGAGGACATCGGAGAGGACACGGTCCTGGTGCGG
AGCGGCCAGGGCACGGCGGCACTGCCCCGGAGCACCAGCCTGGACCGGGACTGGGCCATCA
CCTTCGAGCAGTTCCTGGCCTCCCTCTTAACTGAGCCTGCCCTGGTCAAGTACTTTGACAA
GCCCCGTGTGCATGATGGCCAGGATTACCAGTGCAAAAAACATCCGGATGATGGGCAAGCCC
CTCACCTCGGCCAGTGACTATGAAATCTCGGCCATGTCCGGCTGACACGGGCGCCTTCCCG
GGGGAGTGGGAGGAGAGGGAGGGGAGGGATTTTTTATGTTCTTCTGTGTTGAGTTTTTTCT
TTCTTTCTTTTAAATTAAATATTTATTAGTACCTGGCTTGAAGCCTAGTGTTTTCATAATG
TAATTCAATGAAAACTGTTGGAGAAATATTTAAACACCTCAATGTAGGTACATTACACTCT
TGTTGCGGGGAGGGGATTTACCAGAATACAGTTTATTTCGTGAATTCTAAAAAACAAAAAG

ATGGTTTAAAAAATAGTACTGTTTAATGCTAAGTAAGGCAGCAGTCATTTGTGTATTCAGG
CTTTTTAAATAAAATTAGAGCTGTAAGGAAAATGAAAAGCCACAAATGCAAGACTGTTCTT

FIGURA 78B

AAATGGAAGGCATAGTCAGCGAGGGTAAATCCTATACCACTTTAGGAAGTATTAAAAATAT
TTTTAAGATTTGAAATATATTTCATAGAAGTCCTCTATTCAAAATCATATTCCACAGATGT
TCCCCTTCAAAGGGAAAACATTTGGGGTTCTAAACAGTTATGAAAGTAAGTGATTTTTACA
TGATTCCAGAATAACACTTGTATTGACCAATTTAGACAGATACCAGACCAATTTTGCATTT
AAGAAATTGTTCTGATTATTTACGTCAACTCATTAGAATTCAGTGAAAAGTAACAGTCTTT
TGTCACAGAGAATCTGAAAGTAGCAGCAAAGACAGAGGGCTCATGACAGGTTTTTGCTTTT
GCTTTGCTTTTGTTTTTGAAAGAGTAAAAGTACTGATGCTTCTGATACTGGATGTTTAGCT
TCTTACTGCAAAAACATAAGTAAAACAGTCAACTTTACCATTTCCGTATTCTCCATAGATT
GAAGAAATTTATACCACATATCGCATATGACCATCTTTCCATCAAATCAATGTAGAGATAA
TGTAAACTGAAAAAAAATCTGCAAGATAATGTAACTGAATGTTTTAAAAACAGAACTTGTC
ACTTTATATAAAAGAATAGTATGCTCTATTTCCTGAATGGATGTGGAAATGAAAGCTAGCG
CACCTGCACTTTGAATTCTTGCTTCTTTTTTATTACTGTTATGATTTTGCTTTTTACAGAT
GTTGGACGATTTTTTCTTCTGATTGTTGAATTCATAATCATGGTCTCATTTCCTTTGCTTC
TTTGGAATATTTCTTTCAACACATTCCTTTATTTTATTATACATTGTGTCCTTTTTTTAGC
TATTGCTGCTGTTGTTTTTTATTCTATTTACAGGATGATTTTTAAACTGTCAAATGAAGTA
GTGTTAACCTCAAATAGGCTAAATGTGAACAAATAAAATACAGCAAATACTCAGATACAGC
TTTTTATCTTTGTGCTTGAGTTCCTGCCTAAGGAATAACATTATTCTTTTGACAACTTTTG
CAGGGGAAATTATATCAGGCAACCATTTTGATTAAGTAAATAAATTTTATAGGCAAACATA
TAGAGAGANATACAATTTGTAGTATATCAATGACTATATTTAAAATAAGGANTATAATTGT
TATCAGTTATCTAACTTAAAATGCTTATCCATAATGATCAGTGATATTCAGCTTTTTAAAA
TATGCTTGTTGG

941

FIGURA 79

MRAAYLFLLFLPAGLLAQGQYDLDPLPPFPDHVQYTHYSDQIDNPDYYDYQEVTPRPSEEQ
FQFQSQQQVQQEVIPAPTPEPGNAELEPTEPGPLDCREEQYPCTRLYSIHRPCKQCLNEVC
FYSLRRVYVINKEICVRTVCAHEELLRADLCRDKFSKCGVMASSGLCQSVAASCARSCGSC

Secuencia de señal
Aminoácidos 1-19

Sitios de N-miristoilación
Aminoácidos 161-166, 167-172

942

FIGURA 80

MAARPLPVSPARALLLALAGALLAPCEARGVSLMNEGRADEVVSASVRSGDLNIPVKSFDS
KNHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTVILGHCYYH
GHVRGYSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFPAKKLKSVRGSCGSHH
NTPNLAAKNVFPPPSQTWARRHKRETLKATKYVELVIVADNREFQRQGKDLEKVRQRLIEI
ANHVDKFYRPLNIRIVLVGVEVWNDMDKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNAQL
VSGVYFQGTTIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNFGMNHDTLDRG
CSCQMAVEKGGCIMNASTGYPFPMVFSSCSRKDLETSLEKGMGVCLFNLPEVRESFGGQKC
GNRFVEEGEECDCGEPEECMNRCCNATTCTLKPDAVCAHGLCCEDCQLKPAGTACRDSSNS
CDLPEFCTGASPHCPANVYLHDGHSCQDVDGYCYNGICQTHEQQCVTLWGPGAKPAPGICF
ERVNSAGDPYGNCGKVSKSSFAKCEMRDAKCGKIQCQGGASRPVIGTNAVSIETNIPLQQG
GRILCRGTHVYLGDDMPDPGLVLAGTKCADGKICLNRQCQNISVFGVHECAMQCHGRGVCN
NRKNCHCEAHWAPPFCDKFGFGGSTDSGPIRQADNQGLTIGILVTILCLLAAGFVVYLKRK
TLIRLLFTNKKTTIEKLRCVRPSRPFRGFQPCQAHLGHLGKGLMRKPPDSYPPKDNPRRLL
QCQNVDISRPLEGLNVPQPQSTQRVLPPLHRAPRAPSVPARPLPAKPALRQAQGTCKPNPP
QKPLPADPLARTTRLTHALARTPGQWETGLRLAPLRPAPQYPHQVPRSTHTAYIK

Secuencia de señal.
aminoácidos 1-28

Dominio transmembrana.
aminoácidos 709-729

Sitios de N-glicosilación.
aminoácidos 111-114, 149-152, 381-384, 452-455, 651-654

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.
aminoácidos 206-209, 730-733, 742-745

Sitios de N-miristoilación.
aminoácidos 88-93, 103-108, 177-182, 180-185, 313-318, 343-348, 377-382, 407-412, 423-428, 424-429, 524-529, 630-635, 635-640, 656-661, 668-673, 694-699, 708-713, 712-717, 878-883

Señal
aminoácidos 347-356

Proteínas desintegrinas
aminoácidos 454-504, 410-460, 486-536

FIGURA 81

MASRSMRLLLLLSCLAKTGVLGDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDAELECQSY
GNGAHLASILSLKEASTIAEYISGYQRSQPIWIGLHDPQKRQQWQWIDGAMYLYRSWSGKS
MGGNKHCAEISSNNNNFLTWSSNECNKRQHFLCKYRP

Secuencia de señal.
aminoácidos 1-22

Sitio de N-glicosilación.
aminoácidos 50-53

Sitios de N-miristoilación.
aminoácidos 64-69, 125-130

Señal de dominio de lectina tipo C.
aminoácidos 129-154

Dominio tipo lectina C.
aminoácidos 47-156

FIGURA 82

MATMENKVICALVLVSMLALGTLAEAQTETCTVAPRERQNCGFPGVTPSQCANKGCCFDDT
VRGVPWCFYPNTIDVPPEEECEF

Secuencia de señal.
aminoácidos 1-24

Sitios de N-miristoilación.
aminoácidos 45-50, 64-69

Dominio trefoil (tipo P).
aminoácidos 30-71

FIGURA 83

MLPQIPFLLLVSLNLVHGVFYAERYQTPTGIKGPLPNTKTQFFIPYTIKSKGIAVRGEQGT
PGPPGPAGPRGHPGPSGPPGKPGYGSPGLQGEPGLPGPPGPSAVGKPGVPGLPGKPGERGP
YGPKGDVGPAGLPGPRGPPGPPGIPGPAGISVPGKPGQQGPTGAPGPRGFPGEKGAPGVPG
MNGQKGEMGYGAPGRPGERGLPGPQGPTGPSGPPGVGKRGEMGVPGQPGIKGDRGFPGEMG
PIGPPGPQGPPGERGPEGIGKPGAAGAPGQPGIPGTKGLPGAPGIAGPPGPPGFGKPGLPG
LKGERGPAGLPGGPGAKGEQGPAGLPGKPGLTGPPGNMGPQGPKGIPGSHGLPGPKGETGP
AGPAGYPGAKGERGSPGSDGKPGYPGKPGLDGPKGMPGLPGPKGDPGVGGPPGLPGPVGPA
GAKGMPGHNGEAGPRGAPGIPGTRGPIGPPGIPGFPGSKGDPGSPGPPGPAGIATKGLNGP
TGPPGPPGPRGPSGEPGLPGPPGPPGPPGQAVMPEGFIKAGQRPSLSGTPLVSANQGVTGM
PVSAFTVILSKAYPAIGTPIPFDKILYNRQQHYDPRTGIFTCQIPGIYYFSYHVHVKGTHV
WVGLYKNGTPVMYTYDEYTKGYLDQASGSAIIDLTENDQVWLQLPNAESNGLYSSEYVHSS
FSGFLVAPM

Secuencia de señal.
aminoácidos 1-18

Sitio de fosforilación tirosina cinasa.
aminoácidos 116-123

Sitios de N-miristoilación.
aminoácidos 18-23, 276-281, 317-322, 350-355, 380-385, 415-420, 446-451, 529-534, 548-553, 587-592, 613-618, 661-666

Sitio de amidación.
aminoácidos 219-222

Señal de dominio clq.
aminoácidos 571-601

Dominio clq.
aminoácidos 553-677

Repetición de triple hélice de colágeno (20 copias).
aminoácidos 92-150, 156-214, 223-281, 282-340, 344-403, 404-463, 464-522

FIGURA 84

MPLPPLLLLLLAAPWGRAVPCVSGGLPKPANITFLSINMKNVLQWTPPEGLQGVKVTYTVQ
YFIYGQKKWLNKSECRNINRTYCDLSAETSDYEHQYYAKVKAIWGTKCSKWAESGRFYPFL
ETQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIYSNLKYNVSVLNTKSNRTW
SQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLKDQSSEFKAKIIFWY
VLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILIYGNEFDKRFFVPAEKIVINFITLNI
SDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFCDSEENT
EGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLQEEVSTQGTLLESQAAL
AVLGPQTLQYSYTPQLQDLDPLAQEETDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQ
DSEGCEPSEGDGLGEEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQMEN

Secuencia de señal.
aminoácidos 1-18

Dominio transmembrana.
aminoácidos 239-259

Sitios de N-glicosilación.
aminoácidos 31-34, 72-75, 80-83, 171-174, 180-183, 189-192, 304-307, -523-526 .

Sitio de fosforilación tirosina cinasas.
aminoácidos 385-392, 518-526

Sitios de N-miristoilación.
aminoácidos 53-58, 106-111, 368-373, 492-497

Proteínas 3-fosfatidilinositol y 4-cinasa.
aminoácidos 253-279

Factor de tejido.
aminoácidos 1-278

FIGURA 85

MRAPGRPALRPLPLPPLLLLLLAAPWGRAVPCVSGGLPKPANITFLSINMKNVLQWTPPEG
LQGVKVTYTVQYFIYGQKKWLNKSECRNINRTYCDLSAETSDYEHQYYAKVKAIWGTKCSK
WAESGRFYPFLETQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIYSNLKYNV
SVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLKDQS
SEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILIYGNEFDKRFFVPAE
KIVINFITLNISDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHL
MEIFCDSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQKLSLQEEVST
QGTLLESQAALAVLGPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEEEPSTTLVDWDPQTGR
LCIPSLSSFDQDSEGCEPSEGDGLGEEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYV
QMEN

Secuencia de señal.

aminoácidos 1-29

Dominio transmembrana.

aminoácidos 240-270

FIGURA 86

MAAARLCLSLLLLSTCVALLLQPLLGAQGAPLEPVYPGDNATPEQMAQYAADLRRYINMLT
RPRYGKRHKEDTLAFSEWGSPHAAVPRELSPLDL

Secuencia de señal.
aminoácidos 1-29

Sitio de N-miristoilación.
aminoácidos 80-85

Sitio de amidación.
aminoácidos 65-68

Señal de la familia de hormona pancreática.
aminoácidos 49-65

Péptido de hormona pancreática.
aminoácidos 30-65

FIGURA 87

MGPPSAPPCRLHVPWKEVLLTASLLTFWNPPTTAKLTIESTPFNVAEGKEVLLLAHNLPQN
RIGYSWYKGERVDGNSLIVGYVIGTQQATPGPAYSGRETIYPNASLLIQNVTQNDTGFYTL
QVIKSDLVNEEATGQFHVYPELPKPSISSNNSNPVEDKDAVAFTCEPEVQNTTYLWWVNGQ
SLPVSPRLQLSNGNMTLTLLSVKRNDAGSYECEIQNPASANRSDPVTLNVLYGPDGPTISP
SKANYRPGENLNLSCHAASNPPAQYSWFINGTFQQSTQELFIPNITVNNSGSYMCQAHNSA
TGLNRTTVTMITVSGSAPVLSAVATVGITIGVLARVALI

Secuencia de señal.

aminoácidos 1-34

Dominio transmembrana.

aminoácidos 322-342

Sitios de N-glicosilación.

aminoácidos 104-107, 111-114, 115-119, 152-155, 173-176, 197-200, 224-227, 256-259, 274-277, 288-291, 292-295, 309-312

Sitio de fosforilación tirosina cinasa.

aminoácidos 206-213

Sitios de N-miristoilación.

aminoácidos 85-90, 211-216, 295-300, 307-312, 332-337

Dominio de inmunoglobulina.

aminoácidos 160-217, 252-301

FIGURA 88

MELRVLLCWASLAAALEETLLNTKLETADLKWVTFPQVDGQWEELSGLDEEQHSVRTYEVC
DVQRAPGQAHWLRTGWVPRRGAVHVYATLRFTMLECLSLPRAGRSCKETFTVFYYESDADT
ATALTPAWMENPYIKVDTVAAEHLTRKRPGAEATGKVNVKTLRLGPLSKAGFYLAFQDQGA
CMALLSLHLFYKKCAQLTVNLTRFPETVPRELVVPVAGSCVVDAVPAPGPSPSLYCREDGQ
WAEQPVTGCSCAPGFEAAEGNTKCRACAQGTFKPLSGEGSCQPCPANSHSNTIGSAVCQCR
VGYFRARTDPRGAPCTTPPSAPRSVVSRLNGSSLHLEWSAPLESGGREDLTYALRCRECRP
GGSCAPCGGDLTFDPGPRDLVEPWVVVRGLRPDFTYTFEVTALNGVSSLATGPVPFEPVNV
TTDREVPPAVSDIRVTRSSPSSLSLAWAVPRAPSGAVLDYEVKYHEKGAEGPSSVRFLKTS
ENRAELRGLKRGASYLVQVRARSEAGYGPFGQEHHSQTQLDESEGWREQLALIAGTAVVGV
VLVLVVIVVAVLCLRKQSNGREAEYSDKHGQYLIGHGTKVYIDPFTYEDPNEAVREFAKEI
DVSYVKIEEVIGAGEFGEVCRGRLKAPGKKESCVAIKTLKGGYTERQRREFLSEASIMGQF
EHPNIIRLEGVVTNSMPVMILTEFMENGALDSFLRLNDGQFTVIQLVGMLRGIASGMRYLA
EMSYVHRDLAARNILVNSNLVCKVSDFGLSRFLEENSSDPTYTSSLGGKIPIRWTAPEAIA
FRKFTSASDAWSYGIVMWEVMSFGERPYWDMSNQDVINAIEQDYRLPPPPDCPTSLHQLML
DCWQKDRNARPRFPQVVSALDKMIRNPASLKIVARENGGASHPLLDQRQPHYSAFGSVGEW
LRAIKMGRYEESFAAAGFGSFELVSQISAEDLLRIGVTLAGHQKKILASVQHMKSQAKPGT
PGGTGGPAPQY

Secuencia de señal.

aminoácidos 1-15

Dominio transmembrana.

aminoácidos 539-559

Sitios de N-glicosilación.

aminoácidos 203-206, 335-338, 426-429, 768-771

Sitio de fijación de glicosaminoglicano.

aminoácidos 280-283

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasas.

aminoácidos 564-567, 639-642, 795-798

Sitio de fosforilación tirosina cinasas.

aminoácidos 508-515, 588-596, 608-614, 729-736

Sitios de N-miristoilación.

aminoácidos 152-157, 182-187, 264-269, 298-303, 317-322, 374-379, 496-501, 681-686, 699-704, 719-724, 723-728,

951-956, 975-980

Sitio de amidación.

aminoácidos 637-640

Señal de sitio activo específico de tirosina proteína cinasas.

aminoácidos 736-748

FIGURA 89

MAQLFLPLLAALVLAQAPAALADVLEGDSSEDRAFRVRIAGDAPLQGVLGGALTIPCHVHY
LRPPPSRRAVLGSPRVKWTFLSRGREAEVLVARGVRVKVNEAYRFRVALPAYPASLTDVSL
ALSELRPNDSGIYRCEVQHGIDDSSDAVEVKVKGVVFLYREGSARYAFSFSGAQEACARIG
AHIATPEQLYAAYLGGYEQCDAGWLSDQTVRYPIQTPREACYGDMDGFPGVRNYGVVDPDD
LYDVYCYAEDLNGELFLGDPPEKLTLEEARAYCQERGAEIATTGQLYAAWDGGLDHCSPGW
LADGSVRYPIVTPSQRCGGGLPGVKTLFLFPNQTGFPNKHSRFNVYCFRDSAQPSAIPEAS
NPASNPASDGLEAIVTVTETLEELQLPQEATESESRGAIYSIPIMEDGGGGSSTPEDPAEA
PRTLLEFETQSMVPPTGFSEEEGKALEEEEKYEDEEEKEEEEEEEFVEDEALMANPSELSS
PGPEASLPTSPAAQEKSLSQAPARAVLQPGASPLPDGESEASRPPRVHGPPTETLPTPRER
NLASPSPSTLVEAREVGEATGGPELSGVPRGESEETGSSEGAPSLLPATRAPEGTRELEAP
SEDNSGRTAPAGTSVQAQPVLPTDSASRGGVAVVPASGDCVPSPCHNGGTCLEEEEGVRCL
CLPGYGGDLCDVGLRFCNPGWDAFQGACYKHFSTRRSWEEAETQCRMYGAHLASISTPEEQ
DFINNRYREYQWIGLNDRTIEGDFLWSDGVPLLYENWNPGQPDSYFLSGENCVVMVWHDQG
QWSDVPCNYHLSYTCKMGLVSCGPPPELPLAQVFGRPRLRYEVDTVLRYRCREGLAQRNLP
LIRCQENGRWEAPQISCVPRRPARALHPEEDPEGRQGRLLGRWKALLIPPSSPMPGP

Señal peptídica.

aminoácidos 1-15

Sitios de N-glicosilación.

aminoácidos 130-133, 337-340

Sitio de fosforilación tirosina cinasas.

aminoácidos 128-135, 451-459

Sitios de N-miristoilación.

aminoácidos 47-52, 50-55, 133-138, 142-147, 174-179, 183-188, 281-286, 288-293, 297-302, 324-329, 403-408, 414-419, 415-420, 576-581, 586-591, 677-682, 684-689, 720-725, 772-777, 811-816

Señal de patrón de dominio cisteína tipo EGF.

aminoácidos 670-682

Señal de dominio de lectina tipo C.

aminoácidos 784-809

Señal de inmunoglobulinas y complejos de proteínas de alta histocompatibilidad.

aminoácidos 135-142

Proteínas de ligamiento de dominios.

aminoácidos 166-216, 264-314

Patrones de proteínas de proteínas con fijación de calcio de dominio tipo EGF.

aminoácidos 655-676

FIGURA 90

MAQLFLPLLAALVLAQAPAALADVLEGDSSEDRAFRVRIAGDAPLQGVLGGALTIPCHVHY
LRPPPSRRAVLGSPRVKWTFLSRGREAEVLVARGVRVKVNEAYRFRVALPAYPASLTDVSL
ALSELRPNDSGIYRCEVQHGIDDSSDAVEVKVKGVVFLYREGSARYAFSFSGAQEACARIG
AHIATPEQLYAAYLGGYEQCDAGWLSDQTVRYPIQTPREACYGDMDGFPGVRNYGVVDPDD
LYDVYCYAEDLNGELFLGDPPEKLTLEEARAYCQERGAEIATTGQLYAAWDGGLDHCSPGW
LADGSVRYPIVTPSQRCGGGLPGVKTLFLFPNQTGFPNKHSRFNVYCFRDSAQPSAIPEAS
NPASNPASDGLEAIVTVTETLEELQLPQEATESESRGAIYSIPIMEDGGGGSSTPEDPAEA
PRTLLEFETQSMVPPTGFSEEEGKALEEEEKYEDEEEKEEEEEEEEVEDEALWAWPSELSS
PGPEASLPTEPAAQEESLSQAPARAVLQPGASPLPDGESEASRPPRVHGPPTETLPTPRER
NLASPSPSTLVEAREVGEATGGPELSGVPRGESEETGSSEGAPSLLPATRAPEGTRELEAP
SEDNSGRTAPAGTSVQAQPVLPTDSASRGGVAVVPASGNSAQGSTALSILLLFFPLQLWVT

Secuencia de señal.
aminoácidos 1-15

Dominio transmembrana.
aminoácidos 652-670

Sitios de N-glicosilación.
aminoácidos 130-133, 337-340

Sitio de fosforilación tirosina cinasas.
aminoácidos 128-135, 451-459

Sitios de N-miristoilación.
aminoácidos 47-52, 50-55, 133-138, 142-147, 174-179, 183-188, 281-286, 288-293, 297-302, 324-329, 403-408, 414-419, 415-420, 576-581, 586-591

Inmunoglobulinas y complejos de alta histocompatibilidad.
aminoácidos 135-141

Dominio de ligamiento extracelular.
aminoácidos 156-251, 257-353

Dominio de inmunoglobulina.
aminoácidos 50-139

FIGURA 91

MAVRQWVIALALAALLVVDREVPVAAGKLPFSRMPICEEMVESPTCSQMSWLVCGTDGLTY
TNECQLCLARIKTKQDIQIMKDGKC

Secuencia de señal.

aminoácidos 1-20

Sitio de N-miristoilación.

aminoácidos 58-63

Dominio inhibidor de serina proteasas tipo Kazal.

aminoácidos 37-86

FIGURA 92

MAPLCPSPWLPLLIPAPAPGLTVQLLLSLLLLVPVHPQRLPRMQEDSPLGGGSSGEDDPLG
EEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPEVKPKSEEEGSLKLEDLPTVEAPGDP
QEPQNNAHRDKEGDDQSHWRYGGDPPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPLELL
GFQLPPLPELRLRNNGHSVQLTLPPGLEMALGPGREYRALQLHLHWGAAGRPGSEHTVEGH
RFPAEIHVVHLSTAFARVDEALGRPGGLAVLAAFLEEGPEENSAYEQLLSRLEEIAEEGSE
TQVPGLDISALLPSDFSRYFRYEGSLTTPPCAQGVIWTVFNQTVMLSAKQLHTLSDTLWGP
GDSRLQLNFRATQPLNGRVIEASFPAGVDSSPRAAEPVQLNSCLAAGDILALVFGLLFAVT
SVAFLVQMRRQHRRGTKGGVSYRPAEVAETGA

Secuencia de señal.

aminoácidos 1-37

Dominio transmembrana.

aminoácidos 409-429

Sitio de N-glicosilación.

aminoácidos 346-349

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.

aminoácidos 440-443

Sitios de N-miristoilación.

aminoácidos 50-55, 51-56, 209-214, 236-241, 310-315, 339-344, 421-426, 442-447

Señal de anhidrasas carbónicas tipo eucariota.

aminoácidos 237-253

Anhidrasas carbónicas tipo eucariota.

aminoácidos 141-390

FIGURA 93

MDALQLANSAFAVDLFKQLCEKEPLGNVLFSPICLSTSLSLAQVGAKGDTANEIGQVLHFE
NVKDVPFGFQTVTSDVNKLSSFYSLKLIKRLYVDKSLNLSTEFISSTKRPYAKELETVDFK
DKLEETKGQINNSIKDLTDGHFENILADNSVNDQTKILVVNAAYFVGKWMKKFSESETKEC
PFRVNKTDTKPVQMMNMEATFCMGNIDSINCKIIELPFQNKHLSMFILLPKDVEDESTGLE
KIEKQLNSESLSQWTNPSTMANAKVKLSIPKFKVEKMIDPKACLENLGLKHIFSEDTSDFS
GMSETKGVALSNVIHKVCLEITEDGGDSIEVPGARILQHKDELNADHPFIYIIRHNKTRNI
IFFGKFCSP

Secuencia de señal.
aminoácidos 1-42

Sitios de N-glicosilación.
aminoácidos 99-102, 133-136, 188-191, 361-364

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.
aminoácidos 173-176

Sitios de N-miristoilación.
aminoácidos 130-135, 207-212, 306-311, 312-317

Serpin (inhibidor de serina proteasas).
aminoácidos 3-375

Proteínas de fibrilarina.
aminoácidos 150-169

FIGURA 94

MDRHSSYIFIWLQLELCAMAVLLTKGEIRCYCDAAHCVATGYMCKSELSACFSRLLDPQNS
MSPLTHGCLDSLASTTDICQAKQARNHSGTTIPTLECCHEDMCNYRGLHDVLSPPRGEASG
QGNRYQHDGSRNLITKVQELTSSKELWFRAAVIAVPIAGGLILVLLIMLALRMLRSEMKRL
QDQRQQMLSRLHYSFHGHHSKKGQVAKLDLECMVPVSGHEWCCLTCDKMRQADLSNDKILS
LVHWGMYSGHGKLEFV

Secuencia de señal.
aminoácidos 1-26

Dominio transmembrana.
aminoácidos 152-172

Sitio de N-glicosilación.
aminoácidos 87-90

Sitio de fijación de glicosaminoglicanos.
aminoácidos 121-124, 252-255

Sitios de N-miristoilación.
aminoácidos 68-73, 249-254

FIGURA 95

MRNRRNDTLDSTRTLYSSASRSTDLSYSESDLVNFIQANFKKRECVFFTKDSKATENVCKC
GYAQSQHMEGTQINQSEKWNYKKHTKEFPTDAFGDIQFETLGKKGKYIRLSCDTDAEILYE
LLTQHWHLKTPNLVISVTGGAKNFALKPRMRKIFSRLIYIAQSKGAWILTGGTHYGLTKYI
GEVVRDNTISRSSEENIVAIGIAAWGMVSNRDTLIRNCDAEGYFLAQYLMDDFTRDPLYIL
DNNHTHLLLVDNGCHGHPTVEAKLRNQLEKHISERTIQDSNYGGKIPIVCFAQGGGKETLK
AINTSIKNKIPCVVVEGSGRIADVIASLVEVEDAPTSSAVKEKLVRFLPRTVSRLSEEETE
SWIKWLKEILECSHLLTVIKMEEAGDEIVSNAISYALYKAFSTSEQDKDNWNGQLKLLLEW
NQLDLANDEIFTNDRRWESADLQEVMFTALIKDRPKFVRLFLENGLNLRKFLTHDVLTELF
SNHFSTLVYRNLQIAKNSYNDALLTFVWKLVANFRRGFRKEDRNGRDEMDIELHDVSPITR
HPLQALFIWAILQNKKELSKVIWEQTRGCTLAALGASKLLKTLAKVKNDINAAGESEELAN
EYETRAVELFTECYSSDEDLAEQLLVYSCEANGGSNCLELAVEATDQHFTAQPGVQNFLSK
QWYGEISRDTKNWKIILCLFIIPLVGCGFVSFRKKFVDKHKKLLWYYVAFFTSPFVVFSWN
VVFYIAFLLLFAYVLLMDFHSVPHPPELVLYSLVFVLFCDEVRQWYVNGVNYFTDLWNVMD
TLGLFYFIAGIVFRLHSSNKSSLYSGRVIFCLDYIIFTLRLIHIFTVSRNLGPKIIMLQRM
LIDVFFFLFLFAVWMVAFGVARQGILRQNEQRWRWIFRSVIYEPYLAMFGQVPSDVDGTTY
DFAHCTFTGNESKPLCVELDEHNLPRFPEWITIPLVCIYMLSTNILLVNLLVAMFGYTVGT
VQENNDQVWKFQRYFLVQEYCSRLNIPFPFIVFAYFYMVVKKCFKCCCKEKNMESSVCCFK
NEDNETLAWEGVMKENYLVKINTKANDTSEEMRHRFRQLDTKLNDLKGLLKEIANKIK

Dominio transmembranas.

aminoácidos 681-701, 718-738, 734-754, 757-777, 784-804, 819-839, 853-873, 951-971, 995-1015

Sitios de N-glicosilación.

aminoácidos 6-9, 75-78, 247-250, 308-311, 812-815, 925-928, 1041-1044, 1063-1066

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.

aminoácidos 83-86

Sitio de fosforilación tirosina cinasas.

aminoácidos 21-27, 219-226, 279-286

Sitios de N-miristoilación.

aminoácidos 71-76, 141-146, 174-179, 209-214, 299-304,

577-582, 643-648, 904-909

Sitio de amidación.

aminoácidos 102-105

FIGURA 96

MRNRRNDTLDSTRTLYSSASRSTDLSYSESDLVNFIQANFKKRECVFFTKDSKATENVCKC
GYAQSQHMEGTQINQSEKNNYKKHTKEFPTDAFGDIQFETLGKKGKYIRLSCDTDAEILYE
LLTQHWHLKTPNLVISVTGGAKNFALKPRMRKIFSRLIYIAQSKGAWILTGGTHYGLMKYI
GEVVRDNTISRSSEENIVAIGIAAWGMVSNRDTLIRNCDAEGYFLAQYLMDDFTRDPLYIL
DNNHTHLLLVDNGCHGHPTVEAKLRNQLEKYISERTIQDSNYGGKIPIVCFAQGGGKETLK
AINTSIKNKIPCVVVEGSGQIADVIASLVEVEDALTSSAVKEKLVRFLPRTVSRLPEEETE
SWIKWLKEILECSHLLTVIKMEEAGDEIVSNAISYALYKAFSTSEQDKDNWNGQLKLLLEW
NQLDLANDEIFTNDRRNESADLQEVMFTALIKDRPKFVRLFLENGLNLRKFLTHDVLTKLF
SNHFSTLVYRNLQIAKNSYNDALLTFVWKLVANFRRGFRKEDRNGRDEMDIELHDVSPITR
HPLQALFIWAILQNKKELSKVIWEQTRGCTLAALGASKLLKILAKVKNDINAAGESEELAN
EYETRAVELFTECYSSDEDLAEQLLVYSCEAWGGSNCLELAVEATDQHFIAQPGVQNFLSK
QWYGEISRDTKNWKIILCLFIIPLVGCGFVSFRKKPVDKHKKLLWYYVAFFTSPFVVFSWN
VVFYIAFLLLFAYVLLMDFHSVPHPPELVLYSLVFVLFCDEVRQWYVNGVNYFTDLWNVMD
TLGLFYFIAGIVFRLHSSNKSSLYSGRVIFCLDYIIFTLRLIRIFTVSRNLGPKIIMLQRM
LIDVFFFLFLFAVWMVAFGVARQGILRQNEQRWRWIFRSVIYEPYLAMFGQVPSDVDGTTY
DFAHCTFTGNESKPLCVELDEHNLPRFPEWITIPLVCIYMLSTNILLVNLLVAMFGYTVGT
VQENNDQVWKFQRYFLVQEYCSRLNIPFPFIVFAYFYMVVKKCFKCCCKEKNMESSVCCFK
NEDNETLAWEGVMKENYLVKINTKANDTSEEMRHRFRQLDTKLNDLKGLLKEIANKIK

Dominio transmembranas.

aminoácidos 681-701, 718-738, 734-754, 757-777, 784-804, 819-839, 853-873, 951-971, 995-1015

Sitios de N-glicosilación.

aminoácidos 6-9, 75-78, 247-250, 308-311, 812-815, 925-928, 1041-1044, 1063-1066

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.

aminoácidos 83-86

Sitio de fosforilación tirosina cinasas.

aminoácidos 21-27, 219-226, 279-286

Sitios de N-miristoilación.

aminoácidos 71-76, 141-146, 174-179, 209-214, 299-304,

577-582, 643-648, 904-909

Sitio de amidación.

aminoácidos 102-105

FIGURA 97

RRGRGEVLAVERGSGSGGGGTRSGWPAPAAGADKKPSRCGSGREGEGVSEGHKSMTGLYEL
VWRVLHALLCLHRTLTSWLRVRFGTWNWIWRRCCRAASAAVLAPLGFTLRKPPAVGRNRRH
HRHPRGGSCLAAAEHRMRWRADGRSLEKLPVHMGLVITEVEQEPSFSDIASLVVWCMAVGI
SYISVYDHQGIFKRNNSRLMDEILKQQQELLGLDCSKYSPEFANSNDKDDQVLNCHLAVKV
LSPEDGKADIVRAAQDFCQLVAQKQRRPTDLDVDTLASLLSSNGCPDPDLVLKFGPVDSTL
GFLPWHIRLTEIVSLPSHLNISYEDFFSALRQYAACEQRLGK

Dominio transmembrana.
aminoácidos 166-186

Sitios de N-glicosilación.
aminoácidos 198-201, 325-328

Sitio de fijación de glicosaminoglicanos.
aminoácidos 14-17, 16-19

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasas.
aminoácidos 34-37, 270-273

Sitios de N-miristoilación.
aminoácidos 13-18, 15-20, 17-22, 19-24, 20-25, 47-52, 129-134, 156-161, 215-220

FIGURA 98

MTGLYELVWRVLHALLCLHRTLTSWLRVRFGTWNWIWRRCCRAASAAVLAPLGFTLRKPPA
VGRNRRHHRHPRGGSCLAAAHHRMRWRADGRSLEKLPVHMGLVITEVEQEPSFSDIASLVV
WCMAVGISYISVYDHQGIFKRNNSRLMDEILKQQQELLGLDCSKYSPEFANSNDKDDQVLN
CHLAVKVLSPEDGKADIVRAAQDFCQLVAQKQKRPTDLDVDTLASLLSSNGCPDPDLVLKP
GPVDSTLGFLPWHIRLTEIVSLPSHLNISYEDFFSALRQYAACEQRLGK

Secuencia de señal.
aminoácidos 1-23

Dominio transmembrana.
aminoácidos 115-132

Sitios de N-glicosilación.
aminoácidos 144-147, 271-274

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.
aminoácidos 216-219

Sitios de N-miristoilación.
aminoácidos 75-80, 102-107, 161-166

FIGURA 99

MVQRLWVSRLLRHRKAQLLLVNLLTFGLEVCLAAGITYVPPLLLEVGVEEKFMTMVLGIGP
VLGLVCVPLLGSASDHWRGRYGRRRPFIWALSLGILLSLFLIPRAGWLAGLLCPDPRPLEL
ALLILGVGLLDFCGQVCFTPLEALLSDLFRDPDHCRQAYSVYAFMISLGGCLGYLLPAIDW
DTSALAPYLGTQEECLFGLLTLIFLTCVAATLLVAEEAALGPTEPAEGLSAPSLSPHCCPC
RARLAFRNLGALLPRLHQLCCRMPRTLRRLFVAELCSWMALMTFTLFYTDFVGEGLYQGVP
RAEPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFGTRAVYLASVAAFPVAA
GATCLSHSVAVVTASAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGGASSEDSL
NTSFLPGPKPGAPFPNGHVGAGGSGLLPPPPALCGASACDVSVRVVVGEPTEARVVPGRGI
CLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIYFATQVVFDKSDLA
KYSA

Secuencia de señal.

aminoácidos 1-35

Dominio transmembranas.

aminoácidos 15-35, 52-72, 87-107, 122-142, 160-180, 240-260, 277-297, 321-341, 486-506, 521-541

Sitios de N-miristoilación.

aminoácidos 27-32, 95-100, 171-176, 299-304, 303-308, 320-325, 419-424, 447-452, 462-467

Sitio de amidación.

aminoácidos 82-85

Secuencia de adhesión de células.

aminoácidos 415-417

FIGURA 100

MALRRLGAALLLLPLLAAVEETLMDSTTATAELGWMVHPPSGWEEVSGYDENMNTIRTYQV
CNVFESSQNNWLRTKFIRRRGAHRIHVEMKFSVRDCSSIPSVPGSCKETFNLYYYEADFDS
ATKTFPNWMENPWVKVDTIAADESFSQVDLGGRVMKINTEVRSFGPVSRSGFYLAFQDYGG
CMSLIAVRVFYRKCPRIIQNGAIFQETLSGAESTSLVAARGSCIANAEEVDVPIKLYCNGD
GEWLVPIGRCMCKAGFEAVENGTVCRGCPSGTFKANQGDEACTHCPINSRTTSEGATNCVC
RNGYYRADLDPLDMPCTTIPSAPQAVISSVNETSLMLEWTPPRDSGGREDLVYNIICKSCG
SGRGACTRCGDNVQYAPRQLGLTEPRIYISDLLAHTQYTFEIQAVNGVTDQSPFSPQFASV
NITTNQAAPSAVSIMHQVSRTVDSITLSWSQPDQPNGVILDYELQYYEKELSEYNATAIKS
PTNTVTVQGLKAGAIYVFQVRARTVAGYGRYSGKMYFQTMTEAEYQTSIQEKLPLIIGSSA
AGLVFLIAVVVIAIVCNRRRGFERADSEYTDKLQHYTSGHMTPGMKIYIDPFTYEDPNEAV
REFAKEIDISCVKIEQVIGAGEFGEVCSGHLKLPGKREIFVAIKTLKSGYTEKQRRDFLSE
ASIMGQFDHPNVIHLEGVVTKSTPVMIITEFMENGSLDSFLRQNDGQFTVIQLVGMLRGIA
AGMKYLADMNYVHRDLAARNILVNSNLVCKVSDFGLSRFLEDDTSDPTYTSALGGKIPIRW
TAPEAIQYRKFTSASDVWSYGIVMWEVMSYGERPYWDMTNQDVINAIEQDYRLPPPMDCPS
ALHQLMLDCWQKDRNHRPKFGQIVNTLDKMIRNPNSLKAMAPLSSGINLPLLDRTIPDYTS
FNTVDEWLEAIKMGQYKESFANAGFTSFDVVSQMMMEDILRVGLTLAGHQKKILNSIQVMR
AQMNQIQSVEV

Secuencia de señal.

aminoácidos 1-18

Dominio transmembrana.

aminoácidos 542-562

Sitios de N-glicosilación.

aminoácidos 265-268, 336-339, 428-431, 482-485, 705-708

Sitio de fijación de glicosaminoglicano.

aminoácidos 367-370

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.

aminoácidos 802-805

Sitio de fosforilación tirosina cinasas.

aminoácidos 374-381, 595-603, 736-743

Sitios de N-miristoilación.

aminoácidos 182-187, 213-218, 224-229, 271-276, 275-280, 299-304, 366-371, 497-502, 546-551, 706-711, 726-731, 730-735, 875-880, 958-963

Sitio de amidación - aminoácidos 644-647

Señal de sitio activo específico de tirosina proteína cinasas - aminoácidos 743-755

Señal 1 del receptor de tirosina cinasa clase V - aminoácidos 182-198

Señal 2 del receptor de tirosina cinasa clase V - aminoácidos 241-261

Dominio de unión de ligando efrin receptor - aminoácidos 20-197

Dominio de proteína cinasa - aminoácidos 622-881

Dominio de fibronectina tipo III - aminoácidos 325-421, 436-520

Dominio SAM (Sterile alpha motif) - aminoácidos 912-976

FIGURA 101

MALRRLGAALLLLPLLAAVEETLMDSTTATAELGWMVHPPSGWEEVSGYDENMNTIRTYQV
CNVFESSQNNWLRTKFIRRRGAERIHVEMKFSVRDCSSIPSVPGSCKETFNLYYYEADFDS
ATKTFPNWMENPWVKVDTIAADESFSQVDLGGRVMKINTEVRSFGPVSRSGFYLAFQDYGG
CMSLIAVRVFYRKCPRIIQNGAIFQETLSGAESTSLVAARGSCIANAEEVDVPIKLYCNGD
GEWLVPIGRCMCKAGFEAVENGTVCRGCPSGTFKANQGDEACTHCPINSRTTSEGATNCVC
RNGYYRADLDPLDMPCTTIPSAPQAVISSVNETSLMLEWTPPRDSGGREDLVYNIICKSCG
SGRGACTRCGDNVQYAPRQLGLTEPRIYISDLLAHTQYTFEIQAVNGVTDQSPFSPQFASV
NITTNQAAPSAVSIMHQVSRTVDSITLSWSQPDQPNGVILDYELQYYEKELSEYNATAIKS
PTNTVTVQGLKAGAIYVFQVRARTVAGYGRYSGKMYFQTMTEAEYQTSIQEKLPLIIGSSA
AGLVFLIAVVVIAIVCNRRGFERADSEYTDKLQHYTSGHMTPGMKIYIDPFTYEDPNEAVR
EFAKEIDISCVKIEQVIGAGEFGEVCSGHLKLPGKREIFVAIKTLKSGYTEKQRRDFLSEA
SIMGQFDHPNVIHLEGVVTKSTPVMIITEFMENGSLDSFLRQNDGQFTVIQLVGMLRGIAA
GMKYLADMNYVHRDLAARNILVNSNLVCKVSDFGLSRFLEDDTSDPTYTSALGGKIPIRWT
APEAIQYRKFTSASDVWSYGIVMWEVMSYGERPYWDMTNQDVINAIEQDYRLPPPMDCPSA
LHQLMLDCWQKDRNHRPKFGQIVNTLDKMIRNPNSLKAMAPLSSGINLPLLDRTIPDYTSF
NTVDEWLEAIKMGQYKESFANAGFTSFDVVSQMMMEDILRVGVTLAGHQKKILNSIQVMRA
QMNQIQSVEV

Secuencia de señal.

aminoácidos 1-18

Dominio transmembrana.

aminoácidos 543-563

Sitios de N-glicosilación.

aminoácidos 265-268, 336-339, 428-431, 482-485, 704-707

Sitio de fijación de glicosaminoglicano.

aminoácidos 367-370

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.

aminoácidos 801-804

Sitio de fosforilación tirosina cinasas.

aminoácidos 374-381, 594-602, 735-742

Sitios de N-miristoilación.

aminoácidos 182-187, 213-218, 224-229, 271-276, 275-280, 299-304, 366-371, 497-502, 546-551, 705-710, 725-730, 729-734, 874-879, 957-962

Sitio de amidación.

aminoácidos 643-646

Señal de sitio activo específico de tirosina proteína cinasas.

aminoácidos 742-754

Señal 1 del receptor de tirosina cinasa clase V.

aminoácidos 182-198

Señal 2 del receptor de tirosina cinasa clase V.

aminoácidos 241-264

Dominio de unión de ligando efrin receptor.

aminoácidos 20-197

Dominio de proteína cinasa.

aminoácidos 621-880

Dominio de fibronectina tipo IIIs.

aminoácidos 325-421, 436-520

Dominio SAM (Sterile alpha motif).

aminoácidos 911-975

Dominios (SH2) de homología a Src 2.

aminoácidos 733-769, 787-797, 806-819

FIGURA 102

MNFSTSSSSFAYDREFLRTLPGFLIVAEIVLGLLVWTLIAGTEYFRVPAFGWVMFVAVFYW
VLTVFFLIIYITMTYTRIPQVPWTTVGLCFNGSAFVLYLSAAVVDASSVSPERDSHNFNSW
AASSFFAFLVNICYAGNTYFSFIAWRSRTIQ

Secuencia de señal.

aminoácidos 1-41

Dominio transmembranas.

aminoácidos 53-73, 86-106, 122-142

Sitios de N-glicosilación.

aminoácidos 2-5, 92-95

Sitio de N-miristoilación.

aminoácidos 88-93

Patrón de cremallera de leucina.

aminoácidos 17-38

FIGURA 103

MNFSTSSSSFAYDREFLRTLPGFLIVAEIVLGLLVWTLIAGTEYFRVPAFGWVMFVAVFYW
VLTVFFLIIYITMTYTRIPQVPWTTVGLCFNGSAFVLYLSAAVVDASSVSPERDSHNFNSW
AASSFFAFLVTICYAGNTYFSFIAWRSRTIQ

Dominio transmembranas.
aminoácidos 19-39, 53-73, 86-106, 122-142

Sitios de N-glicosilación.
aminoácidos 2-5, 92-95

Sitio de N-miristoilación.
aminoácidos 88-93

Patrón de cremallera de leucina.
aminoácidos 17-38

FIGURA 104

MPTTVDDVLEHGGEFHFFQKQMFFLLALLSATFAPIYVGIVFLGFTPDHRCRSPGVAELSL
RCGWSPARELNYTVPGPGPAGEASPRQCRRYEVDWNQSTFDCVDPLASLDTNRSRLPLGPC
RDGWVYETPGSSIVTEFNLVCANSWMLDLFQSSVNVGFFIGSMSIGYIADRFGRKLCLLTT
VLINAAAGVLMAISPTYTWMLIFRLIQGLVSKAGWLIGYILITEFVGRRYRRTVGIFYQVA
YTVGLLVLAGVAYALPHWRWLQFTVALPNFFFLLYYWCIPESPRWLISQNKNAEAMRIIKH
IAKKNGKSLPASLQRLRLEEETGKKLNPSFLDLVRTPQIRKHTMILMYNWFTSSVLYQGLI
MEMGLAGDNIYLDFFYSALVEFPAAFMIILTIDRIGRRYPWAASNMVAGAACLASVFIPGD
LQWLKIIISCLGRMGITMAYEIVCLVNAELYPTFIRNLGVHICSSMCDIGGIITPFLVYRL
TNIWLELPLMVFGVLGLVAGGLVLLLPETKGKALPETIEEAENMQRPRKNKEKMIYLQVQK
LDIPLN

Dominio transmembranas.

aminoácidos 20-40, 174-194, 205-225, 238-258, 269-289, 351-371, 376-396, 412-432, 492-512

Sitios de N-glicosilación.

aminoácidos 72-75, 97-100, 113-116

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.

aminoácidos 345-348

Sitio de fosforilación tirosina cinasa.

aminoácidos 536-553

Sitios de N-miristoilación.

aminoácidos 191-196, 254-259, 442-447, 466-471, 504-509

Sitio de amidación.

aminoácidos 174-177, 229-232, 327-330, 401-404

Transportador de azúcar (y de otros).

aminoácidos 23-529

FIGURA 105

MPTTVDDVLEHGGEFHFFQKQMPFLLALLSATFAPIYVGIVFLGFTPDHRCRSPGVAELSL
RCGWSPAEELNYTVPGPGPAGEASPRQCRRYEVDWNQSTFDCVDPLASLDTNRSRLPLGPC
RDGWVYETPGSSIVTEFELVCANSWMLDLFQSSVNVGFFIGSMSIGYIADRFGRKLCLLTT
VLINAAAGVLMAISPTYTWMLIFRLIQGLVSKAGWLIGYILITEFVGRRYRRTVGIFYQVA
YTVGLLVLAGVAYALPHWRWLQFTVSLPNFFFLLYYWCIPESPRWLISQNKNAEAMRIIKH
IAKKNGKSLPASLQRLRLEEETGKKLNPSFLDLVRTPQIRKHTMILMYNWFTSSVLYQGLI
MEMGLAGDNIYLDFFYSALVEFPAAFMIILTIDRIGRRYPWAASNMVAGAACLASVFIPGD
LQWLKIIISCLGRMGITMAYEIVCLVNAELYPTFIRNLGVHICSSMCDIGGIITPFLVYRL
TNIWLELPLMVFGVLGLVAGGLVLLLPETKGKALPETIEEAENMQRPRKNKEKMIYLQVQK
LDIPLN

Dominio transmembranas.

aminoácidos 20-40, 174-194, 205-225, 238-258, 269-289, 351-371, 376-396, 412-432, 492-512

Sitios de N-glicosilación.

aminoácidos 72-75, 97-100, 113-116

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.

aminoácidos 345-348

Sitio de fosforilación tirosina cinasa.

aminoácidos 536-544

Sitios de N-miristoilación.

aminoácidos 191-196, 254-259, 442-447, 466-471, 504-509

Sitios de amidación.

aminoácidos 174-177, 229-232, 327-330, 401-404

Transportador de azúcar (y de otros).

aminoácidos 23-529

FIGURA 106

MAPWPELGDAQPNPDKYLEGAAGQQPTAPDKSKETNKTDNTEAPVTKIELLPSYSTATLID
EPTEVDDPWNLPTLQDSGIKWSERDTKGKILCFFQGIGRLILLLGFLYFFVCSLDILSSAF
QLVGGKMAGQFFSNSSIMSNPLLGLVIGVLVTVLVQSSSTSTSIVVSMVSSSLLTVRAAIP
IIMGANIGTSITNTIVALMQVGDRSEFRRAFAGATVHDFFNWLSVLVLLPVEVATHYLEII
TQLIVESFHFKNGEDAPDLLKVITKPFTKLIVQLDKKVISQIAMNDEKAKNKSLVKIWCKT
FTNKTQINVTVPSTANCTSPSLCWTDGIQNWTMKNVTYKENIAKCQHIFVNFHLPDLAVGT
ILLILSLLVLCGCLIMIVKILGSVLKGQVATVIKKTINTDFPFPFAWLTGYLAILVGAGMT
FIVQSSSVFTSALTPLIGIGVITIERAYPLTLGSNIGTTTTAILAALASPGNALRSSLQIA
LCHFFFNISGILLWYPIPFTRLPIRMAKGLGNISAKYRWFAVFYLIIFFFLIPLTVFGLSL
AGWRVLVGVGVPVVFIIILVLCLRLLQSRCPRVLPKKLQNWNFLPLWMRSLKPWDAVVSKF
TGCFQMRCCYCCRVCCRACCLLCGCPKCCRCSKCCEDLEEAQEGQDVPVKAPETFDNITIS
REAQGEVPASDSKTECTAL

Dominio transmembranas.
aminoácidos 96-116, 136-156, 178-198, 219-239, 356-376, 372-392, 406-426, 445-465, 488-508, 523-543, 549-569, 565-585, 592-612

Sitios de N-glicosilación.
aminoácidos 36-39, 136-139, 295-298, 308-311, 313-316, 321-324, 335-338, 340-343, 495-498, 520-532, 667-670

Sitios de N-miristoilación.
aminoácidos 23-28, 79-84, 126-131, 131-136, 146-151, 150-155, 187-192, 191-196, 393-398, 423-428, 460-465, 464-469, 519-524, 546-551, 634-639

Señal de región de unión hierro-sulfuro ferredoxinas 4Fe-4S.
aminoácidos 635-646

Señal de la familia de la insulina
aminoácidos 621-635

Enterotoxinas estables en calor.

aminoácidos 617-633, 625-641, 613-629

cotransportador Na+/Pi

aminoácidos 118-549

FIGURA 107

MAPWPELGDAQPNPDKYLEGAAGQQPTAPDKSKETNKTDNTEAPVTKIELLPSYSTATLID
EPTEVDDPWNLPTLQDSGIKWSERDTKGKILCFFQGIGRLILLLGFLYFFVCSLDILSSAF
QLVGGKMAGQFFSNSSIMSNPLLGLVIGVLVTVLVQSSSTSTSIVVSMVSSSLLTVRAAIP
IIMGANIGTSITNTIVALMQVGDRSEFRRAFAGATVHDFFNWLSVLVLLPVEVATHYLEII
TQLIVESFHFKNGEDAPDLLKVITKPFTKLIVQLDKKVISQIAMNDEKAKNKSLVKIWCKT
FTNKTQINVTVPSTANCTSPSLCWTDGIQNWTMKNVTYKENIAKCQHIFVNFHLPDLAVGT
ILLILSLLVLCGCLIMIVKILGSVLKGQVATVIKKTINTDFPFPFAWLTGYLAILVGAGMT
FIVQSSSVFTSALTPLIGIGVITIERAYPLTLGSNIGTTTTAILAALASPGNALRSSLQIA
LCHFFFNISGILLWYPIPFTRLPIRMAKGLGNISAKYRWFAVFYLIIFFFLIPLTVFGLSL
AGWRVLVGVGVPVVFIIILVLCLRLLQSRCPRVLPKKLQNWNFLPLWMRSLKPWDAVVSKF
TGCFQMRCCCCCRVCCRACCLLCGCPKCCRCSKCCEDLEEAQEGQDVPVKAPETFDNITIS
REAQGEVPASDSKTECTAL

Dominio transmembranas.

aminoácidos 96-116, 136-156, 178-198, 219-239, 356-376, 372-392, 406-426, 445-465, 488-508, 523-543, 547-567, 563-583, 592-612

Sitios de N-glicosilación.

aminoácidos 36-39, 136-139, 295-298, 308-311, 313-316, 321-324, 335-338, 340-343, 495-498, 520-523, 667-670

Sitios de N-miristoilación.

atnino acids 23-28, 79-84, 126-131, 131-136, 146-151, 150-155, 187-192, 191-196, 393-398, 423-428, 460-465, 464-469, 519-524, 546-551, 634-639

Señal de región de unión hierro-sulfuro ferredoxinas 4Fe-4S.

aminoácidos 635-645

Señal de la familia de la insulina.

aminoácidos 621-635

Cotransportador Na+/Pi.

aminoácidos 118-549

FIGURA 108

MWSGWWLWPLVAVCTADFFRDEAERIMRDSPVIDGHNDLPWQLLDMFNNRLQDERANLTTL
AGTHTNIPKLRAGFVGGQFWSVYTPCDTQNKDAVRRTLEQMDVVHRMCRMYPETFLYVTSS
AGIRQAFREGKVASLIGVEGGHSIDSSLGVLRALYQLGMRYLTLTHSCNTPWADNWLVDTG
DSEPQSQGLSPFGQRVVKELNRLGVLIDLAHVSVATMKATLQLSRAPVIFSHSSAYSVCAS
RRNVPDDVLRLVKQTDSLVMVNFYNNYISCTNKANLSQVADHLDHIKEVAGARAVGFGGDF
DGVPRVPEGLEDVSKYPDLIAELLRRNWTEAEVKGALADNLLRVFQAVEQASNLTQAPEEE
PIPLDQLGGSCRTHYGYSSGASSLHRHWGLLLASLAPLVLCLSLL

Secuencia de señal

aminoácidos 1-16

Dominio transmembrana.

aminoácidos 392-410

Sitios de N-glicosilación.

aminoácidos 57-60, 279-282, 332-335, 358-361

Sitio de fosforilación tirosina cinasa.

aminoácidos 110-118

Sitios de N-miristoilación.

aminoácidos 63-68, 78-83, 124-129, 139-144, 151-156, 375-380, 395-400

Patrón de cremallera de leucina.

aminoácidos 390-411

Sitio activo de dipeptidasa renal.

aminoácidos 140-162

Dipeptidasa renal.

aminoácidos 1-411

FIGURA 109

MSSCNFTHATFVLIGIPGLEKAHFWVGFPLLSMYVVAMFGNCIVVFIVRTERSLHAPMYLF
LCMLAAIDLALSTSTMPKILALFWFDSREISFEACLTQMFFIHALSAIESTILLAMAFDRY
VAICHPLRHAAVLNNTVTAQIGIVAVVRGSLFFFPLPLLIKRLAFCHSNVLSHSYCVHQDV
MKLAYADTLPNVVYGLTAILLVMGVDVMFISLSYFLIIRTVLQLPSKSERAKAFGTCVSHI
GVVLAFYVPLIGLSVVHRFGNSLHPIVRVVMGDIYLLLPPVINPIIYGAKTKQIRTRVLAM
FKISCDKDLQAVGGK

Dominio transmembranas.
aminoácidos 26-46, 57-77, 96-116, 144-164, 197-217, 238-258, 272-292

Sitios de N-glicosilación.
aminoácidos 5-8, 136-139

Sitio de N-miristoilación.
aminoácidos 18-23, 238-243, 245-250

Receptor de transmembrana 7 (familia de rodopsinas).
aminoácidos 40-131, 212-291

FIGURA 110A

MRILKRFLACIQLLCVCRLDWANGYYRQQRKLVEEIGWSYTGALNQKNWGKKYPTCNSPKQ
SPINIDEDLTQVNVNLKKLKFQGWDKTSLENTFIHNTGKTVEINLTNDYRVSGGVSEMVFK
ASKITFHWGKCNMSSDGSEHSLEGQKFPLEMQIYCFDADRFSSFEEAVKGKGKLRALSILF
EVGTEENLDFKAIIDGVESVSRFGKQAALDPFILLNLLPNSTDKYYIYNGSLTSPPCTDTV
DWIVFKDTVSISESQLAVFCEVLTMQQSGYVMLMDYLQNNFREQQYKFSRQVFSSYTGKEE
IHEAVCSSEPENVQADPENYTSLLVTWERPRVVYDTMIEKFAVLYQQLDGEDQTKHEFLTD
GYQDLGAILNNLLPNMSYVLQIVAICTNGLYGKYSDQLIVDMPTDNPELDLFPELIGTEEI
IKEEEEGKDIEEGAIVNPGRDSATNQIRKKEPQISTTTHYNRIGTKYNEAKTNRSPTRGSE
FSGKGDVPNTSLNSTSQPVTKLATEKDISLTSQTVTELPPHTVEGTSASLNDGSKTVLRSP
HMNLSGTAESLNTVSITEYEEESLLTSFKLDTGAEDSSGSSPATSAIPFISENISQGYIFS
SENPETITYDVLIPESARNASEDSTSSGSEESLKDPSMEGNVWFPSSTDITAQPDVGSGRE
SFLQTNYTEIRVDESEKTTKSFSAGPVMSQGPSVTDLEMPHYSTFAYFPTEVTPHAFTPSS
RQQDLVSTVNVVYSQTTQPVYNGETPLQPSYSSEVFPLVTPLLLDNQILNTTPAASSSDSA
LHATPVFPSVDVSFESILSSYDGAPLLPFSSASFSSELFRHLHTVSQILPQVTSATESDKV
PLHASLPVAGGDLLLEPSLAQYSDVLSTTHAASETLEFGSESGVLYKTLMFSQVEPPSSDA
MMHARSSGPEPSYALSDNEGSQHIFTVSYSSAIPVHDSVGVTYQGSLFSGPSHIPIPKSSL
ITPTASLLQPTHALSGDGEWSGASSDSEFLLPDTDGLTALNISSPVSVAEFTYTTSVFGDD
NKALSKSEIIYGNETELQIPSFNEMVYPSESTVMPNMYDNVNKLNASLQETSVSISSTKGM
FPGSLAHTTTKVFDHEISQVPENNFSVQPTHTVSQASGDTSLKPVLSANSEPASSDPASSE
MLSPSTQLLFYETSASFSTEVLLQPSFQASDVDTLLKTVLPAVPSDPILVETPKVDKISST
MLHLIVSNSASSENMLHSTSVPVFDVSPTSHMHSASLQGLTISYASEKYEPVLLKSESSHQ
VVPSLYSNDELFQTANLEINQAHPPKGRHVFATPVLSIDEPLNTLINKLIHSDEILTSTKS
SVTGKVFAGIPTVASDTFVSTDHSVPIGNGHVAITAVSPHRDGSVTSTKLLFPSKATSELS
HSAKSDAGLVGGGEDGDTDDDGDDDDDRDSDGLSIHKCMSCSSYRESQEKVMNDSDTHENS
LMDQNNPISYSLSENSEEDNRVTSVSSDSQTGMDRSPGKSPSANGLSQKHNDGKEENDIQT
GSALLPLSPESKAWAVLTSDEESGSGQGTSDSLNENETSTDFSFADTNEKDADGILAAGDS
EITPGFPQSPTSSVTSENSEVFHVSEAEASNSSHESRIGLAEGLESEKKAVIPLVIVSALT
FICLVVLVGILIYWRKCFQTAHFYLEDSTSPRVISTPPTPIFPISDDVGAIPIKHFPKHVA
DLHASSGFTEEFETLKEFYQEVQSCTVDLGITADSSNHPDNKHKNRYINIVAYDHSRVKLA
QLAEKDGKLTDYINANYVDGYNRPKAYIAAQGPLKSTAEDFWRMIWEHNVEVIVMITNLVE
KGRRKCDQYWPADGSEEYGNFLVTQKSVQVLAYYTVRNFTLRNTKIKKGSQKGRPSGRVVT
QYHYTQWPDMGVPEYSLPVLTFVRKAAYAKRHAVGPVVVHCSAGVGRTGTYIVLDSMLQQI
QHEGTVNIFGFLKHIRSQRNYLVQTEEQYVFIHDTLVEAILSKETEVLDSHIHAYVNALLI
PGPAGKTKLEKQFQLLSQSNIQQSDYSAALKQCNREKNRTSSIIPVERSRVGISSLSGEGT
DYINASYIMGYYQSNEFIITQHPLLHTIKDFWRMIWDHNAQLVVMIPDGQNMAEDEFVYWP

NKDEPINCESFKVTLMAEEHKCLSNEEKLIIQDFILEATQDDYVLEVRHFQCPKNPNPDSP
ISKTFELISVIKEEAANRDGPMIVHDEHGGVTAGTFCALTTLMHQLEKENSVDVYQVAKMI
NLMRPGVFADIEQYQFLYKVILSLVSTRQEENPSTSLDSNGAALPDGNIAESLESLV

Secuencia de señal.
aminoácidos 1-19

Dominio transmembrana.
aminoácidos 1638-1658

Sitios de N-glicosilación.
aminoácidos 105-108, 134-137, 223-226, 232-235, 324-327, 381-384, 497-500, 501-504, 552-555, 602-605, 629-632, 677-680, 1017-1020, 1050-1053, 1082-1085, 1122-1125, 1456-1459, 1561-1564, 1617-1620, 1868-1871, 2051-2054, 2078-2081

Sitio de fijación de glicosaminoglicanos.
aminoácidos 490-493, 991-994, 1548-1551, 1550-1553, 2070-2073

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.
aminoácidos 1877-1880

Sitio de fosforilación tirosina cinasa.
aminoácidos 148-156

FIGURA 110B

Sitios de N-miristoilación.
aminoácidos 139-144, 186-191, 372-377, 471-476, 486-491, 533-538, 555-560, 582-587, 588-593, 638-643, 893-898, 960-965, 1097-1102, 1259-1264, 1385-1390, 1411-1416, 1415-1420, 1549-1554, 1551-1556, 1553-1558, 1579-1584, 1625-1630, 1879-1884, 1935-1940, 2123-2128, 2225-2230, 2226-2231, 2230-2235
Sitios de amidación.
aminoácidos 49-52, 1831-1834
Sitio activo de las proteínas fosfatasas específicas de tirosina.
aminoácidos 1930-1942
Fosfatasa proteína-tirosina.
aminoácidos 1749-1990, 2047-2280
Anhidrasas carbónicas tipo eucariota.
aminoácidos 38-300
Dominio de fibronectina tipo III.
aminoácidos 312-401
Proteínas de dominio MAM.
aminoácidos 1757-1785, 1060-1088, 1813-1853, 2047-2079

FIGURA 111

MCARMAGRTRAAPRGPYGPWLCLLVALALDVVRVDCGQAPLDPVYLHVTAARPAQPTLWTA
KLDRFKGSRHHTTLITCHRAGLTEPDSSSPLELSEFLWVDFVVENSTGGGVAVTRPVTWQL
EYPGQAPEAEKDKMVWEILVSERDIRALIPLAKAEELVNTAPLTGVPQHVPVRLVTVDGGG
ALVEVTEHVGCESANTQVLQVSEACDAVFVAGKESRGARGVRVDFWWRRLRASLRLTVWAP
LLPLRIELTDTTLEQVRGWRVPGPAEGPAEPAAEASDEAERRARGCHLQYQRAGVRFLAPF
AAHPLDGGRRLTHLLGPDWLLDVSHLVAPHARVLDSRVASLEGGRVVVGREPGVTSIEVRS
PLSDSILGEQALAVTDDKVSVLELRVQPVMGISLTLSRGTAHPGEVTATCWAQSALPAPKQ
EVALSLWLSFSDHTVAPAELYDRRDLGLSVSAEEPGAILPAEEQGAQLGVVVSGAGAEGLP
LHVALHPPEPCRRGRHRVPLASGTAWLGLPPASTPAPALPSSPAWSPPATEATMGGKRQVA
GSVGGNTGVRGKFERAEEEARKEETEAREEEEEEEEEMVPAPQHVTELELGMYALLGVFCV
AIFIFLVNGVVFVLRYQRKEPPDSATDPTSPQPHNWVWLGTDQEELSRQLDRQSPGPPKGE
GSCPCESGGGGEAPTLAPGPPGGTTSSSSTLARKEAGGRRKRVEFVTFVPAPPAQSPEEPV
GAPAVQSILVAGEEDIRWVCEDMGLKDPEELRNYMERIRGSS

Sitio de N-glicosilación.
aminoácidos 106-109

Sitio de fijación de glicosaminoglicanos.
aminoácidos 480-483, 678-681

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.
aminoácidos 314-317

Sitio de fosforilación tirosina cinasa.
aminoácidos 758-766

Sitios de N-miristoilación.
aminoácidos 109-114, 111-116, 193-198, 397-402, 454-459, 472-477, 476-481, 516-521, 550-555, 553-558, 672-677, 693-698, 694-694

Sitios de amidación.
aminoácidos 312-315, 543-546, 708-711

FIGURA 112

MATKTELSPTARESKNAQDMQVDETLIPRKGPSLCSARYGIALVLHFCNFTTIAQNVIMNI
TMVAMVNSTSPQSQLNDSSEVLPVDSFGGLSKAPKSLPAKSSILGGQFAIWEKWGPPQERS
RLCSIALSGMLLGCFTAILIGGFISETLGWPFVFYIFGGVGCVCCLLWFVVIYDDPFSYPW
ISTSEKEYIISSLKQQVGSSKQPLPIKAMLRSLPIWSICLGCFSHQWLVSTMVVYIPTYIS
SVYHVNIRDNGLLSALPFIVAWVIGMVGGYLADFLLTKKFRLITVRKIATILGSSLPSSALI
VSLPYLNSGYITATALLTLSCGLSTLCQSGIYINVLDIAPRYSSFLMGASRGFSSIAPVIV
PTVSGFLLSQDPEFGWRNVFFLLFAVNLLGLLFYLIFGEADVQEWAKERKLTRL

Dominio transmembranas.

aminoácidos 124-144, 154-174, 255-275, 287-307, 319-339, 350-370, 385-405

Sitios de N-glicosilación.

aminoácidos 49-52, 60-63, 68-71, 77-80

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.

aminoácidos 4i5-418

Sitios de N-miristoilación.

aminoácidos 106-111, 131-136, 135-140, 143-148, 160-165, 163-168, 255-260, 269-274, 272-277, 297-302, 335-340, 353-358

FIGURA 113

MQVDETLIPRKVPSLCSARYGIALVLHFCNFTTIAQNVIMNITMVAMVNSTSPQSQLNDSS
EVLPVDSFGGLSKAPKSLPTKSSILGGQFAIWEKWGPPQERSRLCSIALSGMLLGCFTAIL
IGGFISETLGWPFVFYIFGGVGCVCCLLWFVVIYDDPVSYPWISTSEKEYIISSLKQQVRS
SKQPLPIKAMLRSLPIWSICLGCFSHQWLVSTMVVYIPTYISSVYHVNIRDNGLLSALPFI
VAWVIGMVGGYLADFLLTKKFRLITVRKIATILGSLPSSALIVSLPYLNSGYITATALLTL
SCGLSTLCQSGIYINVLDIAPRYSSFLMGASRGFSSIAPVIVPTVSGFLLSQDPEFGWRNV
FFLLFAVNLLGLLFYLIFGEADVQEWAKERKLTRL

Dominio transmembranas.

aminoácidos 22-42, 105-125, 137-157, 236-256, 268-288, 300-320, 331-351, 366-386

Sitios de N-glicosilación.

aminoácidos 30-33, 41-44, 49-52, 58-61

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.

aminoácidos 396-399

Sitios de N-miristoilación.

aminoácidos 87-92, 112-117, 116-121, 124-129, 141-146, 144-149, 236-241, 250-255, 253-258, 278-283, 316-321, 334-339

FIGURA 114

MAGAGPKRRALAAPAAEEKEEAREKMLAAKSADGSAPAGEGEGVTLQRNITLLNGVAIIVG
TIIGSGIFVTPTGVLKEAGSPGLALVVWAACGVFSIVGALCYAELGTTISKSGGDYAYMLE
VYGSLPAFLKLWIELLIIRPSSQYIVALVFATYLLKPLFPTCPVPEEAAKLVACLCVLLLT
AVNCYSVKAATRVQDAFAAAKLLALALIILLGFVQIGKGDVSNLDPKFSFEGTKLDVGNIV
LALYSGLFAYGGWNYLNFVTEEMINPYRNLPLAIIISLPIVTLVYVLTNLAYFTTLSTEQM
LSSEAVAVDFGNYHLGVMSWIIPVFVGLSCFGSVNGSLFTSSRLFFVGSREGHLPSILSMI
HPQLLTPVPSLVFTCVMTLLYAFSKDIFSVINFFSFFNWLCVALAIIGMIWLRHRKPELER
PIKVNLALPVFFILACLFLIAVSFWKTPVECGIGFTIILSGLPVYFFGVWWKNKPKWLLQG
IFSTTVLCQKLMQVVPQET

Dominio transmembranas.

aminoácidos 48-68, 83-103, 120-140, 137-157, 165-185, 200-220, 236-256, 274-294, 316-336, 364-384, 394-414, 430-450

Sitios de N-glicosilación.

aminoácidos 49-52, 340-343

Sitio de fosforilación tirosina cinasa.

aminoácidos 112-119

Sitios de N-miristoilación.

aminoácidos 34-39, 61-66, 107-112, 114-119, 125-130, 337-342, 341-346, 353-358, 459-464, 488-493

Permeasa de aminoácidos.

aminoácidos 46-481

988 993

FIGURA 115

MAGAGPKRRALAAPAAEEKEEAREKMLAAKSADGSAPAGEGEGVTLQRNITLLNGVAIIVG
TIIGSGIFVTPTGVLKEAGSPGLALVVWAACGVFSIVGALCYAELGTTISKSGGDYAYMLE
VYGSLPAFLKLWIELLIIRPSSQYIVALVFATYLLKPLFPTCPVPEEAAKLVACLCVLLLT
AVNCYSVKAATRVQDAFAAAKLLALALIILLGFVQIGKGDVSNLDPNFSFEGTKLDVGNIV
LALYSGLFAYGGWNYLNFVTEEMINPYRNLPLAIIISLPIVTLVYVLTNLAYFTTLSTEQM
LSSEAVAVDFGNYHLGVMSWIIPVFVGLSCFGSVNGSLFTSSRLFFVGSREGHLPSILSMI
HPQLLTPVPSLVFTCVMTLLYAFSKDIFSVINFFSFFNWLCVALAIIGMIWLRHRKPELER
PIKVNLALPVFFILACLFLIAVSFWKTPVECGIGFTIILSGLPVYFFGVWWKNKPKWLLQG
IFSTTVLCQKLMQVVPQET

Dominio transmembranas.

aminoácidos 48-68, 83-103, 120-140, 137-157, 165-185, 199-219, 236-256, 274-294, 316-336, 364-384, 394-414, 430-450

Sitios de N-glicosilación.

aminoácidos 49-52, 230-233, 340-343

Sitio de fosforilación tirosina cinasa.

aminoácidos 112-119

Sitios de N-miristoilación.

aminoácidos 34-39, 61-66, 107-112, 114-119, 125-130, 337-342, 341-346, 353-358, 459-464, 488-493

Permeasas de aminoácidos.

aminoácidos 46-481

FIGURA 116

MALPTARPLLGSCGTPALGSLLFLLFSLGWVQPSRTLAGETGQEAAPLDGVLANPPNISSL
SPRQLLGFPCAEVSGLSTERVRELAVALAQKNVKLSTEQLRCLAHRLSEPPEDLDALPLDL
LLFLNPDAFSGPQACTRFFSRITKANVDLLPRGAPERQRLLPAALACWGVRGSLLSEADVR
ALGGLACDLPGRFVAESAEVLLPRLVSCPGPLDQDQQEAARAALQGGGPPYGPPSTWSVST
MDALRGLLPVLGQPIIRSIPQGIVAAWRQRSSRDPSWRQPERTILRPRFRREVEKTACPSG
KKAREIDESLIFYKKWELEACVDAALLATQMDRVNAIPFTYEQLDVLKHKLDELYPQGYPE
SVIQHLGYLFLKMSPEDIRKWNVTSLETLKALLEVNKGHEMSPQVATLIDRFVKGRGQLDK
DTLDTLTAFYPGYLCSLSPEELSSVPPSSIWAVRPQDLDTCDPRQLDVLYPKARLAFQNMN
GSEYFVKIQSFLGGAPTEDLKALSQQNVSMDLATFMKLRTDAVLPLTVAEVQKLLGPHVEG
LKAEERHRPVRDWILRQRQDDLDTLGLGLQGGIPNGYLVLDLSVQEALSGTPCLLGPGPVL
TVLALLLASTLA .

Secuencia de señal.
aminoácidos 1-33

Dominio transmembrana.
aminoácidos 603-621

Sitios de N-glicosilación.
aminoácidos 57-60, 388-391, 488-491, 515-518

Sitio de fosforilación tirosina cinasa.
aminoácidos 353-360, 471-477

Sitios de N-miristoilación.
aminoácidos 171-176, 174-179, 186-191, 266-271, 501-506, 577-582, 580-585, 581-586

Sitio de amidación.
aminoácidos 304-307

Patrón de cremallera de leucina.
aminoácidos 101-122

FIGURA 117

MALPTARPLLGSCGTPALGSLLFLLFSLGWVQPSRTLAGETGQEAAPLDGVLANPPNISSL
SPRQLLGFPCAEVSGLSTERVRELAVALAQKNVKLSTEQLRCLAHRLSEPPEDLDALPLDL
LLFLNPDAFSGPQACTRFFSRITKANVDLLPRGAPERQRLLPAALACWGVRGSLLSEADVR
ALGGLACDLPGRFVAESAEVLLPRLVSCPGPLDQDQQEAARAALQGGGPPYGPPSTWSVST
MDALRGLLPVLGQPIIRSIPQGIVAAWRQRSSRDPSWRQPERTILRPRFRREVEKTACPSG
KKAREIDESLIFYKKWELEACVDAALLATQMDRVNAIPFTYEQLDVLKHKLDELYPQGYPE
SVIQHLGYLFLKMSPEDIRKWNVTSLETLKALLEVNKGHEMSPQVATLIDRFVKGRGQLDK
DTLDTLTAFYPGYLCSLSPEELSSVPPSSIWAVRPQDLDTCDPRQLDVLYPKARLAFQNMN
GSEYFVKIQSFLGGAPTEDLKALSQQNVSMDLATFMKLRTDAVLPLTVAEVQKLLGPHVEG
LKAEERHRPVRDWILRQRQDDLDTLGLGLQGGIPNGYLVLDLSVQEALSGTPCLLGPGPVL
TVLALLLASTLA

Secuencia de señal.

aminoácidos 1-33

Dominio transmembrana.

aminoácidos 603-621

Sitios de N-glicosilación.

aminoácidos 57-60, 388-391, 488-491, 515-518

Sitio de fosforilación tirosina cinasas.

aminoácidos 353-360, 471-477

Sitios de N-miristoilación.

aminoácidos 171-176, 174-179, 186-191, 266-271, 501-506, 577-582, 580-585, 581-586

Sitio de amidación.

aminoácidos 304-307

Patrón de cremallera de leucina.

aminoácidos 101-122

FIGURA 118

MATTVPDGCRNGLKSKYYRLCDKAEAWGIVLETVATAGVVTSVAFMLTLPILVCKVQDSNR
RKMLPTQFLFLLGVLGIFGLTFAFIIGLDGSTGPTRFFLFGILFSICFSCLLAHAVSLTKL
VRGRKPLSLLVILGLAVGFSLVQDVIAIEYIVLTMNRTNVNVFSELSAPRRNEDFVLLLTY
VLFLMALTFLMSSFTFCGSFTGWKRHGAHIYLTMLLSIAIWVAWITLLMLPDFDRRWDDTI
LSSALAANGWVFLLAYVSPEFWLLTKQRNPMDYPVEDAFCKPQLVKKSYGVENRAYSQEEI
TQGFEETGDTLYAPYSTHFQLQNQPPQKEFSIPRAHAWPSPYKDYEVKKEGS

Dominio transmembranas.
aminoácidos 30-50, 66-86, 98-118, 121-141, 137-157, 174-194, 212-232, 244-264

Sitio de N-glicosilación.
aminoácidos 158-161

Sitios de N-miristoilación.
aminoácidos 8-13, 38-43, 80-85, 88-93, 102-107, 136-141, 201-206

Sitio de amidación.
aminoácidos 124-127

Receptor de transmembrana 7.
aminoácidos 27-273

FIGURA 119

MIPFLPMFSLLLLLIVNPINANNHYDKILAHSRIRGRDQGPNVCALQQILGTKKKYFSTCK
NWYKKSICGQKTTVLYECCPGYMRMEGMKGCPAVLPIDHVYGTLGIVGATTTQRYSDASKL
REEIEGKGSFTYFAPSNEAWDNLDSDIRRGLESNVNVELLNALHSHMINKRMLTKDLKNGM
IIPSMYNNLGLFINHYPNGVVTVNCARIIHGNQIATNGVVHVIDRVLTQIGTSIQDFIEAE
DDLSSFRAAAITSDILEALGRDGHFTLFAPTNEAFEKLPRGVLERFMGDKVASEALMKYHI
LNTLQCSESIMGGAVFETLEGNTIEIGCDGDSITVNGIKMVNKKDIVTNNGVIHLIDQVLI
PDSAKQVIELAGKQQTTFTDLVAQLGLASALRPDGEYTLLAPVNNAFSDDTLSMVQRLLKL
ILQNHILKVKVGLNELYNGQILETIGGKQLRVFVYRTAVCIENSCMEKGSKQGRNGAIHIF
REIIKPAEKSLHEKLKQDKRFSTFLSLLEAADLKELLTQPGDWTLFVPTNDAFKGMTSEEK
EILIRDKNALQNIILYHLTPGVFIGKGFEPGVTNILKTTQGSKIFLKEVNDTLLVNELKSK
ESDIMTTNGVIHVVDKLLYPADTPVGNDQLLEILNKLIKYIQIKFVRGSTFKEIPVTVYTT
KIITKVVEPKIKVIEGSLQPIIKTEGPTLTKVKIEGEPEFRLIKEGETITEVIHGEPIIKK
YTKIIDGVPVEITEKETREERIITGPEIKYTRISTGGGETEETLKKLLQEEVTKVTKFIEG
GDGHLFEDEEIKRLLQGDTPVRKLQANKKVQGSRRRLREGRSQ

Secuencia de señal.
aminoácidos 1-21

Sitio de N-glicosilación.
aminoácidos 599-602

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.
aminoácidos 507-510, 731-734

Sitios de N-miristoilación.
aminoácidos 70-75, 106-111, 109-114, 152-157, 193-198, 214-219, 392-397, 476-481, 570-575, 768-773

Dominio de fasciclina.
aminoácidos 94-232, 496-630

993

FIGURA 120

MIPFLPMFSLLLLLIVNPINANNHYDKILAHSRIRGRDQGPNVCALQQILGTKKKYFSTCK
NWYKKSICGQKTTVLYECCPGYMRMEGMKGCPAVLPIDHVYGTLGIVGATTTQRYSDASKL
REEIEGKGSFTYFAPSNEAWDNLDSDIRRGLESNVNVELLNALHSHMINKRMLTKDLKNGM
IIPSMYNNLGLFINHYPNGVVTVNCARIIHGNQIATNGVVHVIDRVLTQIGTSIQDFIEAE
DDLSSFRAAAITSDILEALGRDGHFTLFAPTNEAFEKLPRGVLERIMGDKVASEALMKYHI
LNTLQCSESIMGGAVFETLEGNTIEIGCDGDSITVNGIKMVNKKDIVTNNGVIHLIDQVLI
PDSAKQVIELAGKQQTTFTDLVAQLGLASALRPDGEYTLLAPVNNAFSDDTLSMDQRLLKL
ILQNHILKVKVGLNELYNGQILETIGGKQLRVFVYRTAVCIENSCMEKGSKQGRNGAIHIF
REIIKPAEKSLHEKLKQDKRFSTFLSLLEAADLKELLTQPGDWTLFVPTNDAFKGMTSEEK
EILIRDKNALQNIILYHLTPGVFIGKGFEPGVTNILKTTQGSKIFLKEVNDTLLVNELKSK
ESDIMTTNGVIHVVDKLLYPADTPVGNDQLLEILNKLIKYIQIKFVRGSTFKEIPVTVYRP
TLTKVKIEGEPEFRLIKEGETITEVIHGEPIIKKYTKIIDGVPVEITEKETREERIITGPE
IKYTRISTGGGETEETLKKLLQEEVTKVTKFIEGGDGHLFEDEEIKRLLQGDTPVRKLQAN
KKVQGSRRRLREGRSQ

Secuencia de señal.
aminoácidos 1-21

Sitio de N-glicosilación.
aminoácidos 599-602

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.
aminoácidos 507-510, 704-707

Sitios de N-miristoilación.
aminoácidos 70-75, 106-111, 109-114, 152-157, 193-198, 214-219, 392-397, 476-481, 570-575, 741-746

Dominio de fasciclinas.
aminoácidos 94-232, 234-367, 370-494, 496-630

FIGURA 121

MIPFLPMFSLLLLLIVNPINANNHYDKILAHSRIRGRDQGPNVCALQQILGTKKKYFSTCK
NWYKKSICGQKTTVLYECCPGYMRMEGMKGCPAVLPIDHVYGTLGIVGATTTQRYSDASKL
REEIEGKGSFTYFAPSNEAWDNLDSDIRRGLESNVNVELLNALHSHMINKRMLTKDLKNGM
IIPSMYNNLGLFINHYPNGVVTVNCARIIHGNQIATNGVVHVIDRVLTQIGTSIQDFIEAE
DDLSSFRAAAITSDILEALGRDGHFTLFAPTNEAFEKLPRGVLERIMGDKVASEALMKYHI
LNTLQCSESIMGGAVFETLEGNTIEIGCDGDSITVNGIKMVNKKDIVTNNGVIHLIDQVLI
PDSAKQVIELAGKQQTTFTDLVAQLGLASALRPDGEYTLLAPVNNAFSDDTLSMDQRLLKL
ILQNHILKVKVGLNELYNGQILETIGGKQLRVFVYRTAVCIENSCMEKGSKQGRNGAIHIF
REIIKPAEKSLHEKLKQDKRFSTFLSLLEAADLKELLTQPGDWTLFVPTNDAFKGMTSEEK
EILIRDKNALQNIILYHLTPGVFIGKGFEPGVTNILKTTQGSKIFLKEVNDTLLVNELKSK
ESDIMTTNGVIHVVDKLLYPADTPVGNDQLLEILNKLIKYIQIKFVRGSTFKEIPVTVYKP
IIKKYTKIIDGVPVEITEKETREERIITGPEIKYTRISTGGGETEETLKKLLQEDTPVRKL
QANKKVQGSRRRLREGRSQ

Secuencia de señal.
aminoácidos 1-21

Sitio de N-glicosilación.
aminoácidos 599-602

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasas.
aminoácidos 507-510, 674-677

Sitios de N-miristoilación.
aminoácidos 70-75, 106-111, 109-114, 152-157, 193-198, 214-219, 392-397, 476-481, 570-575, 711-716

Dominio de fasciclinas.
aminoácidos 94-232, 234-367, 370-494, 496-630

995

FIGURA 122

MIPFLPMFSLLLLLIVNPINANNHYDKILAHSRIRGRDQGPNVCALQQILGTKKKYFSTCK
NWYKKSICGQKTTVLYECCPGYMRMEGMKGCPAVLPIDHVYGTLGIVGATTTQRYSDASKL
REEIEGKGSFTYFAPSNEAWDNLDSDIRRGLESNVNVELLNALHSHMINKRMLTKDLKNGM
IIPSMYNNLGLFINHYPNGVVTVNCARIIHGNQIATNGVVHVIDRVLTQIGTSIQDFIEAE
DDLSSFRAAAITSDILEALGRDGHFTLFAPTNEAFEKLPRGVLERIMGDKVASEALMKYHI
LNTLQCSESIMGGAVFETLEGNTIEIGCDGDSITVNGIKMVNKKDIVTNNGVIHLIDQVLI
PDSAKQVIELAGKQQTTFTDLVAQLGLASALRPDGEYTLLAPVNNAFSDDTLSMDQRLLKL
ILQNHILKVKVGLNELYNGQILETIGGKQLRVFVYRTAVCIENSCMEKGSKQGRNGAIHIF
REIIKPAEKSLHEKLKQDKRFSTFLSLLEAADLKELLTQPGDWTLFVPTNDAFKGMTSEEK
EILIRDKNALQNIILYHLTPGVFIGKGFEPGVTNILKTTQGSKIFLKEVNDTLLVNELKSK
ESDIMTTNGVIHVVDKLLYPADTPVGNDQLLEILNKLIKYIQIKFVRGSTFKEIPVTVYRP
TLTKVKIEGEPEFRLIKEGETITEVIHGEPIIKKYTKIIDGVPVEITEKETREERIITGPE
IKYTRISTGGGETEETLKKLLQEDTPVRKLQANKKVQGSRRRLREGRSQ

Secuencia de señal.

aminoácidos 1-21

Sitio de N-glicosilación.

aminoácidos 599-602

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasas.

aminoácidos 507-510, 704-707

Sitios de N-miristoilación.

aminoácidos 70-75, 106-111, 109-114, 152-157, 193-198, 214-219, 392-397, 476-481, 570-575, 741-746

Dominio de fasciclinas.

aminoácidos 94-232, 234-367, 370-494, 496-630

FIGURA 123

MIPFLPMFSLLLLLLIVNPINANNHYDKILAHSRIRGRDQGPNVCALQQILGTKKKYFSTCK
NWYKKSICGQKTTVLYECCPGYMRMEGMKGCPAVLPIDHVYGTLGIVGATTTQRYSDASKL
RREIEGKGSFTYFAPSNEAWDNLDSDIRRGLESNVNVELLNALHSHMINKRMLTKDLKNGM
IIPSMYNNLGLFINHYPNGVVTVNCARIIHGNQIATNGVVHVIDRVLTQIGTSIQDFIEAE
DDLSSFRAAAITSDILEALGRDGHFTLFAPTNEAFEKLPRGVLERIMGDKVASEALMKYHI
LNTLQCSESIMGGAVFETLEGNTIEIGCDGDSITVNGIKMVNKKDIVTNNGVIHLIDQVLI
PDSAKQVIELAGKQQTTFTDLVAQLGLASALRPDGEYTLLAPVNNAFSDDTLSMDQRLLKL
ILQNHILKVKVGLNELYNGQILETIGGKQLRVFVYRTAVCIENSCMEKGSKQGRNGAIHIF
REIIKPAEKSLHEKLKQDKRFSTFLSLLEAADLKELLTQPGDWTLFVPTNDAFKGMTSEEK
EILIRDKNALQNIILYHLTPGVFIGKGFEPGVTNILKTTQGSKIFLKEVNDTLLVNELKSK
ESDIMTTNGVIHVVDKLLYPADTPVGNDQLLEILNKLIKYIQIKFVRGSTFKEIPVTVYKP
IIKKYTKIIDGVPVEITEKETREERIITGPEIKYTRISTGGGETEETLKKLLQEEVTKVTK
FIEGGDGHLFEDEEIKRLLQGDTPVRKLQANKKVQGSRRRLREGRSQ

Secuencia de señal.
aminoácidos 1-21

Sitio de N-glicosilación.
aminoácidos 599-602

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasas.
aminoácidos 507-510, 674-677

Sitios de N-miristoilación.
aminoácidos 70-75, 106-111, 109-114, 152-157, 193-198, 214-219, 392-397, 476-481, 570-575, 711-716

Dominio de fasciclinas.
aminoácidos 94-232, 234-367, 370-494, 496-630

FIGURA 124

MVNYAWAGRSQRKLWWRSVAVLTCKSVVRPGYRGGLQARRSTLLKTCARARATAPGAMKMV
APWTRFYSNSCCLCCHVRTGTILLGVWYLIINAVVLLILLSALADPDQYNFSSSSELGGDFE
FMDDANMCIAIAISLLMILICAMATYGAYKQRAAWIIPFFCYQIFDFALNMLVAITVLIYP
NSIQEYIRQLPPNFPYRDDVMSVNPTCLVLIILLFISIILTFKGYLISCVWNCYRYINGRN
SSDVLVYVTSNDTTVLLPPYDDATVNGAAKEPPPPYVSA

Dominio transmembranas.

aminoácidos 83-103, 124-144, 162-182, 205-225

Sitios de N-glicosilación.

aminoácidos 111-114, 244-227, 255-258

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.

aminoácidos 39-42

Sitio de fosforilación tirosina cinasa.

aminoácidos 243-251

Sitio de N-miristoilación.

aminoácidos 34-39

FIGURA 125

MPGQGGLQARRSTLLKTCARARATAPGAMKMVAPWTRFYSNSCCLCCHVRTGTILLGVWYL
IINAVVLLILLSALADPDQYNFSSSELGGDFEFMDDANMCIAIAISLLMILICAMATYGAY
KQRAAWIIPFFCYQIFDFALNMLVAITVLIYPNSIQEYIRQLPPNFPYRDDVMSVNPTCLV
LIILLFISIILTFKGYLISCVWNCYRYINGRNSSDVLVYVTSNDTTVLLPPYDDATVNGAA
KEPPPPYVSA

...

Dominio transmembranas.

aminoácidos 54-74, 95-115, 133-153, 176-196

Sitios de N-glicosilación.

aminoácidos 82-85, 215-218, 226-229

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.

aminoácidos 10-13

Sitio de fosforilación tirosina cinasa.

aminoácidos 214-222

...

Sitio de N-miristoilación.

aminoácidos 5-10

...

999

FIGURA 126

MGRAGGGGPGRGPPPLLLFLGAALVLASGAVPARRAGSAVEAEELVKGSPAWEPPANDTRE
EAGPPAAGEDEASWTAPGGELAGPEEVLQESAAVTGTAWLEADSPGLGGVTAEAGSGDAQA
LPATLQAPHEVLGQSIMPPAIPEATEASGPPSPTPGDKLSPASELPKESPLEVWLNLGGST
PDPQVPELTYPFQGTLEPQPASDIIDIDYFEGLDGEGRGADLGSFPGSPGTSENHPDTEGE
TPSWSLLDLYDDFTPFDESDFYPTTSFYDDLDEEEEEEDDKDAVGGGDLEDENELLVPTG
KPGLGPGTGQPTSRWHAVPPQHTLGSVPGSSIALRPRPGEPGRDLASSENGTECRSGFVRH
NGSCRSVCDLFPSYCHNGGQCYLVENIGAFCRCNTQDYIWHKGMRCESIITDFQVMCVAVG
SAALVLLLLFMMTVFFAKKLYLLKTENTKLRRTNKFRTPSELHNDNFSLSTIAEGSHPNDD
PSAPHKIQEVLKSCLKEEESFNIQNSMSPKLEGGKGDQADLDVNCLQNNLT

Secuencia de señal.
aminoácidos 1-30

Dominio transmembrana.
aminoácidos 423-443

Sitios de N-glicosilación.
aminoácidos 57-60, 355-358, 367-370, 473-476

Sitios de N-miristoilación.
aminoácidos 6-11, 29-34, 79-84, 109-114, 116-121, 222-227, 226-231, 230-235, 233-238, 308-313, 314-319, 330-335, 334-339, 368-373, 482-487

FIGURA 127

MLPVYQEVKPNPLQDANICSRVFFWWLNPLFKIGHKRRLEEDDMYSVLPEDRSQHLGEELQ
GFWDKEVLRAENDAQKPSLTRAIIKCYWKSYLVLGIFTLIEESAKVIQPIFLGKIINYFEN
YDPMDSVALNTAYAYATVLTFCTLILAILEHLYFYHVQCAGMRLRVAMCHMIYRKALRLSN
MAMGKTTTGQIVNLLSNDVNKFDQVTVFLHFLWAGPLQAIAVTALLWMEIGISCLAGMAVL
IILLPLQSCFGKLFSSLRSKTATFTDARIRTMNEVITGIRIIKMYAWEKSFSNLITNLRKK
EISKILRSSCLRGMNLASFFSASKIIVFVTFTTYVLLGSVITASRVFVAVTLYGAVRLTVT
LFFPSAIERVSEAIVSIRRIQTFLLLDEISQRNRQLPSDGKKMVHVQDFTAFWDKASETPT
LQGLSFTVRPGELLAVVGPVGAGKSSLLSAVLGELAPSHGLVSVHGRIAYVSQQPWVFSGT
LRSNILFGKKYEKERYEKVIKACALKKDLQLLEDGDLTVIGDRGTTLSGGQKARVNLARAV
YQDADIYLLDDPLSAVDAEVSRHLFELCICQILHEKITILVTHQLQYLKAASQILILKDGK
MVQKGTYTEFLKSGIDFGSLLKKDNEESEQPPVPGTPTLRNRTFSESSVWSQQSSRPSLKD
GALESQDTENVPVTLSEENRSEGKVGFQAYKNYFRAGAHWIVFIFLILLNTAAQVAYVLQD
WWLSYWANKQSMLNVTVNGGGNVTEKLDLNWYLGIYSGLTVATVLFGIARSLLVFYVLVNS
SQTLHNKMFESILKAPVLFFDRNPIGRILNRFSKDIGHLDDLLPLTFLDFIQTLLQVVGVV
SVAVAVIPWIAIPLVPLGIIFIFLRRYFLETSRDVKRLESTTRSPVFSHLSSSLQGLWTIR
AYKAEERCQELFDAHQDLHSEAWFLFLTTSRWFAVRLDAICAMFVIIVAFGSLILAKTLDA
GQVGLALSYALTLMGMFQWCVRQSAEVENMMISVERVIEYTDLEKEAPWEYQKRPPPAWPH
EGVIIFDNVNFMYSPGGPLVLKHLTALIKSQEKVGIVGRTGAGKSSLISALFRLSEPEGKI
WIDKILTTEIGLHDLRKKMSIIPQEPVLFTGTMRKNLDPFKEHTDEELWNALQEVQLKETI
EDLPGKMDTELAESGSNFSVGQRQLVCLARAILRKNQILIIDRATANVDPRTDELIQKKIR
EKFAHCTVLTIAHRLNTIIDSDKIMVLDSGRLKEYDEPYVLLQNKESLFYKMVQQLGKAEA
AALTETAKQVYFKRNYPHIGHTDHMVTNTSNGQPSTLTIFETAL

Dominio transmembranas:

aminoácidos 134-154, 233-253, 346-366, 708-728, 767-787, 834-854, 857-877, 948-968, 967-987

Sitios de N-glicosilación.

aminoácidos 651-654, 690-693, 746-749, 754-757, 792-795, 1176-1179, 1309-1312

Sitio de fosforilación cAMP- y cGMP- dependiente de proteina cinasa.

aminoácidos 1115-1118

Sitio de fosforilación tirosina cinasas.

aminoácidos 37-45, 497-504, 1253-1259

Sitios de N-miristoilación.

aminoácidos 192-197, 318-323, 343-348, 430-435, 448-453, 487-492, 532-537, 537-542, 624-629, 672-677, 752-575,

766-771, 770-775, 779-784, 980-985, 1078-1083, 1174-1179, 1313-1318

Sitios de amidación.
aminoácidos 405-410, 495-500

Sitio de unión ATP/GTP motif A (Vuelta- P).
aminoácidos 445-452, 1075-1082

Transportador ABC.
aminoácidos 438-609, 1068-1250

Región transportadora de membrana ABC.
aminoácidos 92-365, 711-997

1002

FIGURA 128

MSALRRKFGDDYQVVTTSSSGSGLQPQGPGQDPQQQLVPKKKRQRFVDKNGRCNVQHGNLG
SETSRYLSDLFTTLVDLKWRWNLFIFILTYTVAWLFMASMWWVIAYTRGDLNKAHVGNYTP
CVANVYNFPSAFLFFIETEATIGYGYRYITDKCPEGIILFLFQSILGSIVDAFLIGCMFIK
MSQPKKRAETLMFSEHAVISMRDGKLTLMFRVGNLRNSHMVSAQIRCKLLKSRQTPEGEFL
PLDQLELDVGFSTGADQLFLVSPLTICHVIDAKSPFYDLSQRSMQTEQFEIVVILEGIVET
TGMTCQARTSYTEDEVLWGHRFFPVISLEEGFFKVDYSQFHATFEVPTPPYSVKEQEEMLL
MSSPLIAPAITNSKERHNSVECLDGLDDITTKLPSKLQKITGREDFPKKLLRMSSTTSEKA
YSLGDLPMKLQRISSVPGNSEEKLVSKTTKMLSDPMSQSVADLPPKLQKMAGGAARMEGNL
PAKLRKMNSDRFT

Dominio transmembranas.
aminoácidos 83-103, 120-140, 158-178

Sitio de fijación de glicosaminoglicano.
aminoácidos 20-23

Sitios de N-miristoilación.
aminoácidos 58-63, 61-66, 216-221, 301-306, 486-491

Secuencia de adhesión de células.
aminoácidos 109-114

Canal de rectificación de entrada de potasio.
aminoácidos 47-393

1603

FIGURA 129

MPLFKNTSVSSLYSGCRLTLLRPEKDGAATRVDAVCTHRPDPKSPGLDRERLYWKLSQLTH
GITELGPYTLDRHSLYVNGFTHQSSMTTTRTPDTSTMHLATSRTPASLSGPTTASPLLVLF
TINFTITNLRYEENMHHPGSRKFNTTERVLQGLLRPVFKNTSVGPLYSGCRLTLLRPKKDG
AATKVDAICTYRPDPKSPGLDREQLYWELSQLTHSITELGPYTLDRDSLYVNGFTQRSSVP
TTSIPGTPTVDLGTSGTPVSKPGPSAASPLLVLFTLNFTITNLRYEENMQHPGSRKFNTTE
RVLQGLLRSLFKSTSVGPLYSGCRLTLLRPEKDGTATGVDAICTHHPDPKSPRLDREQLYW
ELSQLTHNITELGHYALDNDSLFVNGFTHRSSVSTTSTPGTPTVYLGASKTPASIFGPSAA
SHLLILFTLNFTITNLRYEENMWPGSRKFNTTERVLQGLLRPLFKNTSVGPLYSGSRLTLL
RPEKDGEATGVDAICTHRPDPTGPGLDREQLYLELSQLTHSITELGPYTLDRDSLYVNGFT
HRSSVPTTSTGVVSEEPFTLNFTINNLRYMADMGQPGSLKFNITDNVMKHLLSPLFQRSSL
GARYTGCRVIALRSVKNGAETRVDLLCTYLQPLSGPGLPIKQVFHELSQQTHGITRLGPYS
LDKDSLYLNGYNEPGLDEPPTTPKPATTFLPPLSEATTAMGYHLKTLTLNFTISNLQYSPD
MGKGSATFNSTEGVLQHLLRPLFQKSSMGPFYLGCQLISLRPEKDGAATGVDTTCTYHPDP
VGPGLDIQQLYWELSQLTHGVTQLGFYVLDRDSLFINGYAPQNLSIRGEYQINFHIVNWNL
SNPDPTSSEYITLLRDIQDKVTTLYKGSQLHDTFRFCLVTNLTMDSVLVTVKALFSSNLDP
SLVEQVFLDKTLNASFHWLGSTYQLVDIHVTEMESSVYQPTSSSSTQHFYPNFTITNLPYS
QDKAQPGTTNYQRNKRNIEDALNQLFRNSSIKSYFSDCQVSTFRSVPNRHHTGVDSLCNFS
PLARRVDRVAIYEEFLRMTRNGTQLQNFTLDRSSVLVDGYSPNRNEPLTGNSDLPFWAVIF
IGLAGLLGLITCLICGVLVTTRRRKKEGEYNVQQQCPGYYQSHLDLEDLQ

Dominio transmembranas.

aminoácidos 107-127, 418-438, 1093-1113

Sitios de N-glicosilación.

aminoácidos 6-9, 125-128, 146-149, 162-165, 281-284, 302-305, 374-377, 385-388, 437-440, 457-460, 473-476, 570-573, 591-594, 721-724, 741-744, 836-839, 853-856, 895-898, 928-931, 967-970, 1004-1007, 1058-1061, 1064-1067

Sitio de fijación de glicosaminoglicano.

aminoácidos 644-647

Sitio de fosforilación tirosina cinasas.

aminoácidos 358-365, 1121-1128

Sitios de N-miristoilación.

aminoácidos 15-20, 171-176, 260-265, 310-315, 327-332, 339-344, 482-487, 583-588, 611-616, 778-783, 782-787, 1106-1111, 1114-1119

Dominio SEA.

aminoácidos 112-245, 270-391, 661-684

FIGURA 130

MGNWVVNHWFSVLFLVVWLGLNVFLFVDAFLKYEKADKYYYTRKILGSTLACARASALCLN
FNSTLILLPVCRNLLSFLRGTCSFCSRTLRKQLDHNLTFHKLVAYMICLHTAIHIIAHLFN
FDCYSRSRQATDGSLASILSSLSHDEKKGGSWLNPIQSRNTTVEYVTFTSVAGLTGVIMTI
ALILMVTSATEFIRRSYFEVFWYTHHLFIFYILGLGIHGIGGIVRGQTEESMNESHPRKCA
ESFEMWDDRDSRCRRPKFEGHPPESWKWILAPVILYICERILRFYRSQQKVVITKVVMHPS
KVLELQMNKRGFSMEVGQYIFVNCPSISLLEWHPFTLTSAPEEDFFSIHIRAAGDWTENLI
RAFEQQYSPIPRIEVDGPFGTASEDVFQYEVAVLVGAGIGVTPFASILKSIWYKFQCADHN
LKTKKIYFYWICRETGAFSWFNNLLTSLEQEMEELGKVGFLNYRLFLTGWDSNIVGHAALN
FDKATDIVTGLKQKTSFGRPMWDNEFSTIATSHPKSVVGVFLCGPRTLAKSLRKCCHRYSS
LDPRKVQFYFNKENF

Secuencia de señal.

aminoácidos 1-29

Dominio transmembrana.

aminoácidos 44-64, 60-80, 99-119, 168-188, 204-224

Sitios de N-glicosilación.

aminoácidos 63-66, 97-100, 162-165, 236-239

Sitio de fosforilación tirosina cinasas.

aminoácidos 198-206, 367-373

Sitios de N-miristoilación.

aminoácidos 47-52, 135-140, 178-183, 225-230, 402-407, 527-532

De tipo reductasa férrica.

aminoácidos 46-563

FIGURA 131

MGNWVVNHWFSVLFLVVWLGLNVFLFVDAFLKYEKADKYYYTRKILGSTLACARASALCLN
FNSTLILLPVCRNLLSFLRGTCSFCSRTLRKQLDHNLTFHKLVAYMICLHTAIHIIAHLFN
FDCYSRSRQATDGSLASILSSLSHDEKKGGSWLNPIQSRNTTVEYVTFTSIAGLTGVIMTI
ALILMVTSATEFIRRSYFEVFWYTHHLFIFYILGLGIHGIGGIVRGQTEESMNESHPRKCA
ESFEMWDDRDSHCRRPKFEGHPPESWKWILAPVILYICERILRFYRSQQKVVITKVVMHPS
KVLELQMNKRGFSMEVGQYIFVNCPSISLLEWHPFTLTSAPEEDFFSIHIRAAGDWTENLI
RAFEQQYSPIPRIEVDGPFGTASEDVFQYEVAVLVGAGIGVTPFASILKSIWYKFQCADHN
LKTKKIYFYWICRETGAFSWFNNLLTSLEQEMEELGKVGFLNYRLFLTGWDSNIVGHAALN
FDKATDIVTGLKQKTSFGRPMWDNEFSTIATSHPKSVVGVFLCGPRTLAKSLRKCCHRYSS
LDPRKVQFYFNKENF

Dominio transmembranas.

aminoácidos 6-26, 44-64, 60-80, 99-119, 168-188, 204-224

Sitios de N-glicosilación.

aminoácidos 63-66, 97-100, 162-165, 236-239

Sitio de fosforilación tirosina cinasas.

aminoácidos 198-206, 367-373

Sitios de N-miristoilación.

aminoácidos 47-52, 135-140, 178-183, 225-230, 402-407, 527-532

Transmembrana de tipo reductasa férrica.

aminoácidos 46-563

1007 tara

FIGURA 132

MGNWVVNHWFSVLFLVVWLGLNVFLFVDAFLKYEKADKYYYTRKILGSWKWILAPVILYIC
ERILRFYRSQQKVVITKVVMHPSKVLELQMNKRGFSMEVGQYIFVNCPSISLLEWHPFTLT
SAPEEDFFSIHIRAAGDWTENLIRAFEQQYSPIPRIEVDGPFGTASEDVFQYEVAVLVGAG
IGVTPFASILKSIWYKFQCADHNLKTKKVGHAALNFDKATDIVTGLKQKTSFGRPMWDNEF
STIATSHPKSVVGVFLCGPRTLAKSLRKCCHRYSSLDPRKVQFYFNKENFQ

Dominio transmembranas.
aminoácidos 6-26, 44-64

Sitio de fosforilación tirosina cinasa.
aminoácidos 146-152

Sitios de N-miristoilación.
aminoácidos 181-186, 257-262

Transmembrana de tipo reductasa férrica.
aminoácidos 1-293

FIGURA 133

MGNWVVNHWFSVLFLVVWLGLNVFLFVDAFLKYEKADKYYYTRKILGSTLACARASALCLN
FNSTLILLPVCRNLLSFLRGTCSFCSRTLRKQLDHNLTFHKLVAYMICLHTAIHIIAHLFN
FDCYSRSRQATDGSLASILSSLSHDEKKGGSWLNPIQSRNTTVEYVTFTSIAGLTGVIMTI
ALILMVTSATEFIRRSYFEVFWYTHHLFIFYILGLGIHGIGGIVRGQTEESMNESHPRKCA
ESFEMWDDRDSHCRRPKFEGHPPESWKWILAPVILYICERILRFYRSQQKVVITKVVMHPS
KVLELQMNKRGFSMEVGQYIFVNCPSISLLEWHPFTLTSAPEEDFFSIHIRAAGDWTENLI
RAFEQQYSPIPRIEVDGPFGTASEDVFQYEVAVLVGAGIGVTPFASILKSIWYKFQCADHN
LKTKKVGHAALNFDKATDIVTGLKQKTSFGRPMWDNEFSTIATSHPKSVVGVFLCGPRTLA
KSLRKCCHRYSSLDPRKVQFYFNKENF

Dominio transmembranas.

aminoácidos 6-26, 44-64, 60-80, 99-119, 168-188, 204-224

Sitios de N-glicosilación.

aminoácidos 63-66, 97-100, 162-165, 236-239

Sitio de fosforilación tirosina cinasas.

aminoácidos 198-206, 367-373

Sitios de N-miristoilación.

aminoácidos 47-52, 135-140, 178-183, 225-230, 402-407, 478-483

Transmembrana de tipo reductasa férrica.

aminoácidos 46-514

FIGURA 134

MTSGSVFFYILIFGKYFSHGGGQDVKCSLGYFPCGNITKCLPQLLHCNGVDDCGNQADEDN
CGDNNGWSMQFDKYFASYYKMTSQYPFEAETPECLVGSVPVQCLCQGLELDCDETNLRAVP
SVSSNVTAMSLQWNLIRKLPPDCFKNYHDLQKLYLQNNKITSISIYAFRGLNSLTKLYLSH
NRITFLKPGVFEDLHRLEWLIIEDNHLSRISPPTFYGLNSLILLVLMNNVLTRLPDKPLCQ
HMPRLHWLDLEGNHIHNLRNLTFISCSNLTVLVMRKNKINHLNENTFAPLQKLDELDLGSN
KIENLPPLIFKDLKELSQLNLSYNPIQKIQANQFDYLVKLKSLSLEGIEISNIQQRMFRPL
MNLSHIYFKKFQYCGYAPHVRSCKFNTDGISSLENLLASIIQRVFVWVVSAVTCFGNIFVI
CMRPYIRSENKLYAMSIISLCCADCLMGIYLFVIGGFDLKFRGEYNKHAQLWMESTHCQLV
GSLAILSTEVSVLLLTFLTLEKYICIVYPFRCVRPGKCRTITVLILIWITGFIVAFIPLSN
KEFFKNYYGTNGVCFPLHSEDTESIGAQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSA
ITATEIRNQVKKEMILAKRFFFIVFTDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILP
INSALNPILYTLTTRPFKEMIHRFWYNYRQRKSMDSKGQKTYAPSFIWVEMWPLQEMPPEL
MKPDLFTYPCEMSLISQSTRLNSYS

Dominio transmembranas.

aminoácidos 212t232, 404-424, 443-463, 489-509, 526-546, 577-597, 629-649, 665-685

Sitios de N-glicosilación.

aminoácidos 36-39, 127-130, 264-267, 272-275, 325-328, 368-371

Sitios de N-miristoilación.

aminoácidos 49-54, 352-357, 586-591, 659-664

Patrón de cremallera de leucina.

aminoácidos 487-508

Receptor-LDL de clase A.

aminoácidos 40-59

Acoplamiento receptor proteína-G.

aminoácidos 159-189, 413-432, 422-452, 436-466, 478-518, 528-547, 625-644, 673-689

Repeticiones ricas en leucina.

aminoácidos 127-150, 151-174, 175-198, 199-222, 223-246, 248-271, 272-295, 296-319, 320-343, 344-367

Receptor de transmembrana 7 (proteínas de la familia rodopsina).

aminoácidos 485-681

1010

Dominio de receptor de lipoproteína de baja densidad.
aminoácidos 25-64

FIGURA 135

MTSGSVFFYILIFGKYFSHGGGQDVKCSLGYFPCGNITKCLPQLLHCNGVDDCGNQADEDN
CGDNNGWSLQFDKYFASYYKMTSQYPFEAETPECLVGSVFVQCLCQGLELDCDETNLRAVP
SVSSNVTAMSLQWNLIRKLPPDCFKNYHDLQKLYLQNNKITSISIYAFRGLNSLTKLYLSH
NRITFLKPGVFEDLHRLEWLIIEDNHLSRISPPTFYGLNSLILLVLMNNVLTRLPDKPLCQ
HMPRLHWLDLEGNHIHNLRNLTFISCSNLTVLVMRKNKINHLNENTFAPLQKLDELDLGSN
KIENLPPLIFKDLKELSQLNLSYNPIQKIQANQFDYLVKLKSLSLEGIEISNIQQRMFRPL
MNLSHIYFKKFQYCGYAPHVRSCKPNTDGISSLENLLASIIQRVFVWVVSAVTCFGNIFVI
CMRPYIRSENKLYAMSIISLCCADCLMGIYLFVIGGFDLKFRGEYNKHAQLWMESTHCQLV
GSLAILSTEVSVLLLTFLTLEKYICIVYPFRCVRPGKCRTITVLILIWITGFIVAFIPLSN
KEFFKNYYGTNGVCFPLHSEDTESIGAQIYSVAIFLGINLAAFIIIVFSYGSMFYSVHQSA
ITATEIRNQVKKEMILAKRFFFIVFTDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILP
INSALNPILYTLTTRPFKEMIHRFWYNYRQRKSMDSKGQKTYAPSFIWVEMWPLQEMPPEL
MKPDLFTYPCEMSLISQSTRLNSYS

Dominio transmembranas.
aminoácidos 212-232, 404-424, 443-463, 489-509, 526-546, 577-597, 629-S49, 665-685

Sitios de N-glicosilación.
aminoácidos 36-39, 127-130, 264-267, 272-275, 325-328, 368-371

Sitios de N-miristoilación.
aminoácidos 49-54, 352-357, 586-591, 659-664

Patrón de cremallera de leucina.
aminoácidos 487-508

Repeticiones ricas en leucina:
aminoácidos 127-150, 151-174, 175-198, 199-222, 223-246, 248-271, 272-295, 296-319, 320-343, 344-367

Receptor de transmembrana 7 (familia de rodopsinas).
aminoácidos 59-259

Dominio de receptor de lipoproteína de baja densidad.
aminoácidos 25-64

FIGURA 136

MGRLASRPLLLALLSLALCRGRVVRVPTATLVRVVGTELVIPCNVSDYDGPSEQNFDWSFS
SLGSSFVELASTWEVGFPAQLYQERLQRGEILLRRTANDAVELHIKNVQPSDQGHYKCSTP
STDATVQGNYEDTVQVKVLADSLHVGPSARPPPSLSLREGEPFELRCTAASASPLHTHLAL
LWEVHRGPARRSVLALTHEGRFHPGLGYEQRYHSGDVRLDTVGSDAYRLSVSRALSADQGS
YRCIVSEWLAEQGNWQEIQEKAVEVATVVIQPTVLRAAVPKNVSVAEGKELDLTCNITTDR
ADDVRPEVTWSFSRMPDSTLPGSRVLARLDRDSLVHSSPHVALSHVDARSYHLLVRDVSKE
NSGYYYCHVSLWAPGHNRSWHKVAEAVSSPAGVGVTWLEPDYQVYLNASKVPGFADDPTEL
ACRVVDTKSGEANVRFTVSWYYRMNRRSDNVVTSELLAVMDGDWTLKYGERSKQRAQDGDF
IFSKEHTDTFNFRIQRTTEEDRGNYYCVVSAWTKQRNNSWVKSKDVFSKPVNIFWALEDSV
LVVKARQPKPFFAAGNTFEMTCKVSSKNIKSPRYSVLIMAEKPVGDLSSPNETKYIISLDQ
DSVVKLENWTDASRVDGVVLEKVQEDEFRYRMYQTQVSDAGLYRCMVTAWSPVRGSLWREA
ATSLSNPIEIDFQTSGPIFNASVHSDTPSVIRGDLIKLFCIITVEGAALDPDDMAFDVSWF
AVHSFGLDKAPVLLSSLDRKGIVTTSRRDWKSDLSLERVSVLEFLLQVHGSEDQDFGNYYC
SVTPWVKSPTGSWQKEAEIHSKPVFITVKMDVLNAFKYPLLIGIGLSTVIGLLSCLIGYCS
SHWCCKKEVQETRRERRRLMSMEMD

Secuencia de señal.
aminoácidos 1-21

Dominio transmembrana.
aminoácidos 833-853

Sitios de N-glicosilación.
aminoácidos 44-47, 286-289, 300-303, 383-386, 413-416, 525-528, 600-603, 618-621, 691-694

Sitio de fosforilación tirosina cinasa.
aminoácidos 632-640

Sitios de N-miristoilación.
aminoácidos 243-248, 398-403, 511-516, 651-656, 753-758, 789-794, 836-841, 844-849

Secuencia de adhesión de células.
aminoácidos 703-705

Dominio de inmunoglobulina.

aminoácidos 36-121, 162-249, 292-375, 422-517, 564-657, 704-795

1015_1630

FIGURA 137

MSRSRHLGKIRKRLEDVKSQWVRPARADFSDNESARLATDALLDGGSEAYWRVLSQEGEVD
FLSSVEAQYIQAQAREPPCPPDTLGGAEAGPKGLDSSSLQSGTYFPVASEGSEPALLHSWA
SAEKPYLKEKSSATVYFQTVKHNNIRDLVRRCITRTSQVLVILMDVFTDVEIFCDILEAAN
KRGVFVCVLLDQGGVKLFQEMCDKVQISDSHLKNISIRSVEGEIYCAKSGRKFAGQIREKF
IISDWRFVLSGSYSFTWLCGHVHRNILSKFTGQAVELFDEEFRHLYASSKFVMGLKSPRLV
APVPPGAAPANGRLSSSSGSASDRTSSNPFSGRSAGSHPGTRSVSASSGPCSPAAPHPPPP
PRFQPHQGPWGAPSPQAHLSPRPHDGPPAAVYSNLGAYRPTRLQLEQLGLVPRLTPTWRPF
LQASPHF

Dominio transmembrana.

aminoácidos 154-175

Sitios de N-glicosilación.

aminoácidos 32-35, 217-220

Sitio de fosforilación tirosina cinasa.

aminoácidos 221-228

Sitios de N-miristoilación.

aminoácidos 45-50, 86-91, 94-99, 112-117, 186-191, 311-316, 341-346

Sitio de amidación.

aminoácidos 232-235

1016 1024

FIGURA 138

MALPQPPPPAGARDPVTPEHWASGPPSQAQPLLRQEAKEEEGEETGVQGAWGTGTAEQRR
RGWGEAAESAAAEEGQAEVGGAAAAGSGSPAGGAGGGLGSWRPLLAWLQRRQPQCCPCAAP
LSRSAAHCCHGGTKMAALAYNLGKREINHYFSVRSAKVLALVAVLLLAACHLASRRYRGND
SCEYLLSSGRFLGEKVWQPHSCMMHKYKISEAKNCLVDKHIAFIGDSRIRQLFYSFVKIIN
PQFKEEGNKHENIPFEDKTASVKVDFLWHPEVNGSMKQCIKVWTEDSIAKPHVIVAGAATW
SIKIHNGSSEALSQYKMNITSIAPLLEKLAKTSDVYWVLQDPVYEDLLSENRKMITNEKID
AYNEAAVSILNSSTRNSKSNVKMFSVSKLIAQETIMESLDGLHLPESSRETTAMILMNVYC
NKILKPVDGSCCQPRPPVTLIQKLAACFFTLSIIGYLIFYIIHRNAHRKNKPCTDLESGEE
KKNIINTPVSSLEILLQSFCKLGLIMAYFYMCDRANLFMKENKFYTHSSFFIPIIYILVLG
VFYNENTKETKVLNREQTDEWKGWMQLVILIYHISGASTFLPVYMHIRVLVAAYLFQTGYG
HFSYFWIKGDFGIYRVCQVLFRLNFLVVVLCIVMDRPYQFYYFVPLVTVWFMVIYVTLALW
PQIIQKKANGNCFWHFGLLLKLGFLLLFICFLAYSQGAFEKIFSLWPLSKCFELKGNVYEW
WFRWRLDRYVVFHGMLFAFIYLALQKRQILSEGKGEPLFSNKISNFLLFISVVSFLTYSIW
ASSCKNKAECNELHPSVSVVQILAFILIRNIPGYARSVYSSFFAWFGKISLELFICQYHIW
LAADTRGILVLIPGNPMLNIIVSTFIFVCVAHEISQITNDLAQIIIPKDNSSLLKRLACIA
AFFCGLLILSSIQDKSKH

Dominio transmembranas.

aminoácidos 153-173, 447-467, 495-515, 532-552, 577-597, 624-644, 651-671, 677-697, 693-713, 733-753, 771-791, 808-828, 866-886, 904-924

Sitios de N-glicosilación.

aminoácidos 182-185, 277-280, 311-314, 323-326, 377-380, 904-907

Sitio de fosforilación tirosina cinasa.

aminoácidos 180-187

Sitios de N-miristoilación.

aminoácidos 47-52, 50-55, 53-58, 81-86, 82-87, 89-94, 93-98, 94-99, 96-101, 97-102, 134-139, 181-186, 312-317, 511-516, 549-554, 746-751

Sitio de amidación.

1017

Aminoácidos 144-147

FIGURA 139

MYHCHSGSKPTEKGANEYAYAKWKLCSASAICFIFMIAEVVGGHIAGSLAVVTDAAHLLID
LTSFLLSLFSLWLSSKPPSKRLTFGWHRAEILGALLSILCIWVVTGVLVYLACERLLYPDY
QIQATVMIIVSSCAVAANIVLTVVLHQRCLGHNHKEVQANASVRAAFVHALGDLFQSISVL
ISALIIYFKPEYKIADPICTFIFSILVLASTITILKDFSILLMEGVPKSLNYSGVKELILA
VDGVLSVHSLHIWSLTMNQVILSAHVATAASRDSQVVRREIAKALSKSFTMHSLTIQMESP
VDQDPDCLFCEDPCD

Dominio transmembranas.

aminoácidos 53-73, 91-111, 127-147, 166-186, 202-222

Sitios de N-glicosilación.

aminoácidos 162-165, 234-237

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasa.

aminoácidos 81-84

Sitio de fosforilación tirosina cinasa.

aminoácidos 13-20

Sitios de N-miristoilación.

aminoácidos 7-12, 42-47, 94-99, 228-233

Familia de eflujo de cationes.

aminoácidos 78-306

FIGURA 140

MVLSVPVIALGATLGTATSILALCGVTCLCRHMHPKKGLLPRDQDPDLEKAKPSLLGSAQQ
FNVKKSTEPVQPRALLKFPDIYGPRPAVTAPEVINYADYSLRSTEEPTAPASPQPPNDSRL
KRQVTEELFILPQNGVVEDVCVMETWNPEKAASWNQAPKLHYCLDYDCQKAELPVTRLEAV
TSNHDGGCDCYVQGSVANRTGSVEAQTALKKRQLHTTWEEGLVLPLAEEELPTATLTLTLR
TCDRFSRHSVAGELRLGLDGTSVPLGAAQWGELKTSAKEPSAGAGEVLLSISYLPAANRLL
VVLIKAKNLHSNQSKELLGKDVSVKVTLKHQARKLKKKQTKRAKHKINPVWNEMIMFELPD
DLLQASSVELEVLGQDDSGQSCALGHCSLGLHTSGSERSHWEEMLKNPRRQIAMWHQLEL

Secuencia de señal.

aminoácidos 1-24

Dominio transmembrana.

aminoácidos 4-24

Sitios de N-glicosilación.

aminoácidos 118-121, 201-204, 317-320

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasas.

aminoácidos 65-68, 342-345

Sitios de N-miristoilación.

aminoácidos 11-16, 15-20, 189-194, 197-202, 204-209, 261-266, 380-385, 385-390, 396-401, 401-406

Dominio C2.

aminoácidos 304-394

FIGURA 141

MQKIMHISVLLSPVLWGLIFGVSSNSIQIGGLFPRGADQEYSAFRVGMVQFSTSEFRLTPH
IDNLEVANSFAVTNAFCSQFSRGVYAIFGFYDKKSVNTITSFCGTLHVSFITPSFPTDGTH
PFVIQMRPDLKGALLSLIEYYQWDKFAYLYDSDRGLSTLQAVLDSAAEKKWQVTAINVGNI
NNDKKDEMYRSLFQDLELKKERRVILDCERDKVNDIVDQVITIGKHVKGYHYIIANLEFTD
GDLLKIQFGGANVSGFQIVDYDDSLVSKFIERWSTLEEKEYPGAHTTTIKYTSALTYDAVQ
VMTEAFRNLRKQRIEISRRGNAGDCLANPAVPWGQGVEIERALKQVQVEGLSGNIKFDQNG
KRINYTINIMELKTNGPRKIGYWSEVDKMVVTLTELPSGNDTSGLENKTVVVTTILESPYV
MMKKNHEMLEGNERYEGYCVDLAAEIAKHCGFKYKLTIVGDGKYGARDADTKIWNGMVGEL
VYGKADIAIAPLTITLVREEVIDFSKPFMSLGISIMIKKPQKSKPGVFSFLDPLAYEIWMC
IVFAYIGVSVVLFLVSRFSPYEWHTEEFEDGRETQSSESTNEFGIFNSLWFSLGAFMRQGC
DISPRSLSGRIVGGVWWFFTLIIISSYTANLAAFLTVERMVSPIESAEDLSKQTEIAYGTL
DSGSTKEFFRRSKIAVFDKMWTYMRSAEPSVFVRTTAEGVARVRKSKGKYAYLLESTMNEY
IEQRKPCDTMKVGGNLDSKGYGIATPKGSSLGTPVNLAVLKLSEQGVLDKLKNKWWYDKGE
CGAKDSGSKEKTSALSLSNVAGVFYILVGGLGLAMLVALIEFCYKSRAEAKRMKVAKNAQN
INPSSSQNSQNFATYKEGYNVYGIESVKI

Secuencia de señal.
aminoácidos 1-21

Dominio transmembranas.
aminoácidos 4-24, 543-563, 624-644, 811-831

Sitios de N-glicosilación.
aminoácidos 256-259, 370-373, 406-409, 413-416

Sitio de fosforilación tirosina cinasas.
aminoácidos 35-41, 662-667

Sitios de N-miristoilación.
aminoácidos 17-22, 21-26, 134-140, 181-186, 254-259, 287-292, 355-360, 405-410, 472-477, 593-598, 669-674, 746-751, 760-765, 764-769, 795-800, 823-828

Sitio de amidación.
aminoácidos 365-368

1021

Ligandos a puertas de canales iónicos.
aminoácidos 543-834

Región de unión de ligandos de la familia de receptores .
aminoácidos 56-398

FIGURA 142

MATTVPDGCRNGLKSKYYRLCDKAEAWGIVLETVATAGVVTSVAFMLTLPILVCKVQDSNR
RKMLPTQFLFLLGVLGIFGLTFAFIIGLDGSTGPTRFFLFGILFSICFSCLLAHAVSLTKL
VRGRKPLSLLVILGLAVGFSLVQDVIAIEYIVLTMNRTNVNVFSELSAPRRNEDFVLLLTY
VLFLMALTFLMSSFTFCGSFTGWKRHGAMIYLTMLLSIAIWVAWITLLMLPDFDRRWDDTI
LSSALAANGWVFLLAYVSPEFWLLTKQRNPMDYPVEDAFCKPQLVKKSYGVENRAYSQEEI
TQGFEETGDTLYAPYSTHFQLQNQPPQKEFSIPRAHAWPSPYKDYEVKKEGS

Dominio transmembranas.

aminoácidos 30-50, 66-86, 98-118, 121-141, 137-157, 174-194, 212-232, 244-264

Sitio de N-glicosilación.

aminoácidos 158-161

Sitios de N-miristoilación.

aminoácidos 8-13, 38-43, 80-85, 88-93, 102-107, 136-141, 201-206

Sitio de amidación.

aminoácidos 124-127

Receptor de transmembrana 7.

aminoácidos 27-273

FIGURA 143

MLPAQEAAKLYHTNYVRNSRAIGVLWAIFTICFAIVNVVCFIQPYWIGDGVDTPQAGYFGL
FHYCIGNGFSRELTCRGSFTDFSTLPSGAFKAASFFIGLSMMLIIACIICFTLFFFCNTAT
VYKICAWMQLTSAACLVLGCMIFFDGWDSDEVKRMCGEKTDKYTLGACSVRWAYILAIIGI
LDALILSFLAFVLGNRQDSLMAEELKAENKVLLSQYSLE

Dominio transmembranas.

aminoácidos 21-41, 93-113, 127-147, 173-193

Sitios de N-miristoilación.

aminoácidos 23-28, 67-72, 89-94, 182-187

FIGURA 144

MPGAAAAAAAAAAAMLPAQEAAKLYHTNYVRNSRAIGVLWAIFTICFAIVNVVCFIQPYWI
GDGVDTPQAGYFGLFHYCIGNGFSRELTCRGSFTDFSTLPSGAFKAASFFIGLSMMLIIAC
IICFTLFFFCNTATVYKICAWMQLTSAACLVLGCMIFPDGWDSDEVKRMCGEKTDKYTLGA
CSVRWAYILAIIGILDALILSFLAFVLGNRQDSLMAEELKAENKVLLSQYSLE

Dominio transmembranas.

aminoácidos 1-15, 34-54, 107-127, 141-161, 187-207

Sitios de N-miristoilación.

aminoácidos 3-8, 37-42, 81-86, 103-108, 196-201

FIGURA 145

MPGAAAAAAAAAAAMLPAQEAAKLYHTNYVRNSRAIGVLWAIFTICFAIVNVVCFIQPYWI
GDGVDTPQAGYFGLFHYCIGNGFSRELTCRGSFTDFSTLPSGAFKAASFFIGLSMMLIIAC
IICFTLFFFCNTATVYKICAWMQLTSAACLVLGCMIFPDGWDSDEVKRMCGEKTDKYTLGA
CSVRWAYILAIIGILDALILSFLAFVLGNRQDSLMAEELKAENKVLLSQYSLE

Dominio transmembranas.

aminoácidos 1-15, 34-54, 107-127, 141-161, 187-207

Sitios de N-miristoilación.

aminoácidos 3-8, 37-42, 81-86, 103-108, 196-201

FIGURA 146

RPLRLPCRGPRAAGTVSSESLGHPAPGQVTNRSPQFPMELFQSSPRPPLQPVCVCPGLELC
DPPHRLAFSMGAGLFSVVGTLLLPGLAALVQDWRLLQGLGALMSGLLLLFWGFPALFPESP
CWLLATGQVARARKILWRFAEASGVGPGDSSLEENSLATELTMLSARSPQPRYHSPLGLLR
TRVTWRNGLILGFSSLVGGGIRASFRRSLAPQVPTFYLPYFLEAGLEAAALVFLLLTADCC
GRRPVLLLGTMVTGLASLLLLAGAQYLPGWTVLFLSVLGLLASRAVSALSSLFAAEVFPTV
IRGAGLGLVLGAGFLGQAAGPLDTLHGRQGFFLQQVVFASLAVLALLCVLLLPESRSRGLP
QSLQDADRLRRSPLLRGRPRQDHLPLLPPSNSYWAGHTPEQH

Dominio transmembranas.

aminoácidos 67-87, 102-122, 220-240, 249-269, 265-285, 301-321, 338-358

Sitio de fijación de glicosaminoglicano.

aminoácidos 145-148

Sitios de N-miristoilación.

aminoácidos 14-19, 27-32, 57-62, 101-106, 146-151, 180-185, 191-196, 202-207, 203-208, 228-233, 253-258, 283-288, 308-313, 312-317, 364-369

Sitio de amidación.

aminoácidos 244-247

Patrón de cremallera de leucina.

aminoácidos 96-117

Proteínas de transporte de azúcar.

aminoácidos 117-126, 241-253, 276-321

FIGURA 147

MPFEQCHLSWKNVRRLYNAVSKELVGEFLQFVQLDKEASDPFSLNELLDELSRKQKEELWQ
RLKNLLTDVLLESPVDGWQVVEAQGEDNMETEHGSKMRKSIEIIYAITSVILASVSVINES
ENYEALLECVIILNGILYALPESERKLQSSIQDLCVTWWERGLPAKEDTGKTAFVMLLRRS
LETKTGADVCRLWRIHQALYCFDYDLEESGEIKDMLLECFININYIKKEEGRRFLSCLFNW
NINFIKMIHGTIKNQLQGLQKSLMVYIAEIYFRAWKKASGKILEAIENDCIQDFMFHGIHL
PRRSPVHSKVREVLSYFHHQKKVRQGVEEMLYRLYKPILWRGLKARNSEVRSNAALLFVEA
FPIRDPNLHAIEMDSXIQKQFEELYSLLEDFYPMVRSTGILGVCKITSKYWEMMPPTILID
LLKKVTGELAFDTSSADVRCSVFKCLPMILDNKLSHPLLEQLLPALRYSLHDNSEKVRVAF
VDMLLKIKAVRAAKFWKICPMEHILVRLETDSRPVSRRLVSLIFNSFLPVNQPEEVWCERC
VTLVQMNHAAARRFYQYAHEHTACTNIAKLIHVIRHCLNACIQRAVREPPEDEEEEDGREK
ENVTVLDKTLSVNDVACMAGLLEIIVILWKSIDRSMENNKEAKLYTINKFASVLPEYLKVF
KDDRCKIPLFMLMSFMPASAVPPFSCGVISTLRSREEGAVDKSYCTLLDCLCSWGQVGHIL
RLVDNWLPTEHAQAKSNTASKGRVQIHDTRPVKPELALVYIEYLLTHPKNRECLLSAPRKK
LNHLLKALETSKADLESLLQTPGGKPRGFSEAAAPRAFGLECRLSIHLQHKFCSEGKVYLS
MLEDTGFWLESKILSFIQDQEEDYLKLHRVIYQQIIQTYLTVCKDVVMVGLGDHQFQMQLL
QRSLGIMQTVRGFFYVSLLLDILKEITGSSLIQKTDSDEEVAMLLDTVQKVFQKMLECIAR
SFRKQPEEGLRLLYSVQRPLHEFITAVQSRHTDTPVERGVLSTLIAGPVVEISHQLRKVSD
VEELTPPEHLSDLPPFSRCLIGIIIKSSNVVRSFLDELKACVASNDIEGIVCLTAAVEIIL
VINAGKHKSSKVREVAATVHRKLKTFMEITLEEDSIERFLYESSSRTLGELLNS

Dominio transmembrana.

aminoácidos 1081-1101

Sitios de N-glicosilación.

aminoácidos 120-123, 612-615

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasas.

aminoácidos 280-283, 430-433, 1033-1036

Sitio de fosforilación tirosina cinasas.

aminoácidos 99-106, 314-321, 329-337, 385-391, 844-852

Sitios de N-miristoilación.

aminoácidos 137-142, 189-194, 254-259, 262-267, 698-703, 920-925, 1015-1020

Sitio de amidación.

aminoácidos 233-236

Sitio de unión ATP/GTP motif A (Vuelta- P).

aminoácidos 167-174

Patrón de cremallera de leucina.

aminoácidos 128-149, 135-156, 787-808

Señal de transferencia de proteínas mitocondriales de energía.

aminoácidos 982-990

1029 TST

FIGURA 148

MEKRETFVQAVSKELVGEFLQFVQLDKEASDPFSLNELLDELSRKQKEELWQRLKNLLTDV
LLESPVDGWQVVEAQGEDNMETEHGSKMRKSIEIIYAITSVILASVSVINESENYEALLEC
VIILNGILYALPESERKLQSSIQDLCVTWWEKGLPAKEDTGKTAFVMLLRRSLETKTGADV
CRLWRIHQALYCFDYDLEESGEIKDMLLECFININYIKKEEGRRFLSCLFNWNINFIKMIH
GTIKNQLQGLQKSLMVYIAEIYFRAWKKASGKILEAIENDCIQDFMFHGIHLPRRSPVHSK
VREVLSYFHHQKKVRQGVEEMLYRLYKPILWRGLKARNSEVRSNAALLFVEAFPIRDPNLH
AIEMDSEIQKQFEELYSLLEDPYPMVRSTGILGVCKITSKYWEMMPPTILIDLLKKVTGEL
AFDTSSADVRCSVFKCLPMILDNKLSHPLLEQLLPALRYSLHDNSEKVRVAFVDMLLKIKA
VRAAKFWKICPMEHILVRLETDSRPVSRRLVSLIFNSFLPVNQPEEVWCERCVTLVQMNHA
AARRFYQYAHEHTACTNIAKLIHVIRHCLNACIQRAVREPPEDEEEEDGREKENVTVLDKT
LSVNDVACMAGLLEIIVILWKSIDRSMENNKEAKLYTINKFASVLPEYLKVFKDDRCKIPL
FMLMSFMPASAVPPFSCGVISTLRSREEGAVDKSYCTLLDCLCSWGQVGHILELVDNWLPT
EHAQAKSNIASKGRVQIHDTRPVKPELALVYIEYLLTHPKNRECLLSAPRKKLNHLLKALE
TSKADLESLLQTPGGKPRGFSEAAAPRAFGLHCRLSIHLQHKFCSEGKVYLSMLEDTGFWL
ESKILSFIQDQEEDYLKLHRVIYQQIIQTYLTVCKDVVMVGLGDHQFQMQLLQRSLGIMQT
VKGFFYVSLLLDILKEITGSSLIQKTDSDEEVAMLLDTVQKVFQKMLECIARSFRKQPEEG
LRLLYSVQRPLHEFITAVQSRHTDTPVHRGVLSTLIAGPVVEISHQLRKVSDVEELTPPEH
LSDLPPFSRCLIGIIIKSSNVVRSFLDELKACVASNDIEGIVCLTAAVHIILVINAGKHKS
SKVREVAATVHRKLKTFMEITLERDSIERFLYESSSSRTLGELLNS

Dominio transmembrana.

aminoácidos 1072-1092

Sitios de N-glicosilación.

aminoácidos 111-114, 603-606

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasas.

aminoácidos 3-6, 271-274, 421-424, 1024-1027

Sitio de fosforilación tirosina cinasa.

aminoácidos 91-97, 305-312, 320-328, 376-382, 835-843

Sitios de N-miristoilación.

aminoácidos 128-133, 180-185, 245-250, 253-258, 689-694, 911-916, 1006-1011

Sitio de amidación.

aminoácidos 224-227

1030

Sitio de unión ATP/GTP motif A (Vuelta- P).

aminoácidos 158-165

Patrón de cremallera de leucina.

aminoácidos 119-140, 126-147, 778-799

Señal de transferencia de proteínas mitocondriales de energía.

aminoácidos 973-981

FIGURA 149

MNSTLFSQVENHSVHSNFSEKNAQLLAFENDDCHLPLAMIFTLALAYGAVIILGVSGNLAL
IIIILKQKEMRNVTNILIVNLSFSDLLVAIMCLPFTFVYTLMDHWVFGEAMCKLNPFVQCV
SITVSIFSLVLIAVERHQLIINPRGWRPNNRHAYVGIAVIWVLAVASSLPFLIYQVMTDEP
FQNVTLDAYKDKYVCFDQFPSDSHRLSYTTLLLVLQYFGPLCFIFICYFKIYIRLKRRNNM
MDKMRDNKYRSSETKRINIMLLSIVVAFAVCWLPLTIFNTVFDWNHQIIATCNHNLLFLLC
HLTAMISTCVNPIFYGFLNKNFQRDLQFFFNFCDFRSRDDDYETIAMSTMHTDVSKTSLKQ
ASPVAFKKINNNDDNEKI

Dominio transmembranas.
aminoácidos 38-58, 78-98, 117-137, 155-175, 208-228, 260-280, 296-316

Sitios de N-glicosilación.
aminoácidos 2-5, 11-14, 17-20, 73-76, 81-84, 186-189

Sitio de fosforilación tirosina cinasas.
aminoácidos 247-253, 341-347

Sitio de N-miristoilación.
aminoácidos 54-59

Receptor de transmembrana 7 (familia de rodopsinas).
aminoácidos 57-320

1032

FIGURA 150

MVRCDRGLQMLLTTAGAFAAFSLMAIAIGTDYWLYSSAHICNGTNLTMDDGPPPRRARGDL
THSGLWRVCCIEGIYKGHCFRINHFPEDNDYDHDSSEYLLRIVRASSVFPILSTILLLLGG
LCIGAGRIYSRKNNIVLSAGILFVAAGLSNIIGIIVYISSNTGDPSDKRDEDKKNHYNYGW
SFYFGALSFIVAETVGVLAVNIYIEKNKELRFKTKREFLKASSSSPYARMPSYRYRRRRSR
SSSRSTEASPSRDVSPMGLKITGAIPMGELSMYTLSREPLKVTTAASYSPDQEASFLQVHD
FFQQDLKEGFHVSMLNRRTTPV

Secuencia de señal.
aminoácidos 1-19

Dominio transmembranas.
aminoácidos 103-123, 140-160, 182-202

Sitios de N-glicosilación.
aminoácidos 42-45, 45-48

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasas.
aminoácidos 240-243, 322-325

Sitios de N-miristoilación.
aminoácidos 16-21, 43-48, 74-79, 122-127, 262-267

Secuencia de adhesión de células.
aminoácidos 58-60

PMP-22/EMP/MP20/familia Claudin.
aminoácidos 6-202

FIGURA 151

MARGPGLAPPPLRLPLLLLVLAAVTGHTAAQDNCTCPTNKMTVCSPDGPGGRCQCRALGSG
MAVDCSTLTSKCLLLKARMSAPKNARTLVRPSEHALVDNDGLYDPDCDPEGRFKARQCNQT
SVCWCVNSVGVRRTDKGDLSLRCDELVRTHHILIDLRHRPTAGAFNHSDLDAELRRLFRER
YRLHPKFVAAVHYEQPTIQIELRQNTSQKAAGEVDIGDAAYYFERDIKGESLFQGRGGLDL
RVRGEPLQVERTLIYYLDEIPPKFSMKRLTAGLIAVIVVVVVALVAGMAVLVITNRRKSGK
YKKVEIKELGELRKEPSL

Secuencia de señal.
aminoácidos 1-26

Dominio transmembrana.
aminoácidos 274-294

Sitios de N-glicosilación.
aminoácidos 33-36, 120-123, 168-171, 208-211

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasas.
aminoácidos 271-274, 300-303

Sitios de N-miristoilación.
aminoácidos 59-64, 132-137

Repetición de tiroglobulina de tipo 1.
aminoácidos 73-145

1034

FIGURA 152

MGKCSGRCTLVAFCCLQLVAALERQIFDFLGYQWAPILANFLHIMAVILGIFGTVQYRSRY
LILYAAWLVLWVGWNAFIICFYLEVGQLSQDRDFIMTFNTSLHRSWWMENGPGCLVTPVLN
SRLALEDHHVISVTGCLLDYPYIEALSSALQIFLALFGFVFACYVSKVFLEEEDSFDFIGG
FDSYGYQAPQKTSHLQLQPLYTSG

Secuencia de señal.

aminoácidos 1-21

Dominio transmembranas.

aminoácidos 33-53, 62-82, 147-167

Sitio de N-glicosilación.

aminoácidos 100-103

Sitio de fosforilación tirosina cinasa.

aminoácidos 24-32

Sitios de N-miristoilación.

aminoácidos 50-55, 182-187

FIGURA 153

MPAPRAREQPRVPGERQPLLPRGARGPRRWRRAAGAAVLLVEMLERAAFFGVTANLVLYLN
STNFNWTGEQATRAALVFLGASYLLAPVGGWLADVYLGRYRAVALSLLLYLAASGLLPATA
FPDGRSSFCGEMPASPLGPACPSAGCPRSSPSPYCAPVLYAGLLLLGLAASSVRSNLTSFG
ADQVMDLGRDATRRFFNWFYWSINLGAVLSLLVVAFIQQNISFLLGYSIPVGCVGLAFFIF
LFATPVFITKPPMGSQVSSMLKLALQNCCPQLWQRHSARDRQCARVLADERSPQPGASPQE
DIANFQVLVKILPVMVTLVPYWMVYFQMQSTYVLQGLHLHIPNIFPANPANISVALRAQGS
SYTIPEAWLLLANVVVVLILVPLKDRLIDPLLLRCKLLPSALQKMALGMFFGFTSVIVAGV
LEMERLHYIHHNETVSQQIGEVLYNAAPLSIWWQIPQYLLIGISEIFASIPGLEFAYSEAP
RSMQGAIMGIFFCLSGVGSLLGSSLVALLSLPGGWLHCPKDFGNINNCRMDLYFFLLAGIQ
AVTALLFVWIAGRYERASQGPASHSRFSRDRG

Dominio transmembranas.

aminoácidos 35-55, 73-93, 98-118, 157-177, 199-219, 250-270, 341-361, 369-389, 408-428, 457-477, 499-519, 537-557

Sitios de N-glicosilación.

aminoácidos 61-64, 66-69, 178-181, 223-226, 356-359, 439-442

Sitio de fijación de glicosaminoglicano.

aminoácidos 503-506

Sitios de N-miristoilación.

aminoácidos 51-56, 90-95, 116-121, 147-152, 169-174, 209-214, 258-263, 365-370, 414-419, 479-484, 493-498, 497-502, 506-511, 531-536

Familia POT.

aminoácidos 101-503

FIGURA 154

MWVNPEEVLLANALWITERANPYFILQRRKGHAGDGGGGGGLAGLLVGTLDVVLDSSARVA
PYRILYQTPDSLVYWTIACGGSRKEITEHWEWLEQNLLQTLSIFENENDITTFVRGKIQGI
IAEYNKINDVKEDDDTEKFKEAIVKFHRLFGMPEEEKLVNYYSCSYWKGKVPRQGWMYLSI
NHLCFYSFLMGREAKLVIRWVDITQLEKNATLLLPDVIKVSTRSSEHFFSVFLNINETFKL
MEQLANIAMRQLLDNEGFEQDRSLPKLKRKSPKKVSALKRDLDARAKSERYRALFRLPKDE
KLDGHTDCTLWTPFNKMHILGQMFVSTNYICFTSKEENLCSLIIPLREVTIVEKADSSSVL
PSPLSISTRNRMTFLFANLKDRDFLVQRISDFLQQTTSKIYSDKEFAGSYNSSDDEVYSRP
SSLVSSSPQRSTSSDADGERQFNLNGNSVPTATQTLMTMYRRRSPEEFNPKLAKEFLKEQA
WKIHFAEYGQGICMYRTEKTRELVLKGIPESMRGELWLLLSGAINEKATHPGYYEDLVEKS
MGKYNLATEEIERDLHRSLPEHPAFQNEMGIAALRRVLTAYAFRNPNIGYCQAMNIVTSVL
LLYAKEEEAFWLLVALCERMLPDYYNTRVVGALVDQGVFEELARDYVPQLYDCMQDLGVIS
TISLSWFLTLFLSVMPFESAVVVVDCFFYEGIKVIFQLALAVLDANVDKLLNCKDDGEAMT
VLGRYLDSVTNKDSTLPPIPELHSLLSDDVEPYPEVDIFRLIRTSYEKFGTIRADLIEQMR
FKQRLKVIQTLEDTTKRNVVRTIVTETSFTIDELEELYALFKAEHLTSCYWGGSSNALDRH
DPSLPYLEQYRIDFEQFKGMFALLFPWACGTHSDVLASRLFQLLDENGDSLINFREFVSGL
SAACHGDLTEKLKLLYKMHVLPEPSSDQDEPDSAFEATQYFFEDITPECTHVVGLDSRSKQ
GADDGFVTVSLKPDKGKRANSQENRNYLRLWTPENKSKSKNAKDLPKLNQGQFIELCKTMY
NMFSEDPNEQELYHATAAVTSLLLEIGEVGKLFVAQPAKEGGSGGSGPSCHQGIPGVLFPK
KGPGQPYVVESVEPLPASLAFDSEEESLGGQMEDIKLEDSSPRDNGACSSMLISDDDTKDD
SSMSSYSVLSAGSHEEDKLHCEDIGEDTVLVRSGQGTAALPRSTSLDRDWAITFEQFLASL
LTEPALVKYFDKPVCMMARITSAKNIRMMGKPLTSASDYEISAMSG

Dominio transmembranas.

aminoácidos 35-55, 669-689

Sitios de N-glicosilación.

aminoácidos 212-215, 239-242, 417-420, 1011-1014

Sitio de fijación de glicosaminoglicano.

aminoácidos 1192-1195

Sitio de fosforilación cAMP- y cGMP- dependiente de proteína cinasas.

aminoácidos 272-275, 277-280, 468-471

Sitios de N-miristoilación.

aminoácidos 36-41, 37-42, 39-44, 40-45, 44-49, 414-419, 497-502, 515-520, 668-673, 782-787, 845-850, 846-851, 914-919, 977-982, 1078-1083, 1079-1084, 1082-1087, 1144-1149,

1193-1198

Sitio de amidación.

aminoácidos 991-994

Dominio TBC.

aminoácidos 512-724

Calcio de tipo S-100/ ICaBP.

aminoácidos 891-911

Fig. 155

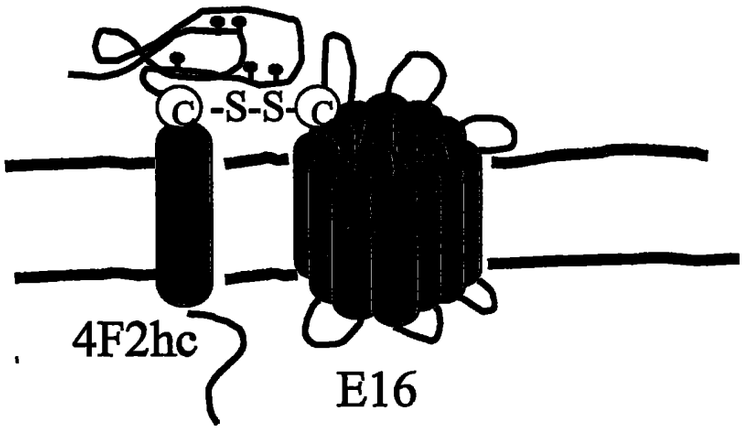
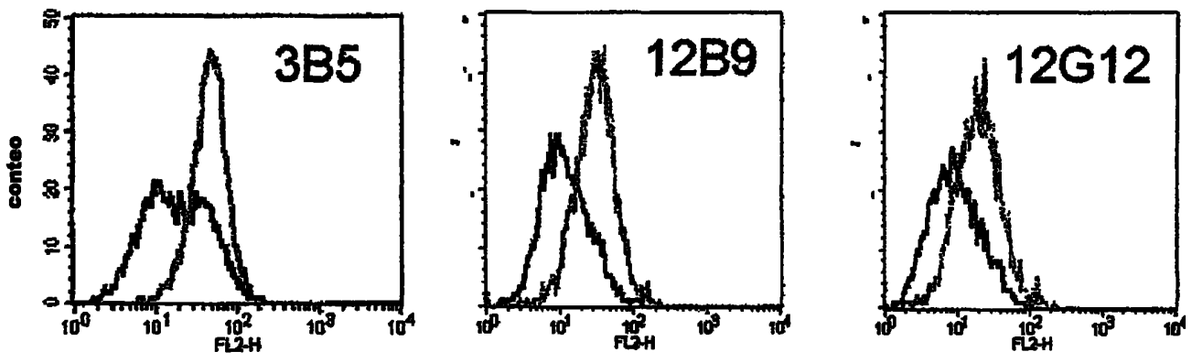


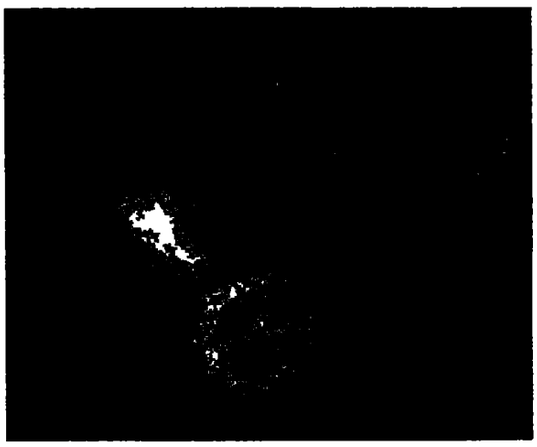
FIG 156



Verde: ARNi de control

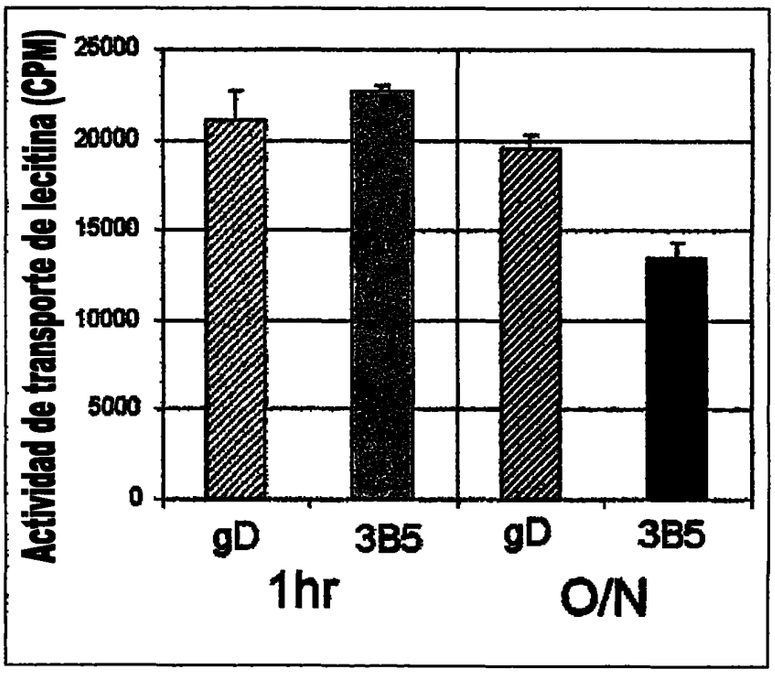
R jo: ARNi E16

FIG 157

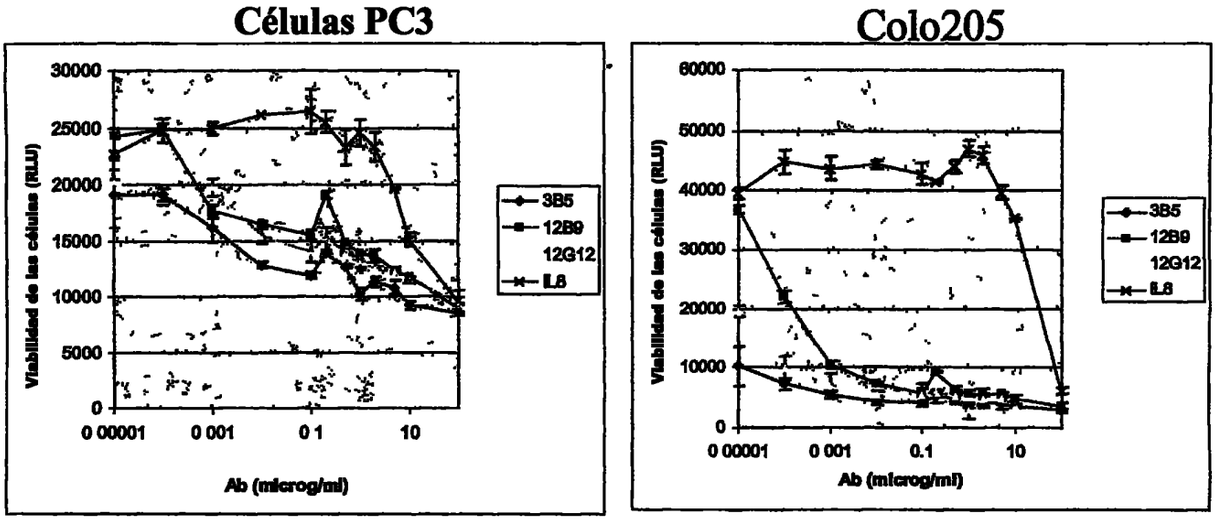


Células PC3 tratadas con E16 mAb (3B5) durante la noche

FIG 158


Actividad de transporte de lecitina (CPM)
25000
20000
15000
10000
5000
0
gD
3B5
gD
3B5
1hr
O/N

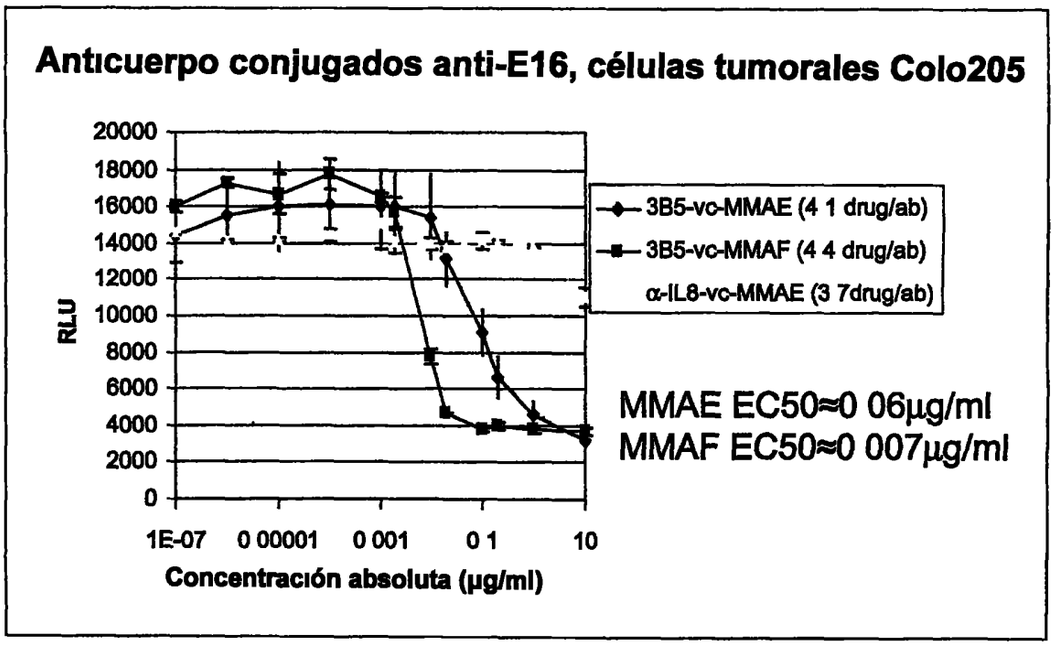
Ef cto de anti-E-16-vc-MMAE sobre el crecimiento de células cancerígenas in vitro



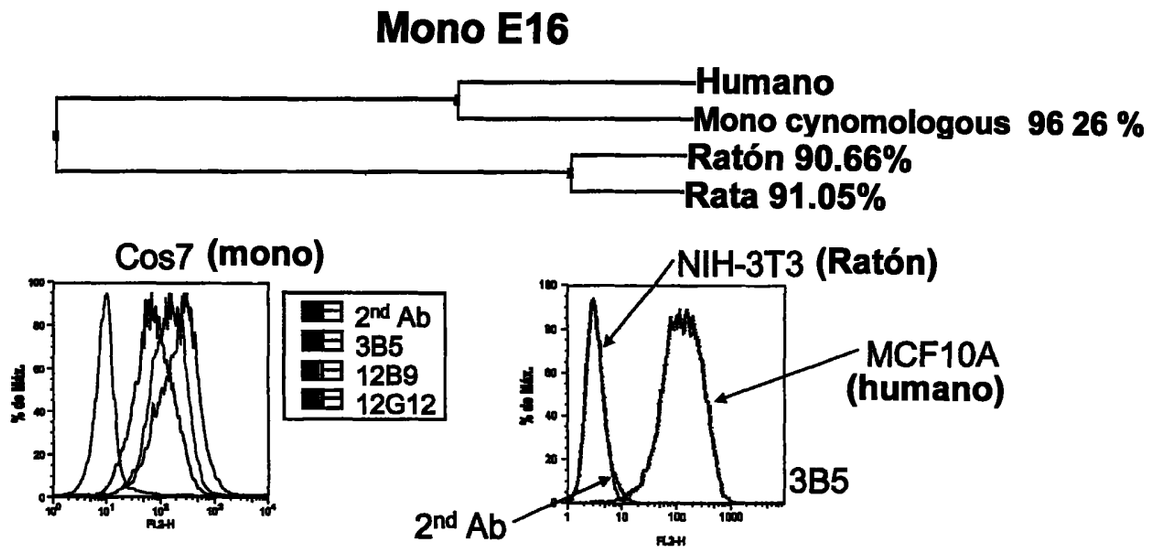
Tratamiento durante 72 h

Fig. 159

FIG. 160



WO 2006/029183
243 / 249
PCT/US2005/031798



El anticuerpo anti-E16 no reacciona con ratón E16, pero si tiene una reacción cruzada con mono E16

FIG 161

Anti u rp s antı-E16 sın conjugar, Xen ınjerto

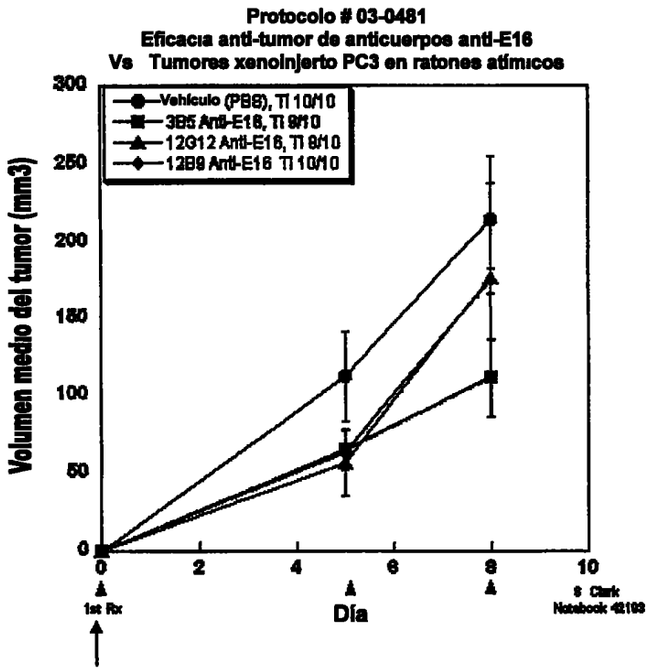


FIG. 162

Día 0 ıny ctar células PC3 & Ab PC3

FIG 163

Anticuerpos anti-E16 conjugados, Xenoinjerto Colo205

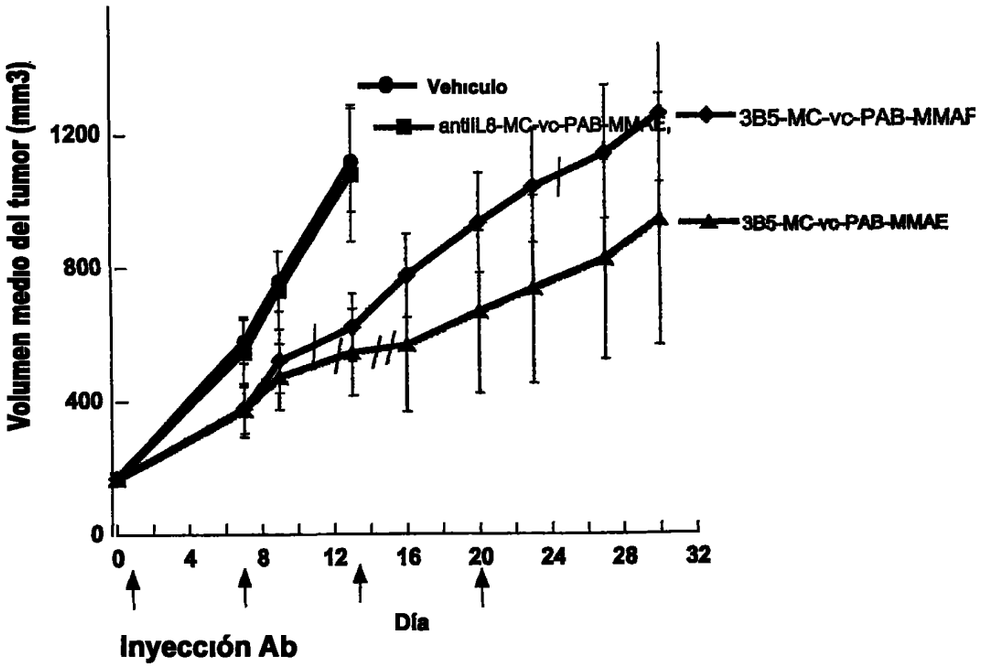


FIG 164

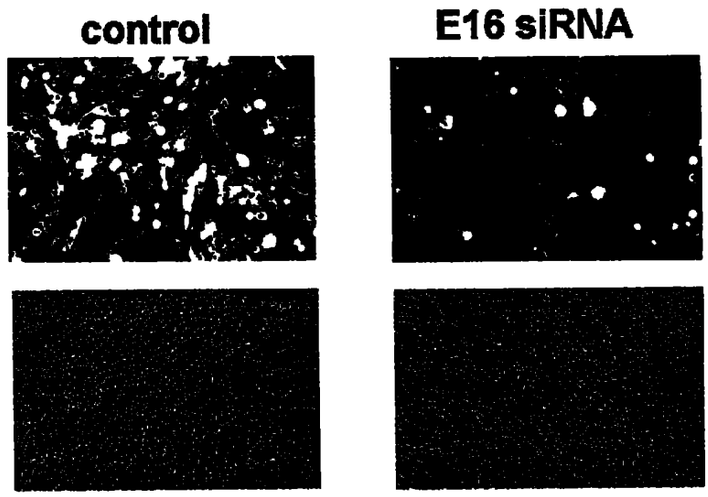


FIG 165

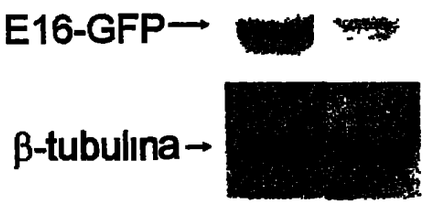
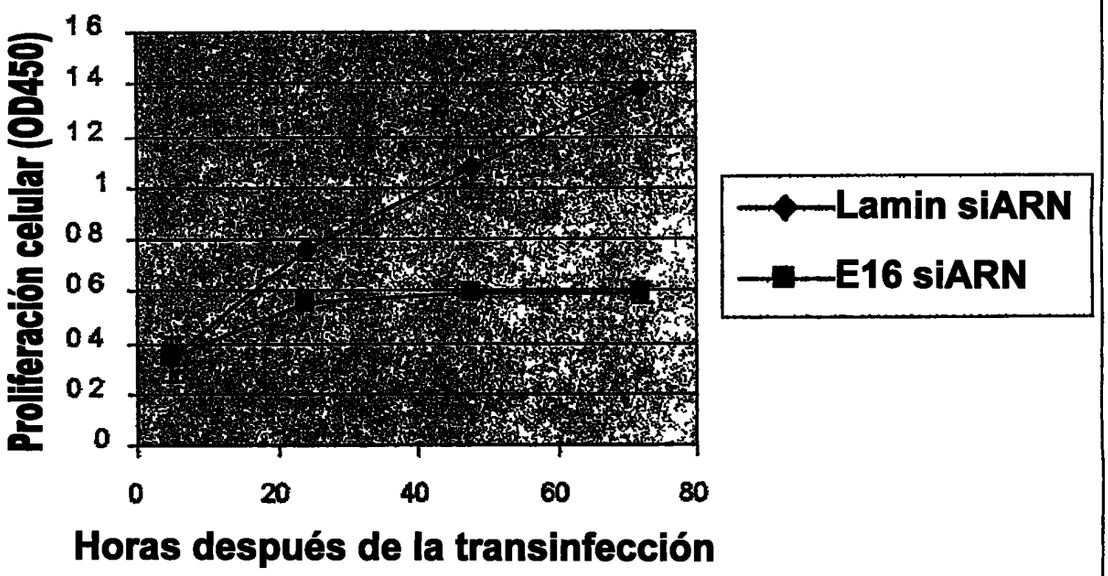


FIG 166

Actividad de transporte de leucina
0.6
0.5
0.4
0.3
0.2
0.1
0
Control
E16 RNAi

FIG 167


Proliferación celular (OD450)
16
14
12
1
08
06
04
02
0
0
20
40
60
80
Horas después de la transinfección
Lamin siARN
E16 siARN

1053
1050

AMENDED

CLAIMS

Amendments to the Claims:

Please amend the claims as follows, wherein strikeout ~~00~~ indicates deleted terminology and underlining 00 indicates added terminology.

WHAT IS CLAIMED IS:

1. (Currently Amended) An isolated antibody that binds to a polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in ~~any one of Figures 79 to 154 (SEQ ID NOS:79-154)~~115 (SEQ ID NO. 115);

(b) the polypeptide shown in ~~any one of Figures 79 to 154~~ (SEQ ID NOS:~~79-154~~, 115), lacking its associated signal peptide;

(c) an extracellular domain of the polypeptide shown in ~~any one of Figures 79 to 154~~ 115 (SEQ ID NOS:~~79-154~~115), with its associated signal peptide;

(d) an extracellular domain of the polypeptide shown in ~~any one of Figures 79 to 154~~ 115(SEQ ID NOS:~~79-154~~,115), lacking its associated signal peptide;

(e) a polypeptide encoded by the nucleotide sequence shown in ~~any one of Figures 1 to 78A-B~~37 (SEQ ID NOS:~~1-78~~;37); or

(f) a polypeptide encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures ~~1 to 78A-B~~37 (SEQ ID NOS:~~1-78~~;37).

2. (Currently amended) An isolated antibody that binds to a polypeptide having:

(a) the amino acid sequence shown in any one of Figures ~~79 to 154~~115 (SEQ ID NOS:~~79-154~~;115);

(b) the amino acid sequence shown in any one of Figures ~~79 to 154~~115 (SEQ ID NOS:~~79 to 154~~;115), lacking its associated signal peptide sequence;

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), with its associated signal peptide sequence;

(d) an amino acid sequence of an extracellular domain of the polypeptide shown in any one of Figures 79 to 154 (SEQ ID NOS:79-154), lacking its associated signal peptide sequence;

(e) an amino acid sequence encoded by the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78); or

(f) an amino acid sequence encoded by the full-length coding region of the nucleotide sequence shown in any one of Figures 1 to 78A-B (SEQ ID NOS:1-78).

3. (Original) The antibody of Claim 1 which is a monoclonal antibody.

4. (Original) The antibody of Claim 1 which is an antibody fragment.

5. (Original) The antibody of Claim 1 which is a chimeric or a humanized antibody.

6. (Original) The antibody of Claim 1 which is conjugated to a growth inhibitory agent.

7. (Original) The antibody of Claim 1 which is conjugated to a cytotoxic agent.

8. (Original) The antibody of Claim 7, wherein the cytotoxic agent is selected from the group consisting of toxins, antibiotics, radioactive isotopes and nucleolytic enzymes.

9. (Original) The antibody of Claim 7, wherein the cytotoxic agent is a toxin.

10. (Original) The antibody of Claim 9, wherein the toxin is selected from the group consisting of auristatin, maytansinoid and calicheamicin.

11. (Original) The antibody of Claim 9, wherein the toxin is a maytansinoid.

12. (Original) The antibody of Claim 1 which is produced in bacteria.

13. (Original) The antibody of Claim 1 which is produced in CHO cells.

14. (Original) The antibody of Claim 1 which induces death of a cell to which it binds.

15. (Original) The antibody of Claim 1 which is detectably labeled.

16. (Original) The antibody of Claim 9, wherein the toxin is selected from the group consisting of MMAE (monomethyl auristatin E), MMAF, MMAF-DMAEA, (MMAF-dimethylaminoethylamine), MMAF-TEG (MMAF-tetraethylene glycol), MMAF-NtBu, and AEVB (auristatin E valeryl benzylhydrazone), and AFP (Apristatin F phenylene diamine).

17. (Original) The antibody of claim 9, wherein the toxin is covalently attached to the antibody by a linker.

18. (Original) The antibody of claim 17, wherein the linker is selected from the group consisting of maleimidocaproyl (MC), valine-citrulline (val-cit, vc), citrulline (2-amino-5-ureido pentanoic acid), PAB (p-

aminobenzylcarbamoyl), Me (N-methyl-valine citrulline), MC(PEG)6-OH (maleimidocaproyl- polyethylene glycol), SPP (N-Succinimidyl 4-(2-pyridylthio) pentanoate), SMCC (N-Succinimidyl 4-(N-maleimidomethyl) cyclohexane-1 carboxylate), and MC-vc-PAB.

19. (Original) The antibody of claim 16, wherein the toxin is MMAE.

20. (Original) The antibody of claim 16, wherein the toxin is MMAF.

21. (Original) The antibody of claim 19 or 20, wherein the linker is MC-vc-PAB.

22. (Original) The antibody of any one of claims 1-21, wherein the antibody binds the TAT188 polypeptide.

23. (Original) The antibody of claim 22, wherein the antibody inhibits proliferation or promotes cell death of a cell expressing TAT188.

24. (Original) The antibody of claim 23, wherein the cell is a cancer cell.

25. (Original) The antibody of claim 24, where the cancer cell is selected from the group of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

26. (Original) An isolated antibody that competes with binding to the epitopes of TAT188 polypeptide bound by an antibody produced by the hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. ____6193), 12B9.1 (ATCC Accession No. ____6194) and 12G12.1 (ATCC Accession No.____6195).

27. (Currently amended) An isolated antibody having the biological activity of an antibody produced by the hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. ____6193), 12B9.1 (ATCC Accession No. ____6194) and 12G12.1 (ATCC Accession No.____6195), wherein the biological activity is inhibition of cell proliferation or promotion of cell death in a cell expressing TAT188.

28. (Currently amended) An isolated antibody comprising in a corresponding complementary determining region (CDR) an amino acid sequence having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid sequence of at least 1, 2, 3, 4, 5, or 6 of the CDR(s) of the antibody produced by a hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. ____6193), 12B9.1 (ATCC Accession No. ____6194) and 12G12.1 (ATCC Accession No. ____6195).

29. (Currently amended) The antibody of any one of claims 1-21, wherein the antibody comprises in a corresponding complementary determining region (CDR) an amino acid sequence having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid sequence of at least 1, 2, 3, 4, 5, or 6 of the CDR(s) of the antibody produced by a hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. ~~6193~~), 12B9.1 (ATCC Accession No. ~~6194~~) and 12G12.1 (ATCC Accession No. ~~6195~~).

30. (Currently amended) The antibody of any one of claims 1-21, wherein the antibody exhibits the biological activity of an antibody produced by the hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. ~~6193~~), 12B9.1 (ATCC Accession No. ~~6194~~) and 12G12.1 (ATCC Accession No. ~~6195~~), wherein the biological activity is inhibition of cell proliferation or promotion of cell death in a cell expressing TAT188.

31. (Original) The antibody of claim 30, wherein the cell is a cancer cell.

32. (Original) The antibody of claim 31, wherein the cancer cell is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

33. (Original) A method of inhibiting growth of a cell expressing TAT188, the method comprising contacting the cell with an antibody of any one of claims 1-21.

34. (Original) The method of claim 33, wherein the cell is a cancer cell.

35. (Original) The method of claim 34, wherein the cancer cell is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

36. (Original) The method of claim 35, wherein the cancer cell is a mammalian cell.

37. (Original) The method of claim 36, wherein the mammalian cell is a human cell.

38. (Original) A method of inhibiting growth of a cell expressing TAT188, the method comprising contacting the cell with an antibody of any one of claims 1-21, wherein the antibody comprises in a corresponding complementary determining region (CDR) an amino acid sequence having at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid sequence of at least 1, 2, 3, 4, 5, or 6 of the CDR(s) of the

antibody produced by a hybridoma selected from the group consisting of 3B5.1 (ATCC Accession No. _____), 12B9.1 (ATCC Accession No. ____) and 12G12.1 (ATCC Accession No. ____).

39. (Original) A method of detecting the level of TAT188 polypeptide expressed in a test cell relative to a control cell, the method comprising:

(a) contacting the test cell and the control cell with an isolated anti-TAT188 antibody of claim 26;

(b) detecting binding of the antibody; and

(c) determining the relative binding of the antibody to the test and control cell.

40. (Original) The method of claim 39, wherein the test cell and control cell are lysed .

41. (Original) The method of claim 39, wherein the test cell is in a tissue.

42. (Original) The method of claim 41, wherein the tissue is a tumor tissue.

43. (Original) The method of claim 42, wherein the tumor tissue is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

44. (Original) A method of detecting the level of TAT188 polypeptide or a polypeptide having at least 80% sequence identity to the amino acid sequence shown in Figure 115 (SEQ ID NO:115) in a test cell relative to a control cell, the method comprising:

(a) contacting the test cell and the control cell with an isolated antibody of any one of claims 1-5, 12-15, and 26-28;

(b) detecting binding of the antibody; and

(c) determining the relative binding of the antibody to the test and control cell.

45. (Original) The method of claim 44, wherein the level of TAT188 polypeptide in the test cell is greater than that in the control cell.

46. (Original) The method of claim 45, wherein the method diagnoses cancer in a tissue containing or having contained the test cell.

47. (Original) A TAT binding interfering RNA (siRNA) which binds to a nucleic acid having at least 80% sequence identity to:

(a) a nucleotide sequence shown in Figure 37 (SEQ ID NO:37; and

(b) the complement of (a), wherein the siRNA reduces expression of TAT188.

48. (Original) An expression vector comprising the siRNA of Claim 47.

49. (Original) The expression vector of Claim 48, wherein said siRNA is operably linked to control sequences recognized by a host cell transfected with the vector.

50. (Original) A host cell comprising the expression vector of Claim 49.

51. (Currently amended) A composition of matter comprising:

(a) the antibody of Claim 1, or

~~(b) the siRNA of claim 47,~~ in combination with a carrier.

52. (Original) The composition of matter of Claim 51, wherein said carrier is a pharmaceutically acceptable carrier.

53. (Original) An article of manufacture:

(a) a container; and

(b) the composition of matter of Claim 51 contained within said container.

54. (Original) The article of manufacture of Claim 53, further comprising a label affixed to said container, or a package insert included with said container, referring to the use of said composition of matter for the therapeutic treatment of or the diagnostic detection of a cancer.

55. (Original) A method of inhibiting the growth of a cancer cell that expresses a polypeptide having at least 80% amino acid sequence identity to the amino acid sequence shown in Figure 115 (SEQ ID NO:115), said method comprising contacting said cancer cell with a siRNA that binds to a nucleic acid encoding the amino acid in said cancer cell, thereby inhibiting the growth of said cancer cell.

56. (Original) The method of claim 55, wherein the nucleic acid has the sequence shown in Figure 37 (SEQ ID NO:37).

57. (Original) The method of Claim 44, wherein the detecting the level of expression of the polypeptide comprises employing an antibody in an immunohistochemistry analysis.

58. (Original) A method for treating or preventing a cell proliferative disorder associated with increased expression or activity of a polypeptide having at least 80% amino acid sequence identity to the amino acid sequence shown in Figure 115 (SEQ ID NO:115), said method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a TAT188 polypeptide.

59. (Original) The method of claim 58, wherein said antagonist is an isolated anti-TAT188 polypeptide antibody of any one of claims 1-21 and 23-28.

60. (Original) The method of claim 59, wherein the cell proliferative disorder is cancer.

61. (Original) The method of claim 60, wherein the cancer is selected from the group consisting of breast, colon, rectum, endometrium, kidney, lung, ovary, skin, and liver.

REMARKS

Original claims 3-26, 31-50, 52-61 are pending in the application. Claims 1-2, 27-30, and 51 are currently amended. No new claims are being added by the amendments. Support for the amendments is found throughout the specification, such as at page 33, lines 4-10, page 34 lines 18-21, and page 36, lines 33-34 and the ATCC deposit receipt enclosed herewith. No new matter is added by the amendments.

Applicants respectfully submit that the claims are in condition for allowance. It is therefore earnestly solicited that such a final favorable disposition is made. The Examiner is invited to telephone Deirdre L. Conley, Ph.D., (Reg. No. 36,487) at (650) 225-2066 if deemed helpful to clarify and advance prosecution.

This Preliminary Amendment is submitted with a transmittal letter. In the unlikely event that this document is separated from the transmittal letter or if fees are required, applicants petition the Commissioner to authorize charging our Deposit Account 07-0630 for any fees required or credits due and any extensions of time necessary to maintain the pendency of this application.

Applicants respectfully request that a timely Notice of Allowance be issued in this case.

Respectfully submitted,
GENENTECH, INC.

Date: January 26, 2007

By: _____
Deirdre L. Conley
Reg. No. 36,487
Telephone No. (650) 225-2066

232889

Leg 3
1059

REIVINDICACIONES

MODIFICADAS

1964
1060

LO QUE SE REIVINDICA ES:

Favor corregir las reivindicaciones como sigue, donde el tachado ~~00~~ indica la terminología suprimida y el subrayado 00 indica la terminología que se agrega.

LO QUE SE REINVINDICA ES:

1. (Se modifica por la presente) Un anticuerpo aislado que se fija a un polipéptido teniendo una identidad de secuencia de amino ácidos por lo menos de 80% con:

a) el polipéptido que aparece en ~~una cualquiera de~~ las Figuras ~~79 a 154 (SEQ ID Nos:79 154)~~ 115 (SEQ ID No.115);

b) el polipéptido que aparece en ~~una cualquiera de~~ las Figuras ~~79 a 154~~ (SEQ ID Nos:~~79 154~~ 115), que carece de su péptido de señal asociado;

c) un dominio extracelular del polipéptido que aparece en ~~una cualquiera de~~ las Figuras ~~79 a 154~~ 115 (SEQ ID Nos:~~79 154~~ 115), con su péptido de señal asociado;

d) un dominio extracelular del polipéptido que aparece en ~~una cualquiera de~~ las Figuras ~~79 a 154~~ 115 (SEQ ID Nos:~~79 154~~ 115), que carece de su péptido de señal asociado;

e) un polipéptido codificado por una secuencia nucleótida que aparece en ~~una cualquiera de~~ las Figuras ~~1 a 78A B~~ 37(SEQ ID Nos: ~~1 78~~ 37); o

f) un polipéptido codificado por la región de codificación de longitud total de una secuencia nucleótida que aparece en ~~una cualquiera de~~ las Figuras ~~1 a 78A B~~ 37(SEQ ID Nos: ~~1 78~~ 37).

2. (Se modifica por la presente) Un anticuerpo aislado que se fija a un polipéptido teniendo:

a) la secuencia de aminoácidos que aparece en ~~una cualquiera de~~ las Figuras ~~79 a 154~~ 115(SEQ ID Nos:~~79 154~~ 115);

b) la secuencia de amino ácidos que aparece en ~~una cualquiera de~~ las Figuras ~~79 a 154~~ 115(SEQ ID Nos:~~79 154~~ 115), que carece de su secuencia de péptido de señal asociado;

c) una secuencia de amino ácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), con su secuencia de péptido de señal asociado;

d) una secuencia de amino ácidos de un dominio extracelular del polipéptido que aparece en una cualquiera de las Figuras 79 a 154 (SEQ ID Nos:79-154), que carece de su secuencia de péptido de señal asociado;

e) una secuencia de amino ácidos codificada por una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78); o

f) una secuencia de aminoácidos codificada por la región de codificación de longitud total de una secuencia nucleótida que aparece en una cualquiera de las Figuras 1 a 78A-B (SEQ ID Nos: 1-78).

3. (Original) El anticuerpo de la Reivindicación 1 el cual es un anticuerpo monoclonal.

4. (Original) El anticuerpo de la Reivindicación 1 el cual es un fragmento de anticuerpo.

5. (Original) El anticuerpo de la Reivindicación 1 el cual es un anticuerpo quimérico o humanizado.

6. (Original) El anticuerpo de la Reivindicación 1 el cual se conjuga con un agente inhibidor de crecimiento.

7. (Original) El anticuerpo de la Reivindicación 1 el cual se conjuga con un agente citotóxico.

8. (Original) El anticuerpo de la Reivindicación 7, donde el agente citotóxico se selecciona entre el grupo conformado por toxinas, antibióticos, isótopos radioactivos y enzimas nucleolíticas.

9. (Original) El anticuerpo de la Reivindicación 7, donde el agente citotóxico es una toxina.

10. (Original) El anticuerpo de la Reivindicación 9, donde la toxina se selecciona entre el grupo conformado por auristatina, maitansinoide y calicheamicina.

11. (Original) El anticuerpo de la Reivindicación 9, donde la toxina es un maitansinoide.

12. (Original) El anticuerpo de la Reivindicación 1 el cual se produce en bacterias.

13. (Original) El anticuerpo de la Reivindicación 1 el cual se produce en células CHO.

14. (Original) El anticuerpo de la Reivindicación 1 que induce la muerte de una célula a la cual se fija.

15. (Original) El anticuerpo de la Reivindicación 1 el cual se marca en forma detectable.

16. (Original) El anticuerpo de la Reivindicación 9, donde la toxina se selecciona entre el grupo conformado por MMAE (mono-metil auristatina E), MMAF, (auristatina E valeril bencihidrazona) y AFP (Auristatina F fenileno diamina).

17. (Original) El anticuerpo de la Reivindicación 9, donde la toxina se adhiere en forma covalente al anticuerpo mediante un enlazante.

18. (Original) El anticuerpo de la Reivindicación 17, donde el enlazante se selecciona entre el grupo conformado por maleimidocaproilo (MC), valina-citrulina (val-cit, vc), citrulina (2-amino-5-ureido ácido pentanoico), PAB (p-aminobencilcarbamoilo), Me (N-metil-valina citrulina), MC(PEG)6-OH (maleimidocaproilo-polietilenglicol), SPP (N-Succinimidil 4-(2-

piridiltio)pentanoato), SMCC (N-Succinimidil 4-(N-maleimidometil) ciclohexano-1 carboxilato) y MC-vc-PAB.

19. (Original) El anticuerpo de la Reivindicación 16, donde la toxina es MMAE.

20. (Original) El anticuerpo de la Reivindicación 16, donde la toxina es MMAF.

21. (Original) El anticuerpo de la Reivindicación 19 0 20, donde el enlazante es MC-vc-PAB.

22. (Original) El anticuerpo de una cualquiera de las Reivindicaciones 1-21 donde el anticuerpo fija el polipéptido TAT188.

23. (Original) El anticuerpo de la Reivindicación 22, donde el anticuerpo inhibe la proliferación o promueve la muerte celular de una célula que expresa el TAT188.

24. (Original) El anticuerpo de la Reivindicación 23, donde la célula es una célula cancerosa.

25. (Original) El anticuerpo de la Reivindicación 24, donde la célula cancerosa se selecciona entre el grupo de mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado.

26. (Se modifica por la presente) Un anticuerpo aislado que compite con la fijación de los epítopos del polipéptido TAT188 unido por un anticuerpo producido por el hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No.— 6193), 12B9.1 (ATCC Acceso No.— 6194) y 12G12.1 (ATCC Acceso No.— 6195).

27. (Se modifica por la presente) Un anticuerpo aislado teniendo la actividad biológica de un anticuerpo producido por el hibridoma seleccionado entre el

grupo conformado por 3B5.1 (ATCC Acceso No.—6193), 12B9.1 (ATCC Acceso No.—6194) y 12G12.1 (ATCC Acceso No.—6195), donde la actividad biológica es la inhibición de la proliferación celular o la promoción de la muerte celular en una célula que expresa el TAT188.

28. (Se modifica por la presente) Un anticuerpo aislado que comprende en una región de determinante complementario (CDR) correspondiente, una secuencia de amino ácidos teniendo por lo menos 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de una secuencia de amino ácidos de por lo menos 1, 2, 3, 4, 5 o 6 de los CDR(s) del anticuerpo producido por un hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No.—6193), 12B9.1 (ATCC Acceso No.—6194) y 12G12.1 (ATCC Acceso No.—6195).

29. (Se modifica por la presente) El anticuerpo de una cualquiera de las Reivindicaciones 1-21, donde el anticuerpo comprende en una región de determinante complementario (CDR) correspondiente, una secuencia de amino ácidos teniendo por lo menos 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de una secuencia de amino ácidos de por lo menos 1, 2, 3, 4, 5 o 6 de los CDR(s) del anticuerpo producido por un hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No.—6193), 12B9.1 (ATCC Acceso No.—6194) y 12G12.1 (ATCC Acceso No.—6195).

30. (Se modifica por la presente) El anticuerpo de una cualquiera de las Reivindicaciones 1-21, donde el anticuerpo exhibe la actividad biológica de un anticuerpo producido por el hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No.—6193), 12B9.1 (ATCC Acceso No.—6194) y 12G12.1 (ATCC Acceso No.—6195), donde la actividad biológica es la inhibición de la proliferación celular o la promoción de la muerte celular en una célula que expresa el TAT188.

31. (Original) El anticuerpo de la Reivindicación 30, donde la célula es una célula cancerosa.

32. (Original) El anticuerpo de la Reivindicación 31, donde la célula cancerosa se selecciona entre el grupo conformado por mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado.

33. (Original) Un método para inhibir el crecimiento de una célula que expresa el TAT188, el método que comprende poner una célula en contacto con un anticuerpo de una cualquiera de las Reivindicaciones 1-21.

34. (Original) El método de la Reivindicación 33, donde la célula es una célula cancerosa.

35. (Original) El método de la Reivindicación 34, donde la célula cancerosa se selecciona entre el grupo conformado por mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado.

36. (Original) El método de la Reivindicación 35, donde la célula cancerosa es una célula de mamífero.

37. (Original) El método de la Reivindicación 36, donde la célula de mamífero es una célula humana.

38. (Original) Un método para inhibir el crecimiento de una célula que expresa el TAT188, el método que comprende poner una célula en contacto con un anticuerpo de una cualquiera de las Reivindicaciones 185-205, donde el anticuerpo comprende en una región de determinante complementario (CDR) correspondiente, una secuencia de amino ácidos teniendo por lo menos 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o 100% de una secuencia de amino ácidos de por lo menos 1, 2, 3, 4, 5 o 6 de las CDR(s) del anticuerpo producido por un hibridoma seleccionado entre el grupo conformado por 3B5.1 (ATCC Acceso No. PTA-6193), 12B9.1 (ATCC Acceso No. PTA-6194) y 12G12.1 (ATCC Acceso No. PTA-6195).

1569
1066

39. (Original) Un método para detectar el nivel del polipéptido TAT188 expresado en una célula de prueba en relación con una célula de control, el método que comprende:

a) poner en contacto la célula de prueba y la célula de control con un anticuerpo anti-TAT188 aislado de la Reivindicación 26;

b) detectar la fijación del anticuerpo; y

c) determinar la fijación relativa del anticuerpo a la célula de prueba y de control.

40. (Original) El método de la Reivindicación 39, donde la célula de prueba y la célula de control se someten a lisis.

41. (Original) El método de la Reivindicación 39, donde la célula de prueba está en un tejido.

42. (Original) El método de la Reivindicación 41, donde el tejido es un tejido de tumor.

43. (Original) El método de la Reivindicación 42, donde el tejido de tumor se selecciona entre el grupo conformado por mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado.

44. (Original) Un método para detectar el nivel del polipéptido TAT188 o un polipéptido teniendo una identidad de secuencia por lo menos de 80% con una secuencia de amino ácidos que se muestra en la Figura 115 (SEQ ID No: 115) en una célula de prueba en relación con una célula de control, el método que comprende:

a) poner en contacto la célula de prueba y la célula de control con un anticuerpo aislado de una cualquiera de las Reivindicaciones 1-5, 12-15 y 26-28

b) detectar la fijación del anticuerpo; y

c) determinar la fijación relativa del anticuerpo a la célula de prueba y de control.

1970
1067

45. (Original) El método de la Reivindicación 44, donde el nivel del polipéptido TAT188 en la célula de prueba es mayor que el de la célula de control.

46. (Original) El método de la Reivindicación 45, donde el método diagnostica el cáncer en un tejido que contiene o que tiene contenida la célula de prueba.

47. (Original) Un ARN de interferencia de unión TAT (ARNsi) que se fija a un ácido nucleico teniendo una identidad de secuencia por lo menos de 80% con:

a) una secuencia nucleótida que se muestra en la Figura 37 (SEQ ID No:37; y

b) el complemento de (a), donde el ARNsi reduce la expresión del TAT188.

48. (Original) Un vector de expresión que comprende el ARNsi de la Reivindicación 47.

49. (Original) El vector de expresión de la Reivindicación 48, donde el citado ARNsi está enlazado operacionalmente a las secuencias de control reconocidas por una célula hospedera transinfectada con el vector.

50. (Original) Una célula hospedera que comprende el vector de expresión de la Reivindicación 49.

51. (Se modifica por la presente) Una composición de sustancias que comprende:

a) el anticuerpo de la Reivindicación 1, ~~o~~

~~b) el ARNsi de la Reivindicación 47,~~

en combinación con un excipiente.

52. (Original) La composición de las sustancias de la Reivindicación 51, donde el citado excipiente es un excipiente farmacéuticamente aceptable.

53. (Original) Un artículo de manufactura:

a) un envase; y

b) la composición de las sustancias de la Reivindicación 51 contenida en el citado envase.

54. (Original) El artículo de manufactura de la Reivindicación 53, que además comprende una etiqueta adherida al citado envase, o un inserto en el paquete incluido con el citado envase, haciendo referencia al uso de la citada composición de sustancias para el tratamiento terapéutico o la detección diagnóstica de un cáncer.

55. (Original) Un método para inhibir el crecimiento de una célula cancerosa que expresa un polipéptido teniendo una identidad de secuencia de amino ácidos por lo menos de 80% con una secuencia de amino ácidos que se muestra en la Figura 115 (SEQ ID No: 115), el citado método que comprende poner en contacto la citada célula cancerosa con un ARNsi que se fija a un ácido nucleico codificando el amino ácido en la citada célula cancerosa, en esta forma inhibiendo el crecimiento de la citada célula cancerosa.

56. (Original) El método de la Reivindicación 55, donde el ácido nucleico tiene la secuencia que se muestra en la Figura 37 (SEQ ID No:37).

57. (Original) El método de la Reivindicación 44, donde la detección del nivel de expresión del polipéptido comprende emplear un anticuerpo en un análisis de inmunohistoquímica.

58. (Original) Un método para tratar o prevenir un desorden de proliferación celular asociado con una expresión o actividad aumentada de un polipéptido teniendo una identidad de secuencia de amino ácidos por lo menos de 80% con una secuencia de amino ácidos que se muestra en la Figura 115 (SEQ ID No: 115), el citado método que comprende administrar a un sujeto que necesite el citado tratamiento una cantidad efectiva de un antagonista de un polipéptido TAT188.

59. (Original) El método de la Reivindicación 58, donde el citado antagonista es un anticuerpo aislado de un polipéptido anti-TAT188 de una cualquiera de las Reivindicaciones 1-21 y 23-28.

60. (Original) El método de la Reivindicación 59, donde el desorden de proliferación celular es cáncer.

61. (Original) El método de la Reivindicación 60, donde el cáncer se selecciona entre el grupo conformado por mama, colon, recto, endometrio, riñón, pulmón, ovario, piel e hígado.

Solicitud No. 11/626,775 Expediente #P5001R1P1C1
Modif. Fechada Enero 26, 2007
Modificación Preliminar

Las reivindicaciones originales 3-26, 31-50, 52-61 están pendientes en la solicitud. Las reivindicaciones 1-2, 27-30 y 51 se modifican en la presente. No se están adicionando reivindicaciones nuevas en la modificación. Las modificaciones se apoyan en la especificación, por ejemplo en la página 33, líneas 4-10, página 34, líneas 18-21 y en la página 36, líneas 33-34 y en el recibo de depósito con el ATCC adjunto. En las modificaciones no se agrega tema.

Respetuosamente los solicitantes indican que las reivindicaciones cumplen los requisitos para ser admitidas. Por lo tanto, se solicita la determinación de una disposición final favorable. Se invita al Examinador a llamar por teléfono a Deirdre L. Conley, Ph.D. (Reg. 36.487) al número (650)225-2066 si considera que puede ser de ayuda para clarificar el tema y proseguir con el asunto.

Esta Modificación Preliminar se presenta con una carta remisoria. En el caso poco probable de que este documento se separe de la carta remisoria o si aplican unas tarifas, los solicitantes solicitan al Comisionado que autorice cobrar a nuestra Cuenta No. 07-0630 cualquier costo o deuda en que se incurra y cualquier prórroga de tiempo que se requiera para mantener esta solicitud pendiente.

1073
1070.

Respetuosamente los solicitantes solicitan el otorgamiento de una Notificación de Permiso para este caso.

Presentado respetuosamente,

GENENTECH, INC.

Fecha: Enero 26, 2007

Por: _____
Deirdre L. Conley
Reg. No. 36.487
Teléfono No. (650) 225-2066

Lot 4
1071

Sequence Listing

<110> GENENTECH, INC.

<120> Compositions and Methods for the Diagnosis and Treatment of Tumor

<130> P5001R1P1 PCT

<140> PCT/US2005/031798
<141> 2005-07-09

<150> US 10/936,626
<151> 2004-09-08

<150> US 10/177,488
<151> 2002-06-19

<150> US 60/299,500
<151> 2001-06-20

<150> US 60/301,880
<151> 2001-06-29

<150> US 60/323,268
<151> 2001-09-18

<160> 159

<210> 1
<211> 1039
<212> DNA
<213> Homo sapiens

<400> 1
ccaagaattc ggcacgaggg tgacccggac agctgtcctc tctgacacca 50

ccccggcctg cctctttgtt gccatgagag ctgcctacct cttcctgcta 100

ttcctgcctg caggcttgct ggctcagggc cagtatgacc tggacccgct 150

gccgccgttc cctgaccacg tccagtacac ccactatagc gaccagatcg 200

acaacccaga ctactatgat tatcaagagg tgactcctcg gccctccgag 250

gaacagttcc agttccagtc ccagcagcaa gtccaacagg aagtcatccc 300

agccccaacc ccagaaccag gaaatgcaga gctggagccc acagagcctg 350

ggcctcttga ctgccgtgag gaacagtacc cgtgcacccg cctctactcc 400

atacacaggc cttgcaaaca gtgtctcaac gaggtctgct tctacagcct 450

ccgccgtgtg tacgtcatta acaaggagat ctgtgttcgt acagtgtgtg 500

cccacgagga gctcctccga gctgacctct gtcgggacaa gttctccaaa 550

tgtggcgtga tggccagcag cggcctgtgc caatccgtgg cggcctcctg 600

tgccaggagc tgtgggagct gctagggtgg tgctggcatc ctgagtcctg 650

gccctcctgg gatctggggc cctcgggccc tgcctgacct ggtgcttttt 700

tccccatccc catgttcctt ttattctgta aaaagttagt ggactgcagc 750

cctgggggtt gcaggctgcg gtgcctcagg cccctccttc agcctgtggc 800

cacctctggg gcacgatggg ggctccccac tgcccagtct gcccctcggg 850

ttgggggagt atcccaggcc tctctgtggg acctgggccc ctgacgggcc 900

ttctcagccc gttttgagga cagacagtcc cccgaggtag gctacatccc 950

cccaccccag ctggtctgct tggatttcct acagcccccg tgggcatgga 1000

ccacctttat tttatacaaa attaaaaaca agtttttac 1039

1073
1072

<210> 2
<211> 5062
<212> DNA
<213> Homo sapiens

<400> 2
cactaacgct cttcctagtc cccgggccaa ctcggacagt ttgctcattt 50

attgcaacgg tcaaggctgg cttgtgccag aacggcgcgc gcgcgacgca 100

cgcacacaca cgggggggaaa ctttttttaaa aatgaaaggc tagaagagct 150

cagcggcggc gcgggccgtg cgcgagggct ccggagctga ctcgccgagg 200

caggaaatcc ctccggtcgc gacgcccggc cccgctcggc gcccgcgtgg 250

gatggtgcag cgctcgccgc cgggcccgag agctgctgca ctgaaggccg 300

gcgacgatgg cagcgcgccc gctgcccgtg tcccccgccc gcgccctcct 350

gctcgccctg gccggtgctc tgctcgcgcc ctgcgaggcc cgaggggtga 400

gcttatggaa cgaaggaaga gctgatgaag ttgtcagtgc ctctgttcgg 450

agtggggacc tctggatccc agtgaagagc ttcgactcca agaatcatcc 500

agaagtgctg aatattcgac tacaacggga aagcaaagaa ctgatcataa 550

atctggaaag aaatgaaggt ctcattgcca gcagtttcac ggaaacccac 600

tatctgcaag acggtactga tgtctccctc gctcgaaatt acacggtaat 650

tctgggtcac tgttactacc atggacatgt acggggatat tctgattcag 700

cagtcagtct cagcacgtgt tctggtctca ggggacttat tgtgtttgaa 750

aatgaaagct atgtcttaga accaatgaaa agtgcaacca acagatacaa 800

actcttccca gcgaagaagc tgaaaagcgt ccggggatca tgtggatcac 850

atcacaacac accaaacctc gctgcaaaga atgtgtttcc accaccctct 900

cagacatggg caagaaggca taaaagagag accctcaagg caactaagta 950

tgtggagctg gtgatcgtgg cagacaaccg agagtttcag aggcaaggaa 1000

aagatctgga aaaagttaag cagcgattaa tagagattgc taatcacgtt 1050

gacaagtttt acagaccact gaacattcgg atcgtgttgg taggcgtgga 1100

agtgtggaat gacatggaca aatgctctgt aagtcaggac ccattcacca 1150

gcctccatga atttctggac tggaggaaga tgaagcttct acctcgcaaa 1200

tcccatgaca atgcgcagct tgtcagtggg gtttatttcc aagggaccac 1250

catcggcatg gccccaatca tgagcatgtg cacggcagac cagtctgggg 1300

gaattgtcat ggaccattca gacaatcccc ttggtgcagc cgtgaccctg 1350

gcacatgagc tgggccacaa tttcgggatg aatcatgaca cactggacag 1400

gggctgtagc tgtcaaatgg cggttgagaa aggaggctgc atcatgaacg 1450

cttccaccgg gtacccattt cccatggtgt tcagcagttg cagcaggaag 1500

gacttggaga ccagcctgga gaaaggaatg ggggtgtgcc tgtttaacct 1550

gccggaagtc agggagtctt tcgggggcca gaagtgtggg aacagatttg 1600

tggaagaagg agaggagtgt gactgtgggg agccagagga atgtatgaat 1650

cgctgctgca atgccaccac ctgtaccctg aagccggacg ctgtgtgcgc 1700

acatgggctg tgctgtgaag actgccagct gaagcctgca ggaacagcgt 1750

gcagggactc cagcaactcc tgtgacctcc cagagttctg cacaggggcc 1800

agccctcact gcccagccaa cgtgtacctg cacgatgggc actcatgtca 1850

ggatgtggac ggctactgct acaatggcat ctgccagact cacgagcagc 1900

agtgtgtcac actctgggga ccaggtgcta aacctgcccc tgggatctgc 1950

tttgagagag tcaattctgc aggtgatcct tatggcaact gtggcaaagt 2000

ctcgaagagt tcctttgcca aatgcgagat gagagatgct aaatgtggaa 2050

aaatccagtg tcaaggaggt gccagccggc cagtcattgg taccaatgcc 2100

gtttccatag aaacaaacat ccccctgcag caaggaggcc ggattctgtg 2150

ccgggggacc cacgtgtact tgggcgatga catgccggac ccagggcttg 2200

tgcttgcagg cacaaagtgt gcagatggaa aaatctgcct gaatcgtcaa 2250

tgtcaaaata ttagtgtctt tggggttcac gagtgtgcaa tgcagtgcca 2300

cggcagaggg gtgtgcaaca acaggaagaa ctgccactgc gaggcccact 2350

gggcacctcc cttctgtgac aagtttggct ttggaggaag cacagacagc 2400

ggccccatcc ggcaagcaga taaccaaggt ttaaccatag gaattctggt 2450

gaccatcctg tgtcttcttg ctgccggatt tgtggtttat ctcaaaagga 2500

agaccttgat acgactgctg tttacaaata agaagaccac cattgaaaaa 2550

ctaaggtgtg tgcgcccttc ccggccaccc cgtggcttcc aaccctgtca 2600

ggctcacctc ggccaccttg gaaaaggcct gatgaggaag ccgccagatt 2650

cctacccacc gaaggacaat cccaggagat tgctgcagtg tcagaatgtt 2700

gacatcagca gacccctcaa cggcctgaat gtccctcagc cccagtcaac 2750

tcagcgagtg cttcctcccc tccaccgggc cccacgtgca cctagcgtcc 2800

ctgccagacc cctgccagcc aagcctgcac ttaggcaggc ccaggggacc 2850

tgtaagccaa acccccctca gaagcctctg cctgcagatc ctctggccag 2900

aacaactcgg ctcactcatg ccttggccag gaccccagga caatgggaga 2950

ctgggctccg cctggcaccc ctcagacctg ctccacaata tccacaccaa 3000

gtgcccagat ccacccacac cgcctatatt aagtgagaag ccgacacctt 3050

ttttcaacag tgaagacaga agtttgcact atctttcagc tccagttgga 3100

gtttttgta ccaactttta ggatttttt taatgtttaa aacatcatta 3150

ctataagaac tttgagctac tgccgtcagt gctgtgctgt gctatggtgc 3200

tctgtctact tgcacaggta cttgtaaatt attaatttat gcagaatgtt 3250

gattacagtg cagtgcgctg tagtaggcat ttttaccatc actgagtttt 3300

ccatggcagg aaggcttgtt gtgcttttag tattttagtg aacttgaaat 3350

atcctgcttg atgggattct ggacaggatg tgtttgcttt ctgatcaagg 3400

ccttattgga aagcagtccc ccaactaccc ccagctgtgc ttatggtacc 3450

agatgcagct caagagatcc caagtagaat ctcagttgat tttctggatt 3500

ccccatctca ggccagagcc aaggggcttc aggtccaggc tgtgtttggc 3550

tttcagggag gccctgtgcc ccttgacaac tggcaggcag gctcccaggg 3600

acacctggga gaaatctggc ttctggccag gaagctttgg tgagaacctg 3650

ggttgcagac aggaatctta aggtgtagcc acaccaggat agagactgga 3700

acactagaca agccagaact tgaccctgag ctgaccagcc gtgagcatgt 3750

ttggaagggg tctgtagtgt cactcaaggc ggtgcttgat agaaatgcca 3800

agcacttctt tttctcgctg tcctttctag agcactgcca ccagtaggtt 3850

atttagcttg ggaaaggtgg tgtttctgta agaaacctac tgcccaggca 3900

ctgcaaaccg ccacctccct atactgcttg gagctgagca aatcaccaca 3950

aactgtaata caatgatcct gtattcagac agatgaggac tttccatggg 4000

accacaacta ttttcagatg tgaaccatta accagatcta gtcaatcaag 4050

tctgtttact gcaaggttca acttattaac aattaggcag actctttatg 4100

cttgcaaaaa ctacaaccaa tggaatgtga tgttcatggg tatagttcat 4150

gtctgctatc attattcgta gatattggac aaagaacctt ctctatgggg 4200

catcctcttt ttccaacttg gctgcaggaa tctttaaaag atgcttttaa 4250

cagagtctga acctatttct taaacacttg caacctacct gttgagcatc 4300

acagaatgtg ataaggaaat caacttgctt atcaacttcc taaatattat 4350

gagatgtggc ttgggcagca tcccccttgaa ctcttcactc ttcaaatgcc 4400

tgactaggga gccatgtttc acaaggtctt taaagtgact aatggcatga 4450

gaaatacaaa aatactcaga taaggtaaaa tgccatgatg cctctgtctt 4500

ctggactggt tttcacatta gaagacaatt gacaacagtt acataattca 4550

ctctgagtgt tttatgagaa agccttcttt tggggtcaac agttttccta 4600

tgctttgaaa cagaaaaata tgtaccaaga atcttggttt gccttccaga 4650

aaacaaaact gcatttcact ttcccggtgt tccccactgt atctaggcaa 4700

catagtattc atgactatgg ataaactaaa cacgtgacac aaacacacac 4750

aaaagggaac ccagctctaa tacattccaa ctcgtatagc atgcatctgt 4800

ttattctata gttattaagt tctttaaaat gtaaagccat gctggaaaat 4850

aatactgctg agatacatac agaattactg taactgatta cacttggtaa 4900

ttgtactaaa gccaaacata tatatactat taaaaaggtt tacagaattt 4950

tatggtgcat tacgtgggca ttgtcttttt agatgcccaa atccttagat 5000

ctggcatgtt agcccttcct ccaattataa gaggatatga accaaaaaaa 5050

aaaaaaaaaa aa 5062

<210> 3
<211> 1150
<212> DNA
<213> Homo sapiens

<400> 3
ggagtttcaa agaagcgcta gtaaggtctc tgagatcctt gcactagcta 50

catcctcagg gtaggaggaa gatggcttcc agaagcatgc ggctgctcct 100

attgctgagc tgcctggcca aaacaggagt cctgggtgat atcatcatga 150

gacccagctg tgctcctgga tggttttacc acaagtccaa ttgctatggt 200

tacttcagga agctgaggaa ctggtctgat gccgagctcg agtgtcagtc 250

ttacggaaac ggagcccacc tggcatctat cctgagttta aaggaagcca 300

gcaccatagc agagtacata agtggctatc agagaagcca gccgatatgg 350

attggcctgc acgacccaca gaagaggcag cagtggcagt ggattgatgg 400

ggccatgtat ctgtacagat cctggtctgg caagtccatg ggtgggaaca 450

agcactgtgc tgagattagc tccaataaca acttttaac ttggagcagc 500

aacgaatgca acaagcgcca acacttcctg tgcaagtacc gaccatagag 550

caagaatcaa gattctgcta actcctgcac agccccgtcc tcttcctttc 600

tgctagcctg gctaaatctg ctcattattt cagaggggaa acctagcaaa 650

ctaagagtga taagggccct actacactgg ctttttagg cttagagaca 700

gaaactttag cattggccca gtagtggctt ctagctctaa atgtttgccc 750

cgccatccct ttccacagta tccttcttcc ctcctcccct gtctctggct 800

gtctcgagca gtctagaaga gtgcatctcc agcctatgaa acagctgggt 850

ctttggccat aagaagtaaa gatttgaaga cagaaggaag aaactcagga 900

gtaagcttct agacccctc agcttctaca cccttctgcc ctctctccat 950

tgcctgcacc ccaccccagc cactcaactc ctgcttgttt ttcctttggc 1000

cataggaagg tttaccagta gaatccttgc taggttgatg tgggccatac 1050

attcctttaa taaaccattg tgtacataaa aaaaaaaaaa aaaaaaaaaa 1100

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1150

<210> 4
<211> 474
<212> DNA
<213> Homo sapiens

<400> 4
ggggagcaga gaggaggcaa tggccaccat ggagaacaag gtgatctgcg 50

ccctggtcct ggtgtccatg ctggccctcg gcaccctggc cgaggcccag 100

acagagacgt gtacagtggc cccccgtgaa agacagaatt gtggttttcc 150

tggtgtcacg ccctcccagt gtgcaaataa gggctgctgt ttcgacgaca 200

ccgttcgtgg ggtcccctgg tgcttctatc ctaataccat cgacgtccct 250

ccagaagagg agtgtgaatt ttagacactt ctgcagggat ctgcctgcat 300

cctgacgcgg tgccatcccc agcacggtga ttagtcccag agctcggctg 350

ccacctccac cggacacctc agacacgctt ctgcagctgt gcctcggctc 400

acaacacaga ttgactgctc tgactttgac tactcaaaat tggcctaaaa 450

attaaaagag ctcgatatta aaaa 474

<210> 5
<211> 3226
<212> DNA
<213> Homo sapiens

<400> 5
aatccatctg agaatatgct gccacaaata ccctttttgc tgctagtatc 50

cttgaacttg gttcatggag tgttttacgc tgaacgatac caaacgccca 100

caggcataaa aggcccacta cccaacacca agacacagtt cttcattccc 150

tacaccataa agagtaaagg tatagcagta agaggagagc aaggtactcc 200

tggtccacca ggccctgctg gacctcgagg gcacccaggt ccttctggac 250

caccaggaaa accaggctac ggaagtcctg gactccaagg agagccaggg 300

ttgccaggac caccgggacc atcagctgta gggaaaccag gtgtgccagg 350

actcccagga aaaccaggag agagaggacc atatggacca aaaggagatg 400

ttggaccagc tggcctacca ggaccccggg gcccaccagg accacctgga 450

atccctggac cggctggaat ttctgtgcca ggaaaacctg gacaacaggg 500

acccacagga gccccaggac ccaggggctt tcctggagaa aagggtgcac 550

caggagtccc tggtatgaat ggacagaaag gggaaatggg atatggtgct 600

cctggtcgtc caggtgagag gggtcttcca ggccctcagg gtcccacagg 650

accatctggc cctcctggag tgggaaaaag aggtgaaaat ggggttccag 700

gacagccagg catcaaaggt gatagaggtt ttccgggaga aatgggacca 750

attggcccac caggtcccca aggccctcct ggggaacgag ggccagaagg 800

cattggaaag ccaggagctg ctggagcccc aggccagcca gggattccag 850

gaacaaaagg tctccctggg gctccaggaa tagctgggcc cccagggcct 900

cctggctttg ggaaaccagg cttgccaggc ctgaagggag aaagaggacc 950

tgctggcctt cctgggggtc caggtgccaa aggggaacaa gggccagcag 1000

gtcttcctgg gaagccaggt ctgactggac cccctgggaa tatgggaccc 1050

caaggaccaa aaggcatccc gggtagccat ggtctcccag gccctaaagg 1100

tgagacaggg ccagctgggc ctgcaggata ccctggggct aagggtgaaa 1150

ggggttcccc tgggtcagat ggaaaaccag ggtacccagg aaaaccaggt 1200

ctcgatggtc ctaagggtaa cccagggtta ccaggtccaa aaggtgatcc 1250

tggagttgga ggacctcctg gtctcccagg ccctgtgggc ccagcaggag 1300

caaagggaat gcccggacac aatggagagg ctggcccaag aggtgccct 1350

ggaataccag gtactagagg ccctattggg ccaccaggca ttccaggatt 1400

ccctgggtct aaaggggatc caggaagtcc cggtcctcct ggcccagctg 1450

gcatagcaac taagggcctc aatggaccca ccgggccacc agggcctcca 1500

ggtccaagag gcccctctgg agagcctggt cttccagggc cccctgggcc 1550

tccaggccca ccaggtcaag cagtcatgcc tgagggtttt ataaaggcag 1600

gccaaaggcc cagtctttct gggacccctc ttgttagtgc caaccagggg 1650

gtaacaggaa tgcctgtgtc tgcttttact gttattctct ccaaagctta 1700

cccagcaata ggaactccca taccatttga taaaattttg tataacaggc 1750

aacagcatta tgacccaagg actggaatct ttacttgtca gataccagga 1800

atatactatt tttcatacca cgtgcatgtg aaagggactc atgtttgggt 1850

aggcctgtat aagaatggca cccctgtaat gtacacctat gatgaataca 1900

ccaaaggcta cctggatcag gcttcaggga gtgccatcat cgatctcaca 1950

gaaaatgacc aggtgtggct ccagcttccc aatgccgagt caaatggcct 2000

atactcctct gagtatgtcc actcctcttt ctcaggattc ctagtggctc 2050

caatgtgagt acaccccaca gagctaatct aaatcttgtg ctagaaaaag 2100

cattctctaa ctctacccca ccctacaaaa tgcatatgga ggtaggctga 2150

aaagaatgta atttttattt tctgaaatac agatttgagc tatcagacca 2200

acaaaccttc cccctgaaaa gtgagcagca acgtaaaaac gtatgtgaag 2250

cctctcttga atttctagtt agcaatctta aggctcttta aggttttctc 2300

caatattaaa aaatatcacc aaagaagtcc tgctatgtta aaaacaaaca 2350

acaaaaaaca aagcaacaaa aaaaaaaatt aaaaaaaaaa acagaaatag 2400

agctctaagt tatgtgaaat ttgatttgag aaactcggca tttccttttt 2450

aaaaaagcct gtttctaact atgaatatga gaacttctag gaaacatcca 2500

ggaggtatca tataactttg tagaacttaa atacttgaat attcaaattt 2550

aaaagacact gtatccccta aaatatttct gatggtgcac tactctgagg 2600

cctgtatggc ccctttcatc aatatctatt caaatataca ggtgcatata 2650

tacttgttaa agctcttata taaaaaagcc ccaaaatatt gaagttcatc 2700

tgaaatgcaa ggtgctttca tcaatgaacc ttttcaaaac ttttctatga 2750

ttgcagagaa gctttttata tacccagcat aacttggaaa caggtatctg 2800

acctattctt atttagttaa cacaagtgtg attaatttga tttctttaat 2850

tccttattga atcttatgtg atatgatttt ctggatttac agaacattag 2900

cacatgtacc ttgtgcctcc cattcaagtg aagttataat ttacactgag 2950

ggtttcaaaa ttcgactaga agtggagata tattatttat ttatgcactg 3000

tactgtattt ttatattgct gtttaaaact tttaagctgt gcctcactta 3050

ttaaagcaca aaatgtttta cctactcctt atttacgaca caataaaata 3100

acatcaatag atttttaggc tgaattaatt tgaaagcagc aatttgctgt 3150

tctcaaccat tctttcaagg cttttcattc gacacaataa aataacatca 3200

atagattttt agggatgggt ggcttt 3226

<210> 6
<211> 1750
<212> DNA
<213> Homo sapiens

<400> 6
catgccgctg ccgccgctgc tgctgttgct cctggcggcg ccttggggac 50

gggcagttcc ctgtgtctct ggtggtttgc ctaaacctgc aaacatcacc 100

ttcttatcca tcaacatgaa gaatgtccta caatggactc caccagaggg 150

tcttcaagga gttaaagtta cttacactgt gcagtatttc atatatgggc 200

aaaagaaatg gctgaataaa tcagaatgca gaaatatcaa tagaacctac 250

tgtgatcttt ctgctgaaac ttctgactac gaacaccagt attatgccaa 300

agttaaggcc atttggggaa caaagtgttc caaatgggct gaaagtggac 350

ggttctatcc ttttttagaa acacaaattg gcccaccaga ggtggcactg 400

actacagatg agaagtccat ttctgttgtc ctgacagctc cagagaagtg 450

gaagagaaat ccagaagacc ttcctgtttc catgcaacaa atatactcca 500

atctgaagta taacgtgtct gtgttgaata ctaaatcaaa cagaacgtgg 550

tcccagtgtg tgaccaacca cacgctggtg ctcacctggc tggagccgaa 600

cactctttac tgcgtacacg tggagtcctt cgtcccaggg ccccctcgcc 650

gtgctcagcc ttctgagaag cagtgtgcca ggactttgaa agatcaatca 700

tcagagttca aggctaaaat catcttctgg tatgttttgc ccatatctat 750

taccgtgttt ctttttttctg tgatgggcta ttccatctac cgatatatcc 800

acgttggcaa agagaaacac ccagcaaatt tgattttgat ttatggaaat 850

gaatttgaca aaagattctt tgtgcctgct gaaaaaatcg tgattaactt 900

tatcaccctc aatatctcgg atgattctaa aatttctcat caggatatga 950

gtttactggg aaaaagcagt gatgtatcca gccttaatga tcctcagccc 1000

agcgggaacc tgaggccccc tcaggaggaa gaggaggtga aacatttagg 1050

gtatgcttcg catttgatgg aaatttttg tgactctgaa gaaaacacgg 1100

aaggtacttc tctcacccag caagagtccc tcagcagaac aataccccg 1150

gataaaacag tcattgaata tgaatatgat gtcagaacca ctgacatttg 1200

tgcggggcct gaagagcagg agctcagttt gcaggaggag gtgtccacac 1250

aaggaacatt attggagtcg caggcagcgt tggcagtctt gggcccgcaa 1300

acgttacagt actcatacac ccctcagctc caagacttag acccccctggc 1350

gcaggagcac acagactcgg aggaggggcc ggaggaagag ccatcgacga 1400

ccctggtcga ctgggatccc caaactggca ggctgtgtat tccttcgctg 1450

tccagcttcg accaggattc agagggctgc gagccttctg agggggatgg 1500

gctcggagag gagggtcttc tatctagact ctatgaggag ccggctccag 1550

acaggccacc aggagaaaat gaaacctatc tcatgcaatt catggaggaa 1600

tgggggttat atgtgcagat ggaaaactga tgccaacact tccttttgcc 1650

ttttgtttcc tgtgcaaaca agtgagtcac ccctttgatc ccagccataa 1700

agtacctggg atgaaagaag ttttttccag tttgtcagtg tctgtgagaa 1750

<210> 7
<211> 1841
<212> DNA
<213> Homo sapiens

<400> 7
agcgggcgtg ggactgagca gtctgctgcc ccccgacatg tgacccagcc 50

ccgccgccca tgcgggctcc cggccgcccg gccctgcggc cgctgccgct 100

gccgccgctg ctgctgttgc tcctggcggc gccttgggga cgggcagttc 150

cctgtgtctc tggtggtttg cctaaacctg caaacatcac cttcttatcc 200

atcaacatga agaatgtcct acaatggact ccaccagagg gtcttcaagg 250

agttaaagtt acttacactg tgcagtattt catatatggg caaaagaaat 300

ggctgaataa atcagaatgc agaaatatca atagaaccta ctgtgatctt 350

tctgctgaaa cttctgacta cgaacaccag tattatgcca aagttaaggc 400

catttgggga acaaagtgtt ccaaatgggc tgaaagtgga cggttctatc 450

ctttttttaga aacacaaatt ggcccaccag aggtggcact gactacagat 500

gagaagtcca tttctgttgt cctgacagct ccagagaagt ggaagagaaa 550

tccagaagac cttcctgttt ccatgcaaca aatatactcc aatctgaagt 600

ataacgtgtc tgtgttgaat actaaatcaa acagaacgtg gtcccagtgt 650

gtgaccaacc acacgctggt gctcacctgg ctggagccga acactcttta 700

ctgcgtacac gtggagtcct tcgtcccagg gccccctcgc cgtgctcagc 750

cttctgagaa gcagtgtgcc aggactttga aagatcaatc atcagagttc 800

aaggctaaaa tcatcttctg gtatgttttg cccatatcta ttaccgtgtt 850

tctttttttct gtgatgggct attccatcta ccgatatatc cacgttggca 900

aagagaaaca cccagcaaat ttgattttga tttatggaaa tgaatttgac 950

aaaagattct ttgtgcctgc tgaaaaaatc gtgattaact ttatcaccct 1000

caatatctcg gatgattcta aaatttctca tcaggatatg agtttactgg 1050

gaaaaagcag tgatgtatcc agccttaatg atcctcagcc cagcgggaac 1100

ctgaggcccc ctcaggagga agaggaggtg aaacatttag ggtatgcttc 1150

gcatttgatg gaaatttttt gtgactctga agaaaacacg gaaggtactt 1200

ctctcaccca gcaagagtcc ctcagcagaa caatacccc ggataaaaca 1250

gtcattgaat atgaatatga tgtcagaacc actgacattt gtgcggggcc 1300

tgaagagcag gagctcagtt tgcaggagga ggtgtccaca caaggaacat 1350

tattggagtc gcaggcagcg ttggcagtct tgggcccgca aacgttacag 1400

tactcataca cccctcagct ccaagactta gacccctgg cgcaggagca 1450

cacagactcg gaggaggggc cggaggaaga gccatcgacg accctggtcg 1500

actgggatcc ccaaactggc aggctgtgta ttccttcgct gtccagcttc 1550

gaccaggatt cagagggctg cgagccttct gaggggatg ggctcggaga 1600

ggagggtctt ctatctagac tctatgagga gccggctcca gacaggccac 1650

caggagaaaa tgaaacctat ctcatgcaat tcatggagga atgggggtta 1700

tatgtgcaga tggaaaactg atgccaacac ttccttttgc cttttgtttc 1750

ctgtgcaaac aagtgagtca ccccttgat cccagccata aagtacctgg 1800

gatgaaagaa gttttttcca gtttgtcagt gtctgtgaga a 1841

<210> 8
<211> 536
<212> DNA
<213> Homo sapiens

<400> 8
gcttggtcta gtgcccattt actctggact ccggatggct gccgcacgcc 50

tctgcctctc cctgctgctc ctgtccacct gcgtggctct gttactacag 100

ccactgctgg gtgcccaggg agccccactg gagccagtgt acccagggga 150

caatgccaca ccagagcaga tggcccagta tgcagctgat ctccgtagat 200

acatcaacat gctgaccagg cctaggtatg ggaaaagaca caaagaggac 250

acgctggcct tctcggagtg ggggtccccg catgctgctg tccccaggga 300

gctcagcccg ctggacttat aatgccacct tctgtctcct acgactccat 350

gagcagcgcc agcccagctc tcccctctgc acccttggct ctggccaaag 400

cttgctccct gctcccacac aggctcaata aagcaagtca aagccaaaaa 450

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 500

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 536

<210> 9
<211> 2588
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 1940
<223> unknown base

<400> 9
gaaggacagc agggccaaca gtcacagcag ccctgaccag agcattcctg 50

gagctcaagc tcctctacaa agaggtggac agagaagaca gcagagacca 100

tgggaccccc ctcagcccct ccctgcagat tgcatgtccc ctggaaggag 150

gtcctgctca cagcctcact tctaaccttc tggaacccac ccaccactgc 200

caagctcact attgaatcca cgccgttcaa tgtcgcagag gggaaggagg 250

ttcttctact cgcccacaac ctgccccaga atcgtattgg ttacagctgg 300

tacaaaggcg aaagagtgga tggcaacagt ctaattgtag gatatgtaat 350

aggaactcaa caagctaccc cagggcccgc atacagtggt cgagagacaa 400

tataccccaa tgcatccctg ctgatccaga acgtcaccca gaatgacaca 450

ggattctata ccctacaagt cataaagtca gatcttgtga atgaagaagc 500

aaccggacag ttccatgtat acccggagct gcccaagccc tccatctcca 550

gcaacaactc caaccccgtg gaggacaagg atgctgtggc cttcacctgt 600

gaacctgagg ttcagaacac aacctacctg tggtgggtaa atggtcagag 650

cctcccggtc agtcccaggc tgcagctgtc caatggcaac atgaccctca 700

ctctactcag cgtcaaaagg aacgatgcag gatcctatga atgtgaaata 750

cagaacccag cgagtgccaa ccgcagtgac ccagtcaccc tgaatgtcct 800

ctatggccca gatggcccca ccatttcccc ctcaaaggcc aattaccgtc 850

caggggaaaa tctgaacctc tcctgccacg cagcctctaa cccacctgca 900

cagtactctt ggtttatcaa tgggacgttc cagcaatcca cacaagagct 950

ctttatcccc aacatcactg tgaataatag cggatcctat atgtgccaag 1000

cccataactc agccactggc ctcaatagga ccacagtcac gatgatcaca 1050

gtctctggaa gtgctcctgt cctctcagct gtggccaccg tcggcatcac 1100

gattggagtg ctggccaggg tggctctgat atagcagccc tggtgtattt 1150

tcgatatttc aggaagactg gcagattgga ccagaccctg aattcttcta 1200

gctcctccaa tcccatttta tcccatggaa ccactaaaaa caaggtctgc 1250

tctgctcctg aagccctata tgctggagat ggacaactca atgaaaattt 1300

aaagggaaaa ccctcaggcc tgaggtgtgt gccactcaga gacttcacct 1350

aactagagac aggcaaactg caaaccatgg tgagaaattg acgacttcac 1400

actatggaca gcttttccca agatgtcaaa acaagactcc tcatcatgat 1450

aaggctctta ccccctttta atttgtcctt gcttatgcct gcctctttcg 1500

cttggcagga tgatgctgtc attagtattt cacaagaagt agcttcagag 1550

ggtaacttaa cagagtatca gatctatctt gtcaatccca acgttttaca 1600

taaaataaga gatcctttag tgcacccagt gactgacatt agcagcatct 1650

ttaacacagc cgtgtgttca aatgtacagt ggtccttttc agagttggac 1700

ttctagactc acctgttctc actccctgtt ttaattcaac ccagccatgc 1750

aatgccaaat aatagaattg ctccctacca gctgaacagg gaggagtctg 1800

tgcagtttct gacacttgtt gttgaacatg gctaaataca atgggtatcg 1850

ctgagactaa gttgtagaaa ttaacaaatg tgctgcttgg ttaaaatggc 1900

tacactcatc tgactcattc tttattctat tttagttggn ttgtatcttg 1950

cctaaggtgc gtagtccaac tcttggtatt accctcctaa tagtcatact 2000

agtagtcata ctccctggtg tagtgtattc tctaaaagct ttaaatgtct 2050

gcatgcagcc agccatcaaa tagtgaatgg tctctctttg gctggaatta 2100

caaaactcag agaaatgtgt catcaggaga acatcataac ccatgaagga 2150

taaaagcccc aaatggtggt aactgataat agcactaatg ctttaagatt 2200

tggtcacact ctcacctagg tgagcgcatt gagccagtgg tgcataaatg 2250

ctacatactc caactgaaat gttaaggaag aagatagatc caattaaaaa 2300

aaattaaaac caatttaaaa aaaaaaagaa cacaggagat tccagtctac 2350

ttgagttagc ataatacaga agtcccctct actttaactt ttacaaaaaa 2400

gtaacctgaa ctaatctgat gttaaccaat gtatttattt ctgtggttct 2450

gtttccttgt tccaatttga caaaacccac tgttcttgta ttgtattgcc 2500

caggggggagc tatcactgta cttgtagagt ggtgctgctt taattcataa 2550

atcacaaata aaagccaatt agctctaaaa aaaaaaaa 2588

<210> 10
<211> 3963
<212> DNA
<213> Homo sapiens

<400> 10
gacctggggg ggcgagggcc ccccaaactc agttcggatc ctacccgagt 50

gaggcggcgc catggagctc cgggtgctgc tctgctgggc ttcgttggcc 100

gcagctttgg aagagaccct gctgaacaca aaattggaaa ctgctgatct 150

gaagtgggtg acattccctc aggtggacgg gcagtgggag gaactgagcg 200

gcctggatga ggaacagcac agcgtgcgca cctacgaagt gtgtgacgtg 250

cagcgtgccc cgggccaggc ccactggctt cgcacaggtt gggtcccacg 300

gcggggcgcc gtccacgtgt acgccacgct gcgcttcacc atgctcgagt 350

gcctgtccct gcctcgggct gggcgctcct gcaaggagac cttcaccgtc 400

ttctactatg agagcgatgc ggacacggcc acggccctca cgccagcctg 450

gatggagaac ccctacatca aggtggacac ggtggccgcg gagcatctca 500

cccggaagcg ccctggggcc gaggccaccg ggaaggtgaa tgtcaagacg 550

ctgcgtctgg gaccgctcag caaggctggc ttctacctgg ccttccagga 600

ccagggtgcc tgcatggccc tgctatccct gcacctcttc tacaaaaagt 650

gcgcccagct gactgtgaac ctgactcgat tcccggagac tgtgcctcgg 700

gagctggttg tgcccgtggc cggtagctgc gtggtggatg ccgtccccgc 750

ccctggcccc agccccagcc tctactgccg tgaggatggc cagtgggccg 800

aacagccggt cacgggctgc agctgtgctc cggggttcga ggcagctgag 850

gggaacacca agtgccgagc ctgtgcccag ggcaccttca agcccctgtc 900

aggagaaggg tcctgccagc catgcccagc caatagccac tctaacacca 950

ttggatcagc cgtctgccag tgccgcgtcg ggtacttccg ggcacgcaca 1000

gacccccggg gtgcaccctg caccacccct ccttcggctc cgcggagcgt 1050

ggtttcccgc ctgaacggct cctccctgca cctggaatgg agtgcccccc 1100

tggagtctgg tggccgagag gacctcacct acgccctccg ctgccgggag 1150

tgccgacccg gaggctcctg tgcgccctgc ggggggagacc tgacttttga 1200

ccccggcccc cgggacctgg tggagccctg ggtggtggtt cgagggctac 1250

gtcctgactt cacctatacc tttgaggtca ctgcattgaa cggggtatcc 1300

tccttagcca cggggcccgt cccatttgag cctgtcaatg tcaccactga 1350

ccgagaggta cctcctgcag tgtctgacat ccgggtgacg cggtcctcac 1400

ccagcagctt gagcctggcc tgggctgttc cccgggcacc cagtggggct 1450

gtgctggact acgaggtcaa ataccatgag aagggcgccg agggtcccag 1500

cagcgtgcgg ttcctgaaga cgtcagaaaa ccgggcagag ctgcgggggc 1550

tgaagcgggg agccagctac ctggtgcagg tacgggcgcg ctctgaggcc 1600

ggctacgggc ccttcggcca ggaacatcac agccagaccc aactggatga 1650

gagcgagggc tggcgggagc agctggccct gattgcgggc acggcagtcg 1700

tgggtgtggt cctggtcctg gtggtcattg tggtcgcagt tctctgcctc 1750

aggaagcaga gcaatgggag agaagcagaa tattcggaca aacacggaca 1800

gtatctcatc ggacatggta ctaaggtcta catcgacccc ttcacttatg 1850

aagaccctaa tgaggctgtg agggaatttg caaaagagat cgatgtctcc 1900

tacgtcaaga ttgaagaggt gattggtgca ggtgagtttg gcgaggtgtg 1950

tcgggggcgg ctcaaggccc cagggaagaa ggagagctgt gtggcaatca 2000

agaccctgaa gggtggctac acggagcggc agcggcgtga gtttctgagc 2050

gaggcctcca tcatgggcca gttcgagcac cccaatatca tccgcctgga 2100

gggcgtggtc accaacagca tgcccgtcat gattctcaca gagttcatgg 2150

agaacggcgc cctggactcc ttcctgcggc taaacgacgg acagttcaca 2200

gtcatccagc tcgtgggcat gctgcggggc atcgcctcgg gcatgcggta 2250

ccttgccgag atgagctacg tccaccgaga cctggctgct cgcaacatcc 2300

tagtcaacag caacctcgtc tgcaaagtgt ctgactttgg cctttcccga 2350

ttcctggagg agaactcttc cgatcccacc tacacgagct ccctgggagg 2400

aaagattccc atccgatgga ctgccccgga ggccattgcc ttccggaagt 2450

tcacttccgc cagtgatgcc tggagttacg ggattgtgat gtgggaggtg 2500

atgtcatttg gggagaggcc gtactgggac atgagcaatc aggacgtgat 2550

caatgccatt gaacaggact accggctgcc cccgcccccа gactgtccca 2600

cctccctcca ccagctcatg ctggactgtt ggcagaaaga ccggaatgcc 2650

cggccccgct tcccccaggt ggtcagcgcc ctggacaaga tgatccggaa 2700

ccccgccagc ctcaaaatcg tggcccggga gaatggcggg gcctcacacc 2750

ctctcctgga ccagcggcag cctcactact cagcttttgg ctctgtgggc 2800

gagtggcttc gggccatcaa aatgggaaga tacgaagaaa gtttcgcagc 2850

cgctggcttt ggctccttcg agctggtcag ccagatctct gctgaggacc 2900

tgctccgaat cggagtcact ctggcgggac accagaagaa aatcttggcc 2950

agtgtccagc acatgaagtc ccaggccaag ccgggaaccc cgggtgggac 3000

aggaggaccg gccccgcagt actgacctgc aggaactccc caccccaggg 3050

acaccgcctc cccattttcc ggggcagagt ggggactcac agaggccccc 3100

agccctgtgc cccgctggat tgcactttga gcccgtgggg tgaggagttg 3150

gcaatttgga gagacaggat ttgggggttc tgccataata ggaggggaaa 3200

atcacccccc agccacctcg gggaactcca gaccaagggt gagggcgcct 3250

ttccctcagg actgggtgtg accagaggaa aaggaagtgc ccaacatctc 3300

ccagcctccc caggtgcccc cctcaccttg atgggtgcgt tcccgcagac 3350

caaagagagt gtgactccct tgccagctcc agagtggggg ggctgtccca 3400

gggggcaaga aggggtgtca gggcccagtg acaaaatcat tggggtttgt 3450

agtcccaact tgctgctgtc accaccaaac tcaatcattt ttttcccttg 3500

taaatgcccc tcccccagct gctgccttca tattgaaggt ttttgagttt 3550

tgtttttggt cttaattttt ctccccgttc ccttttgtt tcttcgtttt 3600

gtttttctac cgtccttgtc ataactttgt gttggaggga acctgtttca 3650

ctatggcctc ctttgcccaa gttgaaacag gggcccatca tcatgtctgt 3700

ttccagaaca gtgccttggt catcccacat ccccggaccc cgcctgggac 3750

ccccaagctg tgtcctatga aggggtgtgg ggtgaggtag tgaaaagggc 3800

ggtagttggt ggtggaaccc agaaacggac gccggtgctt ggaggggttc 3850

ttaaattata tttaaaaaag taacttttg tataaataaa agaaaatggg 3900

acgtgtccca aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3950

aaaaaaaaaa aga 3963

<210> 11
<211> 3476
<212> DNA
<213> Homo sapiens

<400> 11
gctctatgcc gcctaccttg ctctcgccgc tgctgccgga gccgaagcag 50

agaaggcagc gggtcccgtg accgtcccga gagccccgcg ctcccgacca 100

gggggcgggg gcggccccgg ggagggcggg gcaggggcgg ggggaagaaa 150

gggggttttg tgctgcgccg ggagggccgg cgccctcttc cgaatgtcct 200

gcggccccag cctctcctca cgctcgcgca gtctccgccg cagtctcagc 250

tgcagctgca ggactgagcc gtgcacccgg aggagacccc cggaggaggc 300

gacaaacttc gcagtgccgc gacccaaccc cagccctggg tagcctgcag 350

catggcccag ctgttcctgc ccctgctggc agccctggtc ctggcccagg 400

ctcctgcagc tttagcagat gttctggaag gagacagctc agaggaccgc 450

gcttttcgcg tgcgcatcgc gggcgacgcg ccactgcagg gcgtgctcgg 500

cggcgccctc accatccctt gccacgtcca ctacctgcgg ccaccgccga 550

gccgccgggc tgtgctgggc tctccgcggg tcaagtggac tttcctgtcc 600

cggggccggg aggcagaggt gctggtggcg cggggagtgc gcgtcaaggt 650

gaacgaggcc taccggttcc gcgtggcact gcctgcgtac ccagcgtcgc 700

tcaccgacgt ctccctggcg ctgagcgagc tgcgccccaa cgactcaggt 750

atctatcgct gtgaggtcca gcacggcatc gatgacagca gcgacgctgt 800

ggaggtcaag gtcaaagggg tcgtctttct ctaccgagag ggctctgccc 850

gctatgcttt ctccttttct ggggcccagg aggcctgtgc ccgcattgga 900

gcccacatcg ccaccccgga gcagctctat gccgcctacc ttgggggcta 950

tgagcaatgt gatgctggct ggctgtcgga tcagaccgtg aggtatccca 1000

tccagacccc acgagaggcc tgttacggag acatggatgg cttccccggg 1050

gtccggaact atggtgtggt ggacccggat gacctctatg atgtgtactg 1100

ttatgctgaa gacctaaatg gagaactgtt cctgggtgac cctccagaga 1150

agctgacatt ggaggaagca cgggcgtact gccaggagcg gggtgcagag 1200

attgccacca cgggccaact gtatgcagcc tgggatggtg gcctggacca 1250

ctgcagccca gggtggctag ctgatggcag tgtgcgctac cccatcgtca 1300

cacccagcca gcgctgtggt gggggcttgc ctggtgtcaa gactctcttc 1350

ctcttcccca accagactgg cttccccaat aagcacagcc gcttcaacgt 1400

ctactgcttc cgagactcgg cccagccttc tgccatccct gaggcctcca 1450

acccagcctc caacccagcc tctgatggac tagaggctat cgtcacagtg 1500

acagagaccc tggaggaact gcagctgcct caggaagcca cagagagtga 1550

atcccgtggg gccatctact ccatccccat catggaggac ggaggaggtg 1600

gaagctccac tccagaagac ccagcagagg cccctaggac gctcctagaa 1650

tttgaaacac aatccatggt accgcccacg gggttctcag aagaggaagg 1700

taaggcattg gaggaagaag agaaatatga agatgaagaa gagaaagagg 1750

aggaagaaga agaggaggag gtggaggatg aggctctgtg ggcatggccc 1800

agcgagctca gcagcccggg ccctgaggcc tctctcccca ctgagccagc 1850

agcccaggag aagtcactct cccaggcgcc agcaagggca gtcctgcagc 1900

ctggtgcatc accacttcct gatggagagt cagaagcttc caggcctcca 1950

agggtccatg gaccacctac tgagactctg cccactccca gggagaggaa 2000

cctagcatcc ccatcacctt ccactctggt tgaggcaaga gaggtggggg 2050

aggcaactgg tggtcctgag ctatctgggg tccctcgagg agagagcgag 2100

gagacaggaa gctccgaggg tgccccttcc ctgcttccag ccacacgggc 2150

ccctgagggt accagggagc tggaggcccc ctctgaagat aattctggaa 2200

gaactgcccc agcagggacc tcagtgcagg cccagccagt gctgcccact 2250

gacagcgcca gccgaggtgg agtggccgtg gtccccgcat caggtgactg 2300

tgtccccagc ccctgccaca atggtgggac atgcttggag gaggaggaag 2350

gggtccgctg cctatgtctg cctggctatg gggggggacct gtgcgatgtt 2400

ggcctccgct tctgcaaccc cggctgggac gccttccagg gcgcctgcta 2450

caagcacttt tccacacgaa ggagctggga ggaggcagag acccagtgcc 2500

ggatgtacgg cgcgcatctg gccagcatca gcacacccga ggaacaggac 2550

ttcatcaaca accggtaccg ggagtaccag tggatcggac tcaacgacag 2600

gaccatcgaa ggcgacttct tgtggtcgga tggcgtcccc ctgctctatg 2650

agaactggaa ccctgggcag cctgacagct acttcctgtc tggagagaac 2700

tgcgtggtca tggtgtggca tgatcaggga caatggagtg acgtgccctg 2750

caactaccac ctgtcctaca cctgcaagat ggggctggtg tcctgtgggc 2800

cgccaccgga gctgcccctg gctcaagtgt tcggccgccc acggctgcgc 2850

tatgaggtgg acactgtgct tcgctaccgg tgccgggaag gactggccca 2900

gcgcaatctg ccgctgatcc gatgccaaga gaacggtcgt tgggaggccc 2950

cccagatctc ctgtgtgccc agaagacctg cccgagctct gcacccagag 3000

gaggacccag aaggacgtca ggggaggcta ctgggacgct ggaaggcgct 3050

gttgatcccc ccttccagcc ccatgccagg tccctagggg gcaaggcctt 3100

gaacactgcc ggccacagca ctgccctgtc acccaaattt tccctcacac 3150

cttgcgctcc cgccaccaca ggaagtgaca acatgacgag gggtggtgct 3200

ggagtccagg tgacagttcc tgaaggggct tctgggaaat acctaggagg 3250

ctccagccca gcccaggccc tctcccccta ccctgggcac cagatcttcc 3300

atcagggccg gagtaaatcc ctaagtgcct caactgccct ctccctggca 3350

gccatcttgt cccctctatt cctctaggga gcactgtgcc cactctttct 3400

gggttttcca agggaatggg cttgcaggat ggagtgtctg taaaatcaac 3450

aggaaataaa actgtgtatg agccca 3476

<210> 12
<211> 2558
<212> DNA
<213> Homo sapiens

<400> 12
tgtggcactg cctgcgtacc caaccccagc cctgggtagc ctgcagcatg 50

gcccagctgt tcctgccccct gctggcagcc ctggtcctgg cccaggctcc 100

tgcagcttta gcagatgttc tggaaggaga cagctcagag gaccgcgctt 150

ttcgcgtgcg catcgcgggc gacgcgccac tgcagggcgt gctcggcggc 200

gccctcacca tcccttgcca cgtccactac ctgcggccac cgccgagccg 250

ccgggctgtg ctgggctctc cgcgggtcaa gtggactttc ctgtcccggg 300

gccgggaggc agaggtgctg gtggcgcggg gagtgcgcgt caaggtgaac 350

gaggcctacc ggttccgcgt ggcactgcct gcgtacccag cgtcgctcac 400

cgacgtctcc ctggcgctga gcgagctgcg ccccaacgac tcaggtatct 450

atcgctgtga ggtccagcac ggcatcgatg acagcagcga cgctgtggag 500

gtcaaggtca aaggggtcgt ctttctctac cgagagggct ctgcccgcta 550

tgctttctcc ttttctgggg cccaggaggc ctgtgcccgc attggagccc 600

acatcgccac cccggagcag ctctatgccg cctaccttgg gggctatgag 650

caatgtgatg ctggctggct gtcggatcag accgtgaggt atcccatcca 700

gaccccacga gaggcctgtt acggagacat ggatggcttc cccggggtcc 750

ggaactatgg tgtggtggac ccggatgacc tctatgatgt gtactgttat 800

gctgaagacc taaatggaga attgttcctg ggtgaccctc cagagaagct 850

gacattggag gaagcacggg cgtactgcca ggagcggggt gcagagattg 900

ccaccacggg ccaactgtat gcagcctggg atggtggcct ggaccactgc 950

agcccagggt ggctagctga tggcagtgtg cgctacccca tcgtcacacc 1000

cagccagcgc tgtggtgggg gcttgcctgg tgtcaagact ctcttcctct 1050

tccccaacca gactggcttc cccaataagc acagccgctt caacgtctac 1100

tgcttccgag actcggccca gccttctgcc atccctgagg cctccaaccc 1150

agcctccaac ccagcctctg atggactaga ggctatcgtc acagtgacag 1200

agaccctgga ggaactgcag ctgcctcagg aagccacaga gagtgaatcc 1250

cgtggggcca tctactccat ccccatcatg gaggacggag gaggtggaag 1300

ctccactcca gaagacccag cagaggcccc taggacgctc ctagaatttg 1350

aaacacaatc catggtaccg cccacggggt tctcagaaga ggaaggtaag 1400

gcattggagg aagaagagaa atatgaagat gaagaagaga aagaggagga 1450

agaagaagag gaggaggtgg aggatgaggc tctgtgggca tggcccagcg 1500

agctcagcag cccgggccct gaggcctctc tccccactga gccagcagcc 1550

caggaggagt cactctccca ggcgccagca agggcagtcc tgcagcctgg 1600

tgcatcacca cttcctgatg gagagtcaga agcttccagg cctccaaggg 1650

tccatggacc acctactgag actctgccca ctcccaggga gaggaaccta 1700

gcatccccat caccttccac tctggttgag gcaagagagg tgggggaggc 1750

aactggtggt cctgagctat ctggggtccc tcgaggagag agcgaggaga 1800

caggaagctc cgagggtgcc ccttccctgc ttccagccac acgggccccc 1850

gagggtacca gggagctgga ggcccccctct gaagataatt ctggaagaac 1900

tgccccagca gggacctcag tgcaggccca gccagtgctg cccactgaca 1950

gcgccagccg aggtggagtg gccgtggtcc ccgcatcagg taattctgcc 2000

caaggctcaa ctgccctctc tatcctactc cttttcttcc ccctgcagct 2050

ctgggtcacc tgacctgtag tcctttaacc caccatcatc ccaaactctc 2100

ctgtcctttg ccttcattct cttacccacc tctacctatg ggtctccaat 2150

ctcggatatc caccttgtgg gtatctcagc tctccgcgtc tttaccctgt 2200

gatcccagcc ccgccactga ccatctgtga cccttccctg ccattgggcc 2250

ctccacctgt ggctcacatc tcgccagccc cacagagcat cctcaggcct 2300

ctccaagggt cctcatcacc tattgcagcc ttcagggctc ggcctatttt 2350

ccactactcc cttcatccgc ctgtgtgccg tcccctttag ctgcctccta 2400

ttgatctcag ggaagcctgg gagtcccttc tcacccctca acctccggag 2450

tccaggagaa cccgtacccc cacagagcct taagcaacta cttctgtgaa 2500

gtatttttg actgtttcat ggaaaacaag ccttggaaat aaatctctat 2550

taaaccgc 2558

<210> 13
<211> 401
<212> DNA
<213> Homo sapiens

<400> 13
agccagctca ggctacacta tcccaggatc agcatggccg tccgccagtg 50

ggtaatcgcc ctggccttgg ctgccctcct tgttgtggac agggaagtgc 100

cagtggcagc aggaaagctc cctttctcaa gaatgcccat ctgtgaacac 150

atggtagagt ctccaacctg ttcccagatg tccaacctgg tctgcggcac 200

tgatgggctc acatatacga atgaatgcca gctctgcttg gcccggataa 250

aaaccaaaca ggacatccag atcatgaaag atggcaaatg ctgatcccac 300

aggagcacct caagccatga agtgtcagct ggagaacagt ggtgggcatg 350

gagaggatat gacatgaaat aaaagatcca gcccaaaaaa aaaagaaaaa 400

a 401

<210> 14
<211> 1572
<212> DNA
<213> Homo sapiens

<400> 14
cgtttccaat gcacgtacag cccgtacaca ccgtgtgctg ggacacccca 50

cagtcagccg catggctccc ctgtgcccca gcccctggct ccctctgttg 100

atcccggccc ctgctccagg cctcactgtg caactgctgc tgtcactgct 150

gcttctggtg cctgtccatc cccagaggtt gccccggatg caggaggatt 200

ccccccctggg aggaggctct tctggggaag atgacccact gggcgaggag 250

gatctgccca gtgaagagga ttcacccaga gaggaggatc cacccggaga 300

ggaggatcta cctggagagg aggatctacc tggagaggag gatctacctg 350

aagttaagcc taaatcagaa gaagagggct ccctgaagtt agaggatcta 400

cctactgttg aggctcctgg agatcctcaa gaaccccaga ataatgccca 450

cagggacaaa gaaggggatg accagagtca ttggcgctat ggaggcgacc 500

cgccctggcc ccgggtgtcc ccagcctgcg cgggccgctt ccagtccccg 550

gtggatatcc gcccccagct cgccgccttc tgcccggccc tgcgcccccc 600

ggaactcctg ggcttccagc tcccgccgct cccagaactg cgcctgcgca 650

acaatggcca cagtgtgcaa ctgaccctgc ctcctgggct agagatggct 700

ctgggtcccg ggcgggagta ccgggctctg cagctgcatc tgcactgggg 750

ggctgcaggt cgtccgggct cggagcacac tgtggaaggc caccgtttcc 800

ctgccgagat ccacgtggtt cacctcagca ccgcctttgc cagagttgac 850

gaggccttgg ggcgcccggg aggcctggcc gtgttggccg cctttctgga 900

ggagggcccg gaagaaaaca gtgcctatga gcagttgctg tctcgcttgg 950

aagaaatcgc tgaggaaggc tcagagactc aggtcccagg actggacata 1000

tctgcactcc tgccctctga cttcagccgc tacttccgat atgaggggtc 1050

tctgactaca ccgccctgtg cccagggtgt catctggact gtgtttaacc 1100

agacagtgat gctgagtgct aagcagctcc acaccctctc tgacaccctg 1150

tggggacctg gtgactctcg gctacagctg aacttccgag cgacgcagcc 1200

tttgaatggg cgagtgattg aggcctcctt ccctgctgga gtggacagca 1250

gtcctcgggc tgctgagcca gtccagctga attcctgcct ggctgctggt 1300

gacatcctag ccctggtttt tggcctcctt tttgctgtca ccagcgtcgc 1350

gttccttgtg cagatgagaa ggcagcacag aaggggaacc aaaggggtg 1400

tgagctaccg cccagcagag gtagccgaga ctggagccta gaggctggat 1450

cttggagaat gtgagaagcc agccagaggc atctgagggg gagccggtaa 1500

ctgtcctgtc ctgctcatta tgccacttcc ttttaactgc caagaaattt 1550

tttaaaataa atatttataa ta 1572

<210> 15
<211> 2573
<212> DNA
<213> Homo sapiens

<400> 15
tccacatcca ggtctttgtg ctcctcgctt gcctgttcct tttccacgca 50

ttttccagga taactgtgac tccaggcccg caatggatgc cctgcaacta 100

gcaaattcgg cttttgccgt tgatctgttc aaacaactat gtgaaaagga 150

gccactgggc aatgtcctct tctctccaat ctgtctctcc acctctctgt 200

cacttgctca agtgggtgct aaaggtgaca ctgcaaatga aattggacag 250

gttcttcatt ttgaaaatgt caaagatgta ccctttggat ttcaaacagt 300

aacatcggat gtaaacaaac ttagttcctt ttactcactg aaactaatca 350

agcggctcta cgtagacaaa tctctgaatc tttctacaga gttcatcagc 400

tctacgaaga gaccgtatgc aaaggaattg gaaactgttg acttcaaaga 450

taaattggaa gaaacgaaag gtcagatcaa caactcaatt aaggatctca 500

cagatggcca ctttgagaac attttagctg acaacagtgt gaacgaccag 550

accaaaatcc ttgtggttaa tgctgcctac tttgttggca agtggatgaa 600

gaaattttct gaatcagaaa caaaagaatg tcctttcaga gtcaacaaga 650

cagacaccaa accagtgcag atgatgaaca tggaggccac gttctgtatg 700

ggaaacattg acagtatcaa ttgtaagatc atagagcttc cttttcaaaa 750

taagcatctc agcatgttca tcctactacc caaggatgtg gaggatgagt 800

ccacaggctt ggagaagatt gaaaaacaac tcaactcaga gtcactgtca 850

cagtggacta atcccagcac catggccaat gccaaggtca aactctccat 900

tccaaaattt aaggtggaaa agatgattga tcccaaggct tgtctggaaa 950

atctagggct gaaacatatc ttcagtgaag acacatctga tttctctgga 1000

atgtcagaga ccaagggagt ggccctatca aatgttatcc acaaagtgtg 1050

cttagaaata actgaagatg gtggggattc catagaggtg ccaggagcac 1100

ggatcctgca gcacaaggat gaattgaatg ctgaccatcc ctttatttac 1150

atcatcaggc acaacaaaac tcgaaacatc attttctttg gcaaattctg 1200

ttctccttaa gtggcatagc ccatgttaag tcctccctga cttttctgtg 1250

gatgccgatt tctgtaaact ctgcatccag agattcattt tctagataca 1300

ataaattgct aatgttgctg gatcaggaag ccgccagtac ttgtcatatg 1350

tagccttcac acagatagac ctttttttt tttccaattc tatcttttgt 1400

ttccttttt cccataagac aatgacatac gcttttaatg aaaaggaatc 1450

acgttagagg aaaaatattt attcattatt tgtcaaattg tccgggttag 1500

ttggcagaaa tacagtcttc cacaaagaaa attcctataa ggaagatttg 1550

gaagctcttc ttcccagcac tatgctttcc ttctttggga tagagaatgt 1600

tccagacatt ctcgcttccc tgaaagactg aagaaagtgt agtgcatggg 1650

acccacgaaa ctgccctggc tccagtgaaa cttgggcaca tgctcaggct 1700

actataggtc cagaagtcct tatgttaagc cctggcaggc aggtgtttat 1750

taaaattctg aattttgggg attttcaaaa gataatattt tacatacact 1800

gtatgttata gaacttcatg gatcagatct ggggcagcac cctataaatc 1850

aacaccttaa tatgctgcaa caaaatgtag aatattcaga caaaatggat 1900

acataaagac taagtagccc ataaggggtc aaaatttgct gccaaatgcg 1950

tatgccacca acttacaaaa acacttcgtt cgcagagctt ttcagattgt 2000

ggaatgttgg ataaggaatt atagacctct agtagctgaa atgcaagacc 2050

ccaagaggaa gttcagatct taatataaat tcactttcat ttttgatagc 2100

tgtcccatct ggtcatttgg ttggcactag actggtggca ggggcttcta 2150

gctgacttgc acagggattc tcacaatagc cgatatcaga atttgtgttg 2200

aaggaacttg tctcttcatc taatatgata gcgggaaaag gagaggaaac 2250

tactgccttt agaaaatata agtaaagtga ttaaagtgct cacgttacct 2300

tgacacatag tttttcagtc tatgggttta gttactttag atggcaagca 2350

tgtaacttat attaatagta atttgtaaag ttggttggat aagctatccg 2400

tgttgcaggt tcatggatta cttctctata aaaaatatgt atttaccaaa 2450

aattttgtga cattccttct cccatctctt ccttgacctg cattgtaaat 2500

aggttcttct tgttctgaga ttcaatattg aatttttcct atgctattga 2550

caataaaata ttattgaact aca 2573

<210> 16
<211> 1145
<212> DNA
<213> Homo sapiens

<400> 16
gccgagccgg ggctccggaa gccggcgggg gcgccgcggc cgtgcggggc 50

gtcaatggat cgccactcca gctacatctt catctggctg cagctggagc 100

tctgcgccat ggccgtgctg ctcaccaaag gtgaaattcg atgctactgt 150

gatgctgccc actgtgtagc cactggttat atgtgtaaat ctgagctcag 200

cgcctgcttc tctagacttc ttgatcctca gaactcaaat tccccactca 250

cccatggctg cctggactct cttgcaagca cgacagacat ctgccaagcc 300

aaacaggccc gaaaccactc tggcaccacc atacccacat tggaatgctg 350

tcatgaagac atgtgcaatt acagagggct gcacgatgtt ctctctcctc 400

ccaggggtga ggcctcagga caaggaaaca ggtatcagca tgatggtagc 450

agaaacctta tcaccaaggt gcaggagctg acttcttcca aagagttgtg 500

gttccgggca gcggtcattg ccgtgcccat tgctggaggg ctgattttag 550

tgttgcttat tatgttggcc ctgaggatgc ttcgaagtga aaataagagg 600

ctgcaggatc agcggcaaca gatgctctcc cgtttgcact acagctttca 650

cggacaccat tccaaaaagg ggcaggttgc aaagttagac ttggaatgca 700

tggtgccggt cagtgggcac gagaactgct gtctgacctg tgataaaatg 750

agacaagcag acctcagcaa cgataagatc ctctcgcttg ttcactgggg 800

catgtacagt gggcacggga agctggaatt cgtatgacgg agtcttatct 850

gaactacact tactgaacag cttgaaggcc ttttgagttc tgctggacag 900

gagcacttta tctgaagaca aactcattta atcatctttg acagacaaaa 950

tgacctctgc aaacagaatc ttggatattt cttctgaagg attatttgca 1000

cagacttaaa tacagttaaa tgtgttattt gcttttaaaa ttataaaaag 1050

caaagagaag actttgtaca cactgtcacc agggttattt gcatccaagg 1100

gagctggaat tgagtaccta aataaacaaa aatgtgccct aaaaa 1145

<210> 17
<211> 5668
<212> DNA
<213> Homo sapiens

<400> 17
gtcacttagg aaaaggtgtc ctttcgggca gccgggctca gcatgaggaa 50

cagaaggaat gacactctgg acagcacccg gaccctgtac tccagcgcgt 100

ctcggagcac agacttgtct tacagtgaaa gcgacttggt gaattttatt 150

caagcaaatt ttaagaaacg agaatgtgtc ttctttacca aagattccaa 200

ggccacggag aatgtgtgca agtgtggcta tgcccagagc cagcacatgg 250

aaggcaccca gatcaaccaa agtgagaaat ggaactacaa gaaacacacc 300

aaggaatttc ctaccgacgc ctttggggat attcagtttg agacactggg 350

gaagaaaggg aagtatatac gtctgtcctg cgacacggac gcggaaatcc 400

tttacgagct gctgacccag cactggcacc tgaaaacacc caacctggtc 450

atttctgtga ccgggggcgc caagaacttc gccctgaagc cgcgcatgcg 500

caagatcttc agccggctca tctacatcgc gcagtccaaa ggtgcttgga 550

ttctcacggg aggcacccat tatggcctga cgaagtacat cggggaggtg 600

gtgagagata acaccatcag caggagttca gaggagaata ttgtggccat 650

tggcatagca gcttggggca tggtctccaa ccgggacacc ctcatcagga 700

attgcgatgc tgagggctat tttttagccc agtaccttat ggatgacttc 750

acaagggatc cactgtatat cctggacaac aaccacacac atttgctgct 800

cgtggacaat ggctgtcatg gacatcccac tgtcgaagca aagctccgga 850

atcagctaga gaagcatatc tctgagcgca ctattcaaga ttccaactat 900

ggtggcaaga tccccattgt gtgttttgcc caaggaggtg gaaaagagac 950

tttgaaagcc atcaatacct ccatcaaaaa taaaattcct tgtgtggtgg 1000

tggaaggctc gggccggatc gctgatgtga tcgctagcct ggtggaggtg 1050

gaggatgccc cgacatcttc tgccgtcaag gagaagctgg tgcgcttttt 1100

accccgcacg gtgtcccggc tgtctgagga ggagactgag agttggatca 1150

aatggctcaa agaaattctc gaatgttctc acctattaac agttattaaa 1200

atggaagaag ctggggatga aattgtgagc aatgccatct cctacgctct 1250

atacaaagcc ttcagcacca gtgagcaaga caaggataac tggaatgggc 1300

agctgaagct tctgctggag tggaaccagc tggacttagc caatgatgag 1350

attttcacca atgaccgccg atgggagtct gctgaccttc aagaagtcat 1400

gtttacggct ctcataaagg acagacccaa gtttgtccgc ctctttctgg 1450

agaatggctt gaacctacgg aagtttctca cccatgatgt cctcactgaa 1500

ctcttctcca accacttcag cacgcttgtg taccggaatc tgcagatcgc 1550

caagaattcc tataatgatg ccctcctcac gtttgtctgg aaactggttg 1600

cgaacttccg aagaggcttc cggaaggaag acagaaatgg ccgggacgag 1650

atggacatag aactccacga cgtgtctcct attactcggc acccccctgca 1700

agctctcttc atctgggcca ttcttcagaa taagaaggaa ctctccaaag 1750

tcatttggga gcagaccagg ggctgcactc tggcagccct gggagccagc 1800

aagcttctga agactctggc caaagtgaag aacgacatca atgctgctgg 1850

ggagtccgag gagctggcta atgagtacga gacccgggct gttgagctgt 1900

tcactgagtg ttacagcagc gatgaagact tggcagaaca gctgctggtc 1950

tattcctgtg aagcttgggg tggaagcaac tgtctggagc tggcggtgga 2000

ggccacagac cagcatttca ccgcccagcc tggggtccag aatttttttt 2050

ctaagcaatg gtatggagag atttcccgag acaccaagaa ctggaagatt 2100

atcctgtgtc tgtttattat acccttggtg ggctgtggct ttgtatcatt 2150

taggaagaaa cctgtcgaca agcacaagaa gctgctttgg tactatgtgg 2200

cgttcttcac ctccccctte gtggtcttct cctggaatgt ggtcttctac 2250

atcgccttcc tcctgctgtt tgcctacgtg ctgctcatgg atttccattc 2300

ggtgccacac ccccccgagc tggtcctgta ctcgctggtc tttgtcctct 2350

tctgtgatga agtgagacag tggtacgtaa atggggtgaa ttattttact 2400

gacctgtgga atgtgatgga cacgctgggg ctttttttact tcatagcagg 2450

aattgtattt cggctccact cttctaataa aagctctttg tattctggac 2500

gagtcatttt ctgtctggac tacattattt tcactctaag attgatccac 2550

atttttactg taagcagaaa cttaggaccc aagattataa tgctgcagag 2600

gatgctgatc gatgtgttct tcttcctgtt cctctttgcg gtgtggatgg 2650

tggcctttgg cgtggccagg caagggatcc ttaggcagaa tgagcagcgc 2700

tggaggtgga tattccgttc ggtcatctac gagccctacc tggccatgtt 2750

cggccaggtg cccagtgacg tggatggtac cacgtatgac tttgcccact 2800

gcaccttcac tgggaatgag tccaagccac tgtgtgtgga gctggatgag 2850

cacaacctgc cccggttccc cgagtggatc accatccccc tggtgtgcat 2900

ctacatgtta tccaccaaca tcctgctggt caacctgctg gtcgccatgt 2950

ttggctacac ggtgggcacc gtccaggaga acaatgacca ggtctggaag 3000

ttccagaggt acttcctggt gcaggagtac tgcagccgcc tcaatatccc 3050

cttccccttc atcgtcttcg cttacttcta catggtggtg aagaagtgct 3100

tcaagtgttg ctgcaaggag aaaaacatgg agtcttctgt ctgctgtttc 3150

aaaaatgaag acaatgagac tctggcatgg gagggtgtca tgaaggaaaa 3200

ctaccttgtc aagatcaaca caaaagccaa cgacacctca gaggaaatga 3250

ggcatcgatt tagacaactg gatacaaagc ttaatgatct caagggtctt 3300

ctgaaagaga ttgctaataa aatcaaataa aactgtatga aactctaatg 3350

gagaaaaatc taattatagc aagatcatat taaggaatgc tgatgaacaa 3400

ttttgctatc gactactaaa tgagagattt tcagaccgct gggtacatgg 3450

tggatgattt taaatcaccc tagtgtgctg agaccttgag aataaagtgt 3500

gtgattggtt tcatacttga agacggatat aaaggaagaa tatttccttt 3550

atgtgtttct ccagaatggt gcctgtttct ctctgtgtct caatgcctgg 3600

gactggaggt tgatagttta agtgtgttct taccgcctcc tttttccttt 3650

aatcttattt ttgatgaaca catatatagg agaacatcta tcctatgaat 3700

aagaacctgg tcatgcttta ctcctgtatt gttattttgt tcatttccaa 3750

ttgattctct acttttccct tttttgtatt atgtgactaa ttagttggca 3800

tattgttaaa agtctctcaa attaggccag attctaaaac atgctgcagc 3850

aagaggaccc cgctctcttc aggaaaagtg ttttcatttc tcaggatgct 3900

tcttacctgt cagaggaggt gacaaggcag tctcttgctc tcttggactc 3950

accaggctcc tattgaagga accacccccа ttcctaaata tgtgaaaagt 4000

cgcccaaaat gcaaccttga aaggcactac tgactttgtt cttattggat 4050

actcctctta tttattattt ttccattaaa aataatagct ggctattata 4100

gaaaatttag accatacaga gatgtagaaa gaacataaat tgtccccatt 4150

accttaaggt aatcactgct aacaatttct ggatggtttt tcaagtctat 4200

ttttttttcta tgtatgtctc aattctcttt caaaatttta cagaatgtta 4250

tcatactaca tatatacttt ttatgtaagc tttttcactt agtattttat 4300

caaatatgtt tttattatat tcatagcctt cttaaacatt atatcaataa 4350

ttgcataata ggcaacctct agcgattacc ataattttgc tcattgaagg 4400

ctatctccag ttgatcattg ggatgagcat ctttgtgcat gaatcctatt 4450

gctgtatttg ggaaaatttt ccaaggttag attccaataa atatctattt 4500

attattaaat attaaaatat cgatttatta ttaaaaccat ttataaggct 4550

ttttcataaa tgtatagcaa ataggaatta ttaacttgag cataagatat 4600

gagatacatg aacctgaact attaaaataa aatattatat ttaaccctag 4650

tttaagaaga agtcaatatg cttatttaaa tattatggat ggtgggcaga 4700

tcacttgagg tcaggagttc gagaccagcc tggccaacat ggcaaaacca 4750

catctctact aaaaataaaa aaattagctg ggtgtggtgg tgcactcctg 4800

taatcccagc tactcagaag gctgaggtac aagaattgct ggaacctggg 4850

aggcggaggt tgcagtgaac caagattgca ccactgcact ccagccgggg 4900

tgacagagtg agactccgac tgaaaataaa taaataaata aataaataaa 4950

taaataaata aatattatgg atggtgaagg gaatggtata gaattggaga 5000

gattatctta ctgaacacct gtagtcccag ctttctctgg aagtggtggt 5050

atttgagcag gatgtgcaca aggcaattga aatgcccata attagtttct 5100

cagctttgaa tacactataa actcagtggc tgaaggagga aattttagaa 5150

ggaagctact aaaagatcta atttgaaaaa ctacaaaagc attaactaaa 5200

aaagtttatt ttccttttgt ctgggcagta gtgaaaataa ctactcacaa 5250

cattcactat gtttgcaagg aattaacaca aataaaagat gccttttac 5300

ttaaacgcca agacagaaaa cttgcccaat actgagaagc aacttgcatt 5350

agagagggaa ctgttaaatg ttttcaaccc agttcatctg gtggatgttt 5400

ttgcaggtta ctctgagaat tttgcttatg aaaaatcatt atttttagtg 5450

tagttcacaa taatgtattg aacatacttc taatcaaagg tgctatgtcc 5500

ttgtgtatgg tactaaatgt gtcctgtgta cttttgcaca actgagaatc 5550

ctgcggcttg gtttaatgag tgtgttcatg aaataaataa tggaggaatt 5600

gtcaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 5650

aaaaaaaaaa aaaaaaaa 5668

<210> 18
<211> 3740
<212> DNA
<213> Homo sapiens

<400> 18
gaattcgccc ttggctcagc atgaggaaca gaaggaatga cactctggac 50

agcacccgga ccctgtactc cagcgcgtct cggagcacag acttgtctta 100

cagtgaaagc gacttggtga attttattca agcaaatttt aagaaacgag 150

aatgtgtctt ctttaccaaa gattccaagg ccacggagaa tgtgtgcaag 200

tgtggctatg cccagagcca gcacatggaa ggcacccaga tcaaccaaag 250

tgagaaatgg aactacaaga aacacaccaa ggaatttcct accgacgcct 300

ttggggatat tcagtttgag acactgggga agaaagggaa gtatatacgt 350

ctgtcctgcg acacggacgc ggaaatcctt tacgagctgc tgacccagca 400

ctggcacctg aaaacaccca acctggtcat ttctgtgacc gggggcgcca 450

agaacttcgc cctgaagccg cgcatgcgca agatcttcag ccggctcatc 500

tacatcgcgc agtccaaagg tgcttggatt ctcacgggag gcacccatta 550

tggcctgatg aagtacatcg gggaggtggt gagagataac accatcagca 600

ggagttcaga ggagaatatt gtggccattg gcatagcagc ttggggcatg 650

gtctccaacc gggacaccct catcaggaat tgcgatgctg agggctattt 700

tttagcccag taccttatgg atgacttcac aagagatcca ctgtatatcc 750

tggacaacaa ccacacacat ttgctgctcg tggacaatgg ctgtcatgga 800

catcccactg tcgaagcaaa gctccggaat cagctagaga agtatatctc 850

tgagcgcact attcaagatt ccaactatgg tggcaagatc cccattgtgt 900

gttttgccca aggaggtgga aaagagactt tgaaagccat caatacctcc 950

atcaaaaata aaattccttg tgtggtggtg gaaggctcgg gccagatcgc 1000

tgatgtgatc gctagcctgg tggaggtgga ggatgccctg acatcttctg 1050

ccgtcaagga gaagctggtg cgctttttac cccgcacggt gtcccggctg 1100

cctgaggagg agactgagag ttggatcaaa tggctcaaag aaattctcga 1150

atgttctcac ctattaacag ttattaaaat ggaagaagct ggggatgaaa 1200

ttgtgagcaa tgccatctcc tacgctctat acaaagcctt cagcaccagt 1250

gagcaagaca aggataactg gaatgggcag ctgaagcttc tgctggagtg 1300

gaaccagctg gacttagcca atgatgagat tttcaccaat gaccgccgat 1350

gggagtctgc tgaccttcaa gaagtcatgt ttacggctct cataaaggac 1400

agacccaagt ttgtccgcct ctttctggag aatggcttga acctacggaa 1450

gtttctcacc catgatgtcc tcactgaact cttctccaac cacttcagca 1500

cgcttgtgta ccggaatctg cagatcgcca agaattccta taatgatgcc 1550

ctcctcacgt ttgtctggaa actggttgcg aacttccgaa gaggcttccg 1600

gaaggaagac agaaatggcc gggacgagat ggacatagaa ctccacgacg 1650

tgtctcctat tactcggcac cccctgcaag ctctcttcat ctgggccatt 1700

cttcagaata agaaggaact ctccaaagtc atttgggagc agaccagggg 1750

ctgcactctg gcagccctgg gagccagcaa gcttctgaag actctggcca 1800

aagtgaagaa cgacatcaat gctgctgggg agtccgagga gctggctaat 1850

gagtacgaga cccgggctgt tgagctgttc actgagtgtt acagcagcga 1900

tgaagacttg gcagaacagc tgctggtcta ttcctgtgaa gcttggggtg 1950

gaagcaactg tctggagctg gcggtggagg ccacagacca gcatttcatc 2000

gcccagcctg gggtccagaa ttttctttct aagcaatggt atggagagat 2050

ttcccgagac accaagaact ggaagattat cctgtgtctg tttattatac 2100

ccttggtggg ctgtggcttt gtatcattta ggaagaaacc tgtcgacaag 2150

cacaagaagc tgctttggta ctatgtggcg ttcttcacct ccccttcgt 2200

ggtcttctcc tggaatgtgg tcttctacat cgccttcctc ctgctgtttg 2250

cctacgtgct gctcatggat ttccattcgg tgccacaccc ccccgagctg 2300

gtcctgtact cgctggtctt tgtcctcttc tgtgatgaag tgagacagtg 2350

gtacgtaaat ggggtgaatt attttactga cctgtggaat gtgatggaca 2400

cgctggggct tttttacttc atagcaggaa ttgtatttcg gctccactct 2450

tctaataaaa gctctttgta ttctggacga gtcattttct gtctggacta 2500

cattattttc actctaagat tgatccacat ttttactgta agcagaaact 2550

taggacccaa gattataatg ctgcagagga tgctgatcga tgtgttcttc 2600

ttcctgttcc tctttgcggt gtggatggtg gcctttggcg tggccaggca 2650

agggatcctt aggcagaatg agcagcgctg gaggtggata ttccgttcgg 2700

tcatctacga gccctacctg gccatgttcg gccaggtgcc cagtgacgtg 2750

gatggtacca cgtatgactt tgcccactgc accttcactg ggaatgagtc 2800

caagccactg tgtgtggagc tggatgagca caacctgccc cggttccccg 2850

agtggatcac catccccctg gtgtgcatct acatgttatc caccaacatc 2900

ctgctggtca acctgctggt cgccatgttt ggctacacgg tgggcaccgt 2950

ccaggagaac aatgaccagg tctggaagtt ccagaggtac ttcctggtgc 3000

aggagtactg cagccgcctc aatatccccc tccccttcat tgtcttcgct 3050

tacttctaca tggtggtgaa gaagtgcttc aagtgttgct gcaaggagaa 3100

aaacatggag tcttctgtct gctgtttcaa aaatgaagac aatgagactc 3150

tggcatggga gggtgtcatg aaggaaaact accttgtcaa gatcaacaca 3200

aaagccaacg acacctcaga ggaaatgagg catcgattta gacaactgga 3250

tacaaagctt aatgatctca agggtcttct gaaagagatt gctaataaaa 3300

tcaaataaaa ctgtatgaac tctaatggag aaaaatctaa ttatagcaag 3350

atcatattaa ggaatgctga tgaacaattt tgctatcgac tactaaatga 3400

gagattttca gacccctggg tacatggtgg atgattttaa atcaccctag 3450

tgtgctgaga ccttgagaat aaagtgtgtg attggtttca tacttgaaga 3500

cggatataaa ggaagaatat ttcctttatg tgtttctcca gaatggtgcc 3550

tgtttctctc tgtgtctcaa tgcctgggac tggaggttga tagtttaagt 3600

gtgttcttac cgcctccttt ttcctttaat cttatttttg atgaacacat 3650

atataggaga acatctatcc tatgaataag aacctggtca tgctttaaaa 3700

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3740

<210> 19
<211> 3809
<212> DNA
<213> Homo sapiens

<400> 19
gcggcggggg cgaggtgagg tgttggcagt ggaaaggggt tcgggctcgg 50

ggggcggggg gacgcggagc ggatggcccg cgccggccgc aggggcggat 100

aaaaagccgt cgcgctgcgg gagtgggcgg gagggagagg gggtgtccga 150

gggccacaag agtatgacgg ggctgtacga gctggtgtgg cgggtgctgc 200

acgcgctgct ctgtctgcac cgcacgctca cctcctggct ccgcgttcgg 250

ttcggcacct ggaactggat ctggcggcgc tgctgccgcg ccgcctctgc 300

cgcggtccta gcgccgctcg gcttcacgct ccgcaagccc ccggcagtcg 350

gcaggaaccg ccgtcaccac cggcacccgc gcgggggtc gtgcctggca 400

gccgcacacc accggatgcg ctggcgcgcg gacggtcgtt ccttggagaa 450

gctgcctgtg catatgggcc tggtgatcac cgaggtggag caggaaccca 500

gcttctcgga catcgcgagc ctcgtggtgt ggtgtatggc cgtgggcatc 550

tcctacatta gcgtctacga ccaccaaggt attttcaaaa gaaataattc 600

cagattgatg gatgaaattt taaaacaaca gcaagaactt ctgggcctag 650

attgttcaaa atactcacca gaatttgcaa atagtaatga caaagatgat 700

caagttttaa attgccattt ggcagtgaag gtgctgtctc cggaagatgg 750

aaaagcagat attgtaagag ctgctcagga cttttgccag ttagtagccc 800

agaagcaaag gagacccaca gatttggatg tagatacgtt agccagttta 850

cttagttcaa atggttgtcc tgatcctgat ttagtattga agttcggtcc 900

tgtggacagc acattaggct ttcttccctg gcacatcaga ttgactgaga 950

ttgtctcttt gccttcccac ctaaacatca gttatgagga ctttttctct 1000

gcccttcgtc aatatgcagc ctgtgaacag cgtctgggaa agtagtggtc 1050

attggttgca taatttgatt tgaggcttgt ggaggaaagg aaccaagtga 1100

ctctgatgtt tacaaagcac ctatgaaacc ctgtacacac ctatgaaacc 1150

ctgtacacac ctagttcata atcctcataa tttatcaaca aacacaaaga 1200

agtgtcttac ttgagagtga gtgtgtgcgt gtgtgcgtgc acacatgtgc 1250

acgtttgtat gtgtggaaat aaacataaat ggggacgtgt tggagaagga 1300

aatacataga cctacaactt tgagcatata gcagtgatgt tttaggaact 1350

gaaatgtcac acttaataaa gtcttcagcc cagctacttc cctgttttcg 1400

tggggagaag ggggcctgat tagaactgtt ctggttgtgt ttggcgggag 1450

gggaataaat tttgttcagt ccttcttagt gaccaaactt taatttttag 1500

gaataacata ttgacttact gaactgaagc attctgagtt gaaaggagcc 1550

ccagaggaaa ggagttctgt gttgctcaca tgttaaaagc ttgctcacct 1600

tcagagcaga gggaatacct atcttcagat atccgcccat tttcatctct 1650

tcattatagt caaacagtgt gacttgagag tgttgctctg gtgtctgtat 1700

tctggcttat gaagattatt tgaaaaagaa ctcttactac attgaaatgc 1750

agacttttaa aaatttaaat attggattag gcagtcaaaa aaccaaacaa 1800

gcataaaagg tcagtaagtt gtaatcttaa aagtaaaggt ggaaaactca 1850

ttataaatgg aagaaaagtt ttgatttcct tttttgtttg atgggcagta 1900

tgccatatta tatccaaagt tggtttaaaa aatacttcca tcaactattt 1950

ttatttaaaa taaacatttg agggaagtta ccaaggcagc ttttttcctc 2000

aaaagtaacc tgttcctctt tggaacagca cattttaggg gcatggttaa 2050

tacctgagat ttttactcag taaacctgat ggttaccgtg tgtaaaatat 2100

ctttaagtag gattgaaggc ctctgtgggg aatataatgt taccaaagcc 2150

tataaaaata aattttacat gttctttttt atgacagaga gcagcactgg 2200

ttctgttatt ttttaaaatg aataactgat ttcttgacag gtgtttagta 2250

tttcttccct cactgctgat tcttggatag aaaccattct ttatatttga 2300

tagactgttt tcagaaaact cttatcaaca agtgtacaat agttatctaa 2350

aactgtacat ttagaatgga gcagtttaat actagatctc agaaatttga 2400

aaaatagcaa agaagactgg atttggaaag catggtctac aattgcttgt 2450

caattctgaa gctatgaaga ataaatgttt cagctttgga ttatgaaacc 2500

ccgtttatga tttttaaaac acacttgaaa taaaaatgat tgaactaaat 2550

1099
1097.

tttggtccgg tgacatcatt ttgcactgca tagcccatta tacattgtaa 2600

gtctttttgt tttaatttat tactgagagt tttctgtgaa gctacagcat 2650

atctaatcag agaatttctg atttgtttcc ttatactgtg attatgttat 2700

attaaggcat ttgtagtgca gcttaagact gaatttatgc tttttgtaaa 2750

catgatcggt gtaaatgtct tataaacatt ctggagtatg tataacttta 2800

atgaatgaaa tttaatggac ctgattaaaa taaagggatt taatctttgt 2850

taaagtgaag ttagtggaat aaattaccta ctggaatata gtccaagtca 2900

ctaaaggttt aatatttgca tttttgtgct tttatttttct ccttccattc 2950

ataattatat acttgaaagt acatctgtag cctatgactt gagtctcttg 3000

aacttctagg aagaggcaaa ctacaaacta ctaggattct gatttcagat 3050

ataggcattc cagaatcttc tctttacgag ttcacctgct agtataatct 3100

ccacaacttg aatggccttg gttgttctgt aattgctgcc aaaatcatca 3150

caagctgtac gtcatcaagg ctccctttgc actcccaaga agaactgttc 3200

attttaaaca aaagtgtatg tctttatttg tattggaaaa tattgtcttt 3250

aaattgtttc ttgttgacac ttcccacaat ggaaaaatta ccaaattaaa 3300

cctgttttat ggatggcagc ttggagcata gcaagaattt ggaggatttg 3350

aattccattc ccagttctca ttgcgttttg tttcttaaaa ctataataat 3400

tggttactgt tataaagttt aaaaggtggt tttaatatga atagcaaatt 3450

ctggtatatt gtgactaacg cttaagaatg cctgtctttg agaggaaggt 3500

gttataatat taatcaacag tgccaaatac actgtgcata tctacaattt 3550

aatctttgaa tgtttgttac tggattagct ccctcctcct tctgtgtgat 3600

ggtaccatgc atagagtcaa tcaaatcctt gtgatgtttt gtatggactt 3650

tgacaacatg taactaatgt gtaaagcaag tttttatgat taaggaatca 3700

aatttattga attttattat tgaaagttga aacttaacat gtataaacaa 3750

aaaacaataa aataataaac tattttcatt gactataaaa aaaagcccac 3800

gcgtccgcc 3809

<210> 20
<211> 1048
<212> DNA
<213> Homo sapiens

<400> 20
gaggtgaggt gttggcagtg gaaaggggtt cgggctcggg gggcgggggg 50

acgcggagcg atggcccgcg ccggccgcag gggcggataa aaagccgtcg 100

cgctgcggga gtgggcggga gggagagggg gtgtccgagg gccacaagag 150

tatgacgggg ctgtacgagc tggtgtggcg ggtgctgcac gcgctgctct 200

gtctgcaccg cacgctcacc tcctggctcc gcgttcggtt cggcacctgg 250

aactggatct ggcggcgctg ctgccgcgcc gcctctgccg cggtcctagc 300

gccgctcggc ttcacgctcc gcaagccccc ggcagtcggc aggaaccgcc 350

gtcaccaccg gcacccgcgc ggggggtcgt gcctggcagc cgcacaccac 400

cggatgcgct ggcgcgcgga cggtcgttcc ttggagaagc tgcctgtgca 450

tatgggcctg gtgatcaccg aggtggagca ggaacccagc ttctcggaca 500

tcgcgagcct cgtggtgtgg tgtatggccg tgggcatctc ctacattagc 550

gtctacgacc accaaggtat tttcaaaaga aataattcca gattgatgga 600

tgaaatttta aaacaacagc aagaacttct gggcctagat tgttcaaaat 650

actcaccaga atttgcaaat agtaatgaca aagatgatca agttttaaat 700

tgccatttgg cagtgaaggt gctgtctccg gaagatggaa aagcagatat 750

tgtaagagct gctcaggact tttgccagtt agtagcccag aagcaaaaga 800

gacccacaga tttggatgta gatacgttag ccagtttact tagttcaaat 850

ggttgtcctg atcctgattt agtattgaag ttcggtcctg tggacagcac 900

attaggcttt cttccctggc acatcagatt gactgagatt gtctctttgc 950

cttcccacct aaacatcagt tatgaggact ttttctctgc ccttcgtcaa 1000

tatgcagcct gtgaacagcg tctgggaaag tagtggtcat tggttgca 1048

<210> 21
<211> 3332
<212> DNA
<213> Homo sapiens

<400> 21
gggggggcgga accagcctgc acgcgctggc tccgggtgac agccgcgcgc 50

ctcggccagg atctgagtga tgagacgtgt ccccactgag gtgccccaca 100

gcagcaggtg ttgagcatgg gctgagaagc tggaccggca ccaaagggct 150

ggcagaaatg ggcgcctggc tgattcctag gcagttggcg gcagcaagga 200

ggagaggccg cagcttctgg agcagagccg agacgaagca gttctggagt 250

gcctgaacgg ccccctgagc cctacccgcc tggcccacta tggtccagag 300

gctgtgggtg agccgcctgc tgcggcaccg gaaagcccag ctcttgctgg 350

tcaacctgct aacctttggc ctggaggtgt gtttggccgc aggcatcacc 400

tatgtgccgc ctctgctgct ggaagtgggg gtagaggaga agttcatgac 450

catggtgctg ggcattggtc cagtgctggg cctggtctgt gtcccgctcc 500

taggctcagc cagtgaccac tggcgtggac gctatggccg ccgccggccc 550

ttcatctggg cactgtcctt gggcatcctg ctgagcctct ttctcatccc 600

aagggccggc tggctagcag ggctgctgtg cccggatccc aggcccctgg 650

agctggcact gctcatcctg ggcgtggggc tgctggactt ctgtggccag 700

gtgtgcttca ctccactgga ggccctgctc tctgacctct tccgggaccc 750

ggaccactgt cgccaggcct actctgtcta tgccttcatg atcagtcttg 800

ggggctgcct gggctacctc ctgcctgcca ttgactggga caccagtgcc 850

ctggcccccc acctgggcac ccaggaggag tgcctctttg gcctgctcac 900

cctcatcttc ctcacctgcg tagcagccac actgctggtg gctgaggagg 950

cagcgctggg ccccaccgag ccagcagaag ggctgtcggc cccctccttg 1000

tcgccccact gctgtccatg ccgggcccgc ttggctttcc ggaacctggg 1050

cgccctgctt ccccggctgc accagctgtg ctgccgcatg ccccgcaccc 1100

tgcgccggct cttcgtggct gagctgtgca gctggatggc actcatgacc 1150

ttcacgctgt tttacacgga tttcgtgggc gaggggctgt accagggcgt 1200

gcccagagct gagccgggca ccgaggcccg gagacactat gatgaaggcg 1250

ttcggatggg cagcctgggg ctgttcctgc agtgcgccat ctccctggtc 1300

ttctctctgg tcatggaccg gctggtgcag cgattcggca ctcgagcagt 1350

ctatttggcc agtgtggcag ctttccctgt ggctgccggt gccacatgcc 1400

tgtcccacag tgtggccgtg gtgacagctt cagccgccct caccgggttc 1450

accttctcag ccctgcagat cctgccctac acactggcct ccctctacca 1500

ccgggagaag caggtgttcc tgcccaaata ccgaggggac actggaggtg 1550

ctagcagtga ggacagcctg atgaccagct tcctgccagg ccctaagcct 1600

ggagctccct tccctaatgg acacgtgggt gctggaggca gtggcctgct 1650

cccacctcca cccgcgctct gcggggcctc tgcctgtgat gtctccgtac 1700

gtgtggtggt gggtgagccc accgaggcca gggtggttcc gggccggggc 1750

atctgcctgg acctcgccat cctggatagt gccttcctgc tgtcccaggt 1800

ggccccatcc ctgtttatgg gctccattgt ccagctcagc cagtctgtca 1850

ctgcctatat ggtgtctgcc gcaggcctgg gtctggtcgc catttacttt 1900

gctacacagg tagtatttga caagagcgac ttggccaaat actcagcgta 1950

gaaaacttcc agcacattgg ggtggagggc ctgcctcact gggtcccagc 2000

tccccgctcc tgttagcccc atggggctgc cgggctggcc gccagtttct 2050

gttgctgcca aagtaatgtg gctctctgct gccaccctgt gctgctgagg 2100

tgcgtagctg cacagctggg ggctggggcg tccctctcct ctctccccag 2150

tctctagggc tgcctgactg gaggccttcc aaggggttt cagtctggac 2200

ttatacaggg aggccagaag ggctccatgc actggaatgt ggggactctg 2250

caggtggatt acccaggctc agggttaaca gctagcctcc tagttgagac 2300

acacctagag aagggttttt gggagctgaa taaactcagt cacctggttt 2350

cccatctcta agccccttaa cctgcagctt cgtttaatgt agctcttgca 2400

tgggagtttc taggatgaaa cactcctcca tgggatttga acatatgaaa 2450

gttatttgta ggggaagagt cctgaggggc aacacacaag aaccaggtcc 2500

cctcagccca cagcactgtc tttttgctga tccacccccc tcttaccttt 2550

tatcaggatg tggcctgttg gtccttctgt tgccatcaca gagacacagg 2600

catttaaata tttaacttat ttatttaaca aagtagaagg gaatccattg 2650

ctagcttttc tgtgttggtg tctaatattt gggtagggtg ggggatcccc 2700

aacaatcagg tcccctgaga tagctggtca ttgggctgat cattgccaga 2750

atcttcttct cctggggtct ggcccccccaa aatgcctaac ccaggacctt 2800

ggaaattcta ctcatcccaa atgataattc caaatgctgt tacccaaggt 2850

tagggtgttg aaggaaggta gagggtgggg cttcaggtct caacggcttc 2900

cctaaccacc cctcttctct tggcccagcc tggttccccc cacttccact 2950

cccctctact ctctctagga ctgggctgat gaaggcactg cccaaaattt 3000

cccctacccc caactttccc ctacccccaa ctttccccac cagctccaca 3050

accctgtttg gagctactgc aggaccagaa gcacaaagtg cggtttccca 3100

agcctttgtc catctcagcc cccagagtat atctgtgctt ggggaatctc 3150

acacagaaac tcaggagcac cccctgcctg agctaaggga ggtcttatct 3200

ctcaggggg tttaagtgcc gtttgcaata atgtcgtctt atttatttag 3250

cggggtgaat attttatact gtaagtgagc aatcagagta taatgtttat 3300

ggtgacaaaa ttaaagcctt tcttatatgt tt 3332

<210> 22
<211> 3768
<212> DNA
<213> Homo sapiens

<400> 22
aacaaaagct ggagctccac cgcggtggcg gccgctctag cccctggttc 50

ggcccacctc tgaaggttcc agaatcgata gtgaattcgt ggggaagcgc 100

agccatggct ctgcggaggc tgggggccgc gctgctgctg ctgccgctgc 150

tcgccgccgt ggaagaaacg ctaatggact ccactacagc gactgctgag 200

ctgggctgga tggtgcatcc tccatcaggg tgggaagagg tgagtggcta 250

cgatgagaac atgaacacga tccgcacgta ccaggtgtgc aacgtgtttg 300

agtcaagcca gaacaactgg ctacggacca agtttatccg gcgccgtggc 350

gcccaccgca tccacgtgga gatgaagttt tcggtgcgtg actgcagcag 400

catccccagc gtgcctggct cctgcaagga gaccttcaac ctctattact 450

atgaggctga ctttgactcg gccaccaaga ccttccccaa ctggatggag 500

aatccatggg tgaaggtgga taccattgca gccgacgaga gcttctccca 550

ggtggacctg ggtggccgcg tcatgaaaat caacaccgag gtgcggagct 600

tcggacctgt gtcccgcagc ggcttctacc tggccttcca ggactatggc 650

ggctgcatgt ccctcatcgc cgtgcgtgtc ttctaccgca agtgcccccg 700

catcatccag aatggcgcca tcttccagga aaccctgtcg ggggctgaga 750

gcacatcgct ggtggctgcc cggggcagct gcatcgccaa tgcggaagag 800

gtggatgtac ccatcaagct ctactgtaac ggggacggcg agtggctggt 850

gcccatcggg cgctgcatgt gcaaagcagg cttcgaggcc gttgagaatg 900

gcaccgtctg ccgaggttgt ccatctggga ctttcaaggc caaccaaggg 950

gatgaggcct gtacccactg tcccatcaac agccggacca cttctgaagg 1000

ggccaccaac tgtgtctgcc gcaatggcta ctacagagca gacctggacc 1050

ccctggacat gccctgcaca accatccccc ccgcgcccca ggctgtgatt 1100

tccagtgtca atgagacctc cctcatgctg gagtggaccc ctccccgcga 1150

ctccggaggc cgagaggacc tcgtctacaa catcatctgc aagagctgtg 1200

gctcgggccg gggtgcctgc acccgctgcg gggacaatgt acagtacgca 1250

ccacgccagc taggcctgac cgagccacgc atttacatca gtgacctgct 1300

ggcccacacc cagtacacct tcgagatcca ggctgtgaac ggcgttactg 1350

accagagccc cttctcgcct cagttcgcct ctgtgaacat caccaccaac 1400

caggcagctc catcggcagt gtccatcatg catcaggtga gccgcaccgt 1450

ggacagcatt accctgtcgt ggtcccagcc agaccagccc aatggcgtga 1500

tcctggacta tgagctgcag tactatgaga aggagctcag tgagtacaac 1550

gccacagcca taaaaagccc caccaacacg gtcaccgtgc agggcctcaa 1600

agccggcgcc atctatgtct tccaggtgcg ggcacgcacc gtggcaggct 1650

acgggcgcta cagcggcaag atgtacttcc agaccatgac agaagccgag 1700

taccagacaa gcatccagga gaagttgcca ctcatcatcg gctcctcggc 1750

cgctggcctg gtcttcctca ttgctgtggt tgtcatcgcc atcgtgtgta 1800

acagaagacg ggggtttgag cgtgctgact cggagtacac ggacaagctg 1850

caacactaca ccagtggcca catgacccca ggcatgaaga tctacatcga 1900

tcctttcacc tacgaggacc ccaacgaggc agtgcgggag tttgccaagg 1950

aaattgacat ctcctgtgtc aaaattgagc aggtgatcgg agcaggggag 2000

tttggcgagg tctgcagtgg ccacctgaag ctgccaggca agagagagat 2050

ctttgtggcc atcaagacgc tcaagtcggg ctacacggag aagcagcgcc 2100

gggacttcct gagcgaagcc tccatcatgg gccagttcga ccatcccaac 2150

gtcatccacc tggagggtgt cgtgaccaag agcacacctg tgatgatcat 2200

caccgagttc atggagaatg gctccctgga ctcctttctc cggcaaaacg 2250

atgggcagtt cacagtcatc cagctggtgg gcatgcttcg gggcatcgca 2300

gctggcatga agtacctggc agacatgaac tatgttcacc gtgacctggc 2350

tgcccgcaac atcctcgtca acagcaacct ggtctgcaag gtgtcggact 2400

ttgggctctc acgctttcta gaggacgata cctcagaccc cacctacacc 2450

agtgccctgg gcggaaagat ccccatccgc tggacagccc cggaagccat 2500

ccagtaccgg aagttcacct cggccagtga tgtgtggagc tacggcattg 2550

tcatgtggga ggtgatgtcc tatggggagc ggccctactg ggacatgacc 2600

aaccaggatg taatcaatgc cattgagcag gactatcggc tgccaccgcc 2650

catggactgc ccgagcgccc tgcaccaact catgctggac tgttggcaga 2700

aggaccgcaa ccaccggccc aagttcggcc aaattgtcaa cacgctagac 2750

aagatgatcc gcaatcccaa cagcctcaaa gccatggcgc ccctctcctc 2800

tggcatcaac ctgccgctgc tggaccgcac gatccccgac tacaccagct 2850

ttaacacggt ggacgagtgg ctggaggcca tcaagatggg gcagtacaag 2900

gagagcttcg ccaatgccgg cttcacctcc tttgacgtcg tgtctcagat 2950

gatgatggag gacattctcc gggttgggct cactttggct ggccaccaga 3000

aaaaaatcct gaacagtatc caggtgatgc gggcgcagat gaaccagatt 3050

cagtctgtgg aggtttgaca ttcacctgcc tcggctcacc tcttcctcca 3100

agccccgccc cctctgcccc acgtgccggc cctcctggtg ctctatccac 3150

tgcagggcca gccactcgcc aggaggccac gggccacggg aagaaccaag 3200

cggtgccagc cacgagacgt caccaagaaa acatgcaact caaacgacgg 3250

aaaaaaaaag ggaatgggaa aaaagaaaac agatcctggg aggggcgggg 3300

aaatacaagg aatatttttt aaagaggatt ctcataagga aagcaatgac 3350

tgttcttgcg ggggataaaa aagggcttgg gagattcatg cgatgtgtcc 3400

aatcggagac aaaagcagtt tctctccaac tccctctggg aaggtgacct 3450

ggccagagcc aagaaacact ttcagaaaaa caaatgtgaa ggggagagac 3500

aggggccgcc cttggctcct gtccctgctg ctcctctagg cctcactcaa 3550

caaccaagcg cctggaggac gggacagatg gacagacagc caccctgaga 3600

acccctctgg gaaaatctat tcctgccacc actgggcaaa cagaagaatt 3650

tttctgtctt tggagagtat tttagaaact ccaatgaaag acactgtttc 3700

tcctgttggc tcacagggct gaaaggggct tttgtcctcc tgggtcaggg 3750

agaacgcggg gaccccag 3768

<210> 23
<211> 3800
<212> DNA
<213> Homo sapiens

<400> 23
cgggaagcgc agccatggct ctgcggaggc tgggggccgc gctgctgctg 50

ctgccgctgc tcgccgccgt ggaagaaacg ctaatggact ccactacagc 100

gactgctgag ctgggctgga tggtgcatcc tccatcaggg tgggaagagg 150

tgagtggcta cgatgagaac atgaacacga tccgcacgta ccaggtgtgc 200

aacgtgtttg agtcaagcca gaacaactgg ctacggacca agtttatccg 250

gcgccgtggc gcccaccgca tccacgtgga gatgaagttt tcggtgcgtg 300

actgcagcag catccccagc gtgcctggct cctgcaagga gaccttcaac 350

ctctattact atgaggctga ctttgactcg gccaccaaga ccttccccaa 400

ctggatggag aatccatggg tgaaggtgga taccattgca gccgacgaga 450

gcttctccca ggtggacctg ggtggccgcg tcatgaaaat caacaccgag 500

gtgcggagct tcggacctgt gtcccgcagc ggcttctacc tggccttcca 550

ggactatggc ggctgcatgt ccctcatcgc cgtgcgtgtc ttctaccgca 600

agtgcccccg catcatccag aatggcgcca tcttccagga aaccctgtcg 650

ggggctgaga gcacatcgct ggtggctgcc cggggcagct gcatcgccaa 700

tgcggaagag gtggatgtac ccatcaagct ctactgtaac ggggacggcg 750

agtggctggt gcccatcggg cgctgcatgt gcaaagcagg cttcgaggcc 800

gttgagaatg gcaccgtctg ccgaggttgt ccatctggga ctttcaaggc 850

caaccaaggg gatgaggcct gtacccactg tcccatcaac agccggacca 900

cttctgaagg ggccaccaac tgtgtctgcc gcaatggcta ctacagagca 950

gacctggacc ccctggacat gccctgcaca accatccccc ccgcgcccca 1000

ggctgtgatt tccagtgtca atgagacctc cctcatgctg gagtggaccc 1050

ctccccgcga ctccggaggc cgagaggacc tcgtctacaa catcatctgc 1100

aagagctgtg gctcgggccg gggtgcctgc acccgctgcg gggacaatgt 1150

acagtacgca ccacgccagc taggcctgac cgagccacgc atttacatca 1200

gtgacctgct ggcccacacc cagtacacct tcgagatcca ggctgtgaac 1250

ggcgttactg accagagccc cttctcgcct cagttcgcct ctgtgaacat 1300

caccaccaac caggcagctc catcggcagt gtccatcatg catcaggtga 1350

gccgcaccgt ggacagcatt accctgtcgt ggtcccagcc ggaccagccc 1400

aatggcgtga tcctggacta tgagctgcag tactatgaga aggagctcag 1450

tgagtacaac gccacagcca taaaaagccc caccaacacg gtcaccgtgc 1500

agggcctcaa agccggcgcc atctatgtct tccaggtgcg ggcacgcacc 1550

gtggcaggct acgggcgcta cagcggcaag atgtacttcc agaccatgac 1600

agaagccgag taccagacaa gcatccagga gaagttgcca ctcatcatcg 1650

gctcctcggc cgctggcctg gtcttcctca ttgctgtggt tgtcatcgcc 1700

atcgtgtgta acagacgggg gtttgagcgt gctgactcgg agtacacgga 1750

caagctgcaa cactacacca gtggccacat gaccccaggc atgaagatct 1800

acatcgatcc tttcacctac gaggacccca acgaggcagt gcgggagttt 1850

gccaaggaaa ttgacatctc ctgtgtcaaa attgagcagg tgatcggagc 1900

aggggagttt ggcgaggtct gcagtggcca cctgaagctg ccaggcaaga 1950

gagagatctt tgtggccatc aagacgctca agtcgggcta cacggagaag 2000

cagcgccggg acttcctgag cgaagcctcc atcatgggcc agttcgacca 2050

tcccaacgtc atccacctgg agggtgtcgt gaccaagagc acacctgtga 2100

tgatcatcac cgagttcatg gagaatggct ccctggactc ctttctccgg 2150

caaaacgatg ggcagttcac agtcatccag ctggtgggca tgcttcgggg 2200

catcgcagct ggcatgaagt acctggcaga catgaactat gttcaccgtg 2250

acctggctgc ccgcaacatc ctcgtcaaca gcaacctggt ctgcaaggtg 2300

tcggactttg ggctctcacg ctttctagag gacgatacct cagaccccac 2350

ctacaccagt gccctgggcg gaaagatccc catccgctgg acagccccgg 2400

aagccatcca gtaccggaag ttcacctcgg ccagtgatgt gtggagctac 2450

ggcattgtca tgtgggaggt gatgtcctat ggggagcggc cctactggga 2500

catgaccaac caggatgtaa tcaatgccat tgagcaggac tatcggctgc 2550

caccgcccat ggactgcccg agcgccctgc accaactcat gctggactgt 2600

tggcagaagg accgcaacca ccggcccaag ttcggccaaa ttgtcaacac 2650

gctagacaag atgatccgca atcccaacag cctcaaagcc atggcgcccc 2700

tctcctctgg catcaacctg ccgctgctgg accgcacgat ccccgactac 2750

accagcttta acacggtgga cgagtggctg gaggccatca agatggggca 2800

gtacaaggag agcttcgcca atgccggctt cacctccttt gacgtcgtgt 2850

ctcagatgat gatggaggac attctccggg ttggggtcac tttggctggc 2900

caccagaaaa aaatcctgaa cagtatccag gtgatgcggg cgcagatgaa 2950

ccagattcag tctgtggagg tttgacattc acctgcctcg gctcacctct 3000

tcctccaagc cccgcccccc ctgccccacg tgccggccct cctggtgctc 3050

tatccactgc agggccagcc actcgccagg aggccacggg ccacgggaag 3100

aaccaagcgg tgccagccac gagacgtcac caagaaaaca tgcaactcaa 3150

acgacggaaa aaaaaaggga atgggaaaaa agaaaacaga tcctgggagg 3200

gggcgggaaa tacaaggaat atttttaaa gaggattctc ataaggaaag 3250

caatgactgt tcttgcgggg gataaaaaag ggcttgggag attcatgcga 3300

tgtgtccaat cggagacaaa agcagtttct ctccaactcc ctctgggaag 3350

gtgacctggc cagagccaag aaacactttc agaaaaacaa atgtgaaggg 3400

gagagacagg ggccgccctt ggctcctgtc cctgctgctc ctctaggcct 3450

cactcaacaa ccaagcgcct ggaggacggg acagatggac agacagccac 3500

cctgagaacc cctctgggaa aatctattcc tgccaccact gggcaaacag 3550

aagaattttt ctgtctttgg agagtatttt agaaactcca atgaaagaca 3600

ctgtttctcc tgttggctca cagggctgaa aggggctttt gtcctcctgg 3650

gtcagggaga acgcggggac cccagaaagg tcagccttcc tgaggatggg 3700

caacccccag gtctgcagct ccaggtacat atcacgcgca cagcctggca 3750

gcctggccct cctggtgccc actcccgcca gcccctgcct cgaggtcgac 3800

<210> 24
<211> 853
<212> DNA
<213> Homo sapiens

<400> 24
ccgccagctc cttcggcaat gaacttctcc accagcagca gcagcttcgc 50

ctacgaccgg gagttcctcc gcaccctgcc cggcttcctc atcgtggccg 100

agatcgttct ggggctgctg gtatggacgc ttattgctgg aactgagtac 150

ttccgggtcc ccgcatttgg ctgggtcatg tttgtagctg tattttactg 200

ggtcctcacc gtcttcttcc tcattatcta cataacaatg acctacacca 250

ggattcccca ggtgccctgg acaacagtgg gcctgtgctt taacggcagt 300

gccttcgtct tgtacctctc tgccgctgtt gtagatgcat cttccgtctc 350

ccctgagagg gacagtcaca acttcaacag ctgggcggcc tcatcgttct 400

ttgccttcct ggtcaacatc tgctacgctg gaaatacata tttcagtttt 450

atagcatgga gatccaggac catacagtga tttaccattt tgataattaa 500

aaggaaaaaa aaaggaagac tctcactgta aaaacagctg taggtataat 550

gtatattccc agagaattgt atttaactaa ttaatgtttt ttatattctt 600

aaatttgctc acaaattgtg gtttgttaca attaaactgg atacttattt 650

gcaaagtgtt gtagcttata atgaactctt aagtatctta ttaatgtatt 700

aatgtcttca tagatcatat tttcttagac aatgtttaaa tagataaatt 750

gctaatattg agaatgtgtc aagtttgtaa acctaacttt taagatgcca 800

gattcttttt tgattaaatg ttgcaaaatc ccaaaaaaaa aaaaaaaaaa 850

aaa 853

<210> 25
<211> 485
<212> DNA
<213> Homo sapiens

<400> 25
gcccttggca atgaacttct ccaccagcag cagcagcttc gcctacgacc 50

gggagttcct ccgcaccctg cccggcttcc tcatcgtggc cgagatcgtt 100

ctggggctgc tggtatggac gcttattgct ggaactgagt acttccgggt 150

ccccgcattt ggctgggtca tgtttgtagc tgtattttac tgggtcctca 200

ccgtcttctt cctcattatc tacataacaa tgacctacac caggattccc 250

caggtgccct ggacaacagt gggcctgtgc tttaacggca gtgccttcgt 300

cttgtacctc tctgccgctg ttgtagatgc atcttccgtc tcccctgaga 350

gggacagtca caacttcaac agctgggcgg cctcatcgtt ctttgccttc 400

ctggtcacca tctgctacgc tggaaataca tatttcagtt ttatagcatg 450

gagatccagg accatacagt gatttaccaa agggc 485

<210> 26
<211> 2257
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 2159, 2231
<223> unknown base

<400> 26
ggccctgccc tgaaggctgg tcacttgcag aggtaaactc ccctctttga 50

cttctggcca gggtttgtgc tgagctggct gcagccgctc tcagcctcgc 100

tccgggcacg tcgggcagcc tcgggccctc ctgcctgcag gatcatgccc 150

accaccgtgg acgatgtcct ggagcatgga ggggagtttc acttttccca 200

gaagcaaatg tttttcctct tggctctgct ctcggctacc ttcgcgccca 250

tctacgtggg catcgtcttc ctgggcttca cccctgacca ccgctgccgg 300

agccccggag tggccgagct gagtctgcgc tgcggctgga gtcctgcaga 350

ggaactgaac tacacggtgc cgggcccagg acctgcgggc gaagcctccc 400

caagacagtg taggcgctac gaggtggact ggaaccagag cacctttgac 450

tgcgtggacc ccctggccag cctggacacc aacaggagcc gcctgccact 500

gggcccctgc cgggacggct gggtgtacga gacgcctggc tcgtccatcg 550

tcaccgagtt taacctggta tgtgccaact cctggatgtt ggacctattc 600

cagtcatcag tgaatgtagg attctttatt ggctctatga gtatcggcta 650

catagcagac aggtttggcc gtaagctctg cctcctaact acagtcctca 700

taaatgctgc agctggagtt ctcatggcca tttccccaac ctatacgtgg 750

atgttaattt ttcgcttaat ccaaggactg gtcagcaaag caggctggtt 800

aataggctac atcctgatta cagaatttgt tgggcggaga tatcggagaa 850

cagtggggat tttttaccaa gttgcctata cagttgggct cctggtgcta 900

gctggggtgg cttacgcact tcctcactgg aggtggttgc agttcacagt 950

tgctctgccc aacttcttct tcttgctcta ttactggtgc atacctgagt 1000

ctcccaggtg gctgatctcc cagaataaga atgctgaagc catgagaatc 1050

attaagcaca tcgcaaagaa aaatggaaaa tctctacccg cctcccttca 1100

gcgcctgaga cttgaagagg aaactggcaa gaaattgaac ccttcatttc 1150

ttgacttggt cagaactcct cagataagga aacatactat gatattgatg 1200

tacaactggt tcacgagctc tgtgctctac cagggcctca tcatgcacat 1250

gggccttgca ggtgacaata tctacctgga tttcttctac tctgccctgg 1300

ttgaattccc agctgccttc atgatcatcc tcaccatcga ccgcatcgga 1350

cgccgttacc cttgggctgc atcaaatatg gttgcagggg cagcctgtct 1400

ggcctcagtt tttatacctg gtgatctaca atggctaaaa attattatct 1450

catgcttggg aagaatgggg atcacaatgg cctatgagat agtctgcctg 1500

gtcaatgctg agctgtaccc cacattcatt aggaatcttg gcgtccacat 1550

ctgttcctca atgtgtgaca ttggtggcat catcacgcca ttcctggtct 1600

accggctcac taacatctgg cttgagctcc cgctgatggt tttcggcgta 1650

cttggcttgg ttgctggagg tctggtgctg ttgcttccag aaactaaagg 1700

gaaagctttg cctgagacca tcgaggaagc cgaaaatatg caaagaccaa 1750

gaaaaaataa agaaaagatg atttacctcc aagttcagaa actagacatt 1800

ccattgaact aagaagagag accgttgctg ctgtcatgac ctagctttga 1850

tggcagcaag accaaaagta gaaatccctg cactcatcac aaagcccata 1900

caactcaacc aaacttaccc ctgagcccta tcaacctagg tctacagcca 1950

gtggagtcta ttgtacactg tggaaaaata cccatgggac cagatcctgc 2000

caaattcttc cagctcactt tattctcagc attcctagga cattggacat 2050

tggttttctg gagggttttt tttccgatct ttgtattttt ttaaatttga 2100

ttcttttctt tgcaatgcta gcaaccagaa tacatagggg aactgtgggc 2150

taggcaaana aaatagaaaa agtgtgaaaa acagtaaagt tgggagagga 2200

gcatctattt tcttaaagaa ataaaacacc naaaacaaaa aaaaaaaaaa 2250

aaaaaaa 2257

<210> 27
<211> 1725
<212> DNA
<213> Homo sapiens

<400> 27
gcccttctcc tgcctgcagg atcatgccca ccaccgtgga cgatgtcctg 50

gagcatggag gggagtttca ctttttccag aagcaaatgt ttttcctctt 100

ggctctgctc tcggctacct tcgcgcccat ctacgtgggc atcgtcttcc 150

tgggcttcac ccctgaccac cgctgccgga gccccggagt ggccgagctg 200

agtctgcgct gcggctggag tcctgcagag gaactgaact acacggtgcc 250

gggcccagga cctgcgggcg aagcctcccc aagacagtgt aggcgctacg 300

aggtggactg gaaccagagc accttcgact gcgtggaccc cctggccagc 350

ctggacacca acaggagccg cctgccactg ggcccctgcc gggacggctg 400

ggtgtacgag acgcctggct cgtccatcgt caccgagttt aacctggtat 450

gtgccaactc ctggatgttg gacctattcc agtcatcagt gaatgtagga 500

ttctttattg gctctatgag tatcggctac atagcagaca ggtttggccg 550

taagctctgc ctcctaacta cagtcctcat aaatgctgca gctggagttc 600

tcatggccat ttccccaacc tatacgtgga tgttaatttt tcgcttaatc 650

caaggactgg tcagcaaagc aggctggtta ataggctaca tcctgattac 700

agaatttgtt gggcggagat atcggagaac agtggggatt ttttaccaag 750

ttgcctatac agttgggctc ctggtgctag ctggggtggc ttacgcactt 800

cctcactgga ggtggttgca gttcacagtt tctctgccca acttcttctt 850

cttgctctat tactggtgca tacctgagtc tcccaggtgg ctgatctccc 900

agaataagaa tgctgaagcc atgagaatca ttaagcacat cgcaaagaaa 950

aatggaaaat ctctacccgc ctcccttcag cgcctgagac ttgaagagga 1000

aactggcaag aaattgaacc cttcatttct tgacttggtc agaactcctc 1050

agataaggaa acatactatg atattgatgt acaactggtt cacgagctct 1100

gtgctctacc agggcctcat catgcacatg ggccttgcag gtgacaatat 1150

ctacctggat ttcttctact ctgccctggt tgaattccca gctgccttca 1200

tgatcatcct caccatcgac cgcatcggac gccgttaccc ttgggctgca 1250

tcaaatatgg ttgcaggggc agcctgtctg gcctcagttt ttatacctgg 1300

tgatctacaa tggctaaaaa ttattatctc atgcttggga agaatgggga 1350

tcacaatggc ctatgagata gtctgcctgg tcaatgctga gctgtacccc 1400

acattcatta ggaatcttgg cgtccacatc tgttcctcaa tgtgtgacat 1450

tggtggcatc atcacgccat tcctggtcta ccggctcact aacatctggc 1500

ttgagctccc gctgatggtt ttcggcgtgc ttggcttggt tgctggaggt 1550

ctggtgctgt tgcttccaga aactaaaggg aaagctttgc ctgagaccat 1600

cgaggaagcc gaaaatatgc aaagaccaag aaaaaataaa gaaaagatga 1650

tttacctcca agttcagaaa ctagacattc cattgaacta agaagagaga 1700

ccgttgctgc tgtcatgaca agggc 1725

<210> 28
<211> 2280
<212> DNA
<213> Homo sapiens

<400> 28
cgggccaggt ttccaggctc ggccgccgcc tccatcccag cacctgcgga 50

gggagcgctg accatggctc cctggcctga attgggagat gcccagccca 100

accccgataa gtacctcgaa ggggccgcag gtcagcagcc cactgccccc 150

gataaaagca aagagaccaa caaaacagat aacactgagg cacctgtaac 200

caagattgaa cttctgccgt cctactccac ggctacactg atagatgagc 250

ccactgaggt ggatgacccc tggaacctac ccactcttca ggactcgggg 300

atcaagtggt cagagagaga caccaaaggg aagattctct gtttcttcca 350

agggattggg agattgattt tacttctcgg atttctctac tttttcgtgt 400

gctccctgga tattcttagt agcgccttcc agctggttgg aggaaaaatg 450

gcaggacagt tcttcagcaa cagctctatt atgtccaacc ctttgttggg 500

gctggtgatc ggggtgctgg tgaccgtctt ggtgcagagc tccagcacct 550

caacgtccat cgttgtcagc atggtgtcct cttcattgct cactgttcgg 600

gctgccatcc ccattatcat gggggccaac attggaacgt caatcaccaa 650

cactattgtt gcgctcatgc aggtgggaga tcggagtgag ttcagaagag 700

cttttgcagg agccactgtc catgacttct tcaactggct gtccgtgttg 750

gtgctcttgc ccgtggaggt ggccacccat tacctcgaga tcataaccca 800

gcttatagtg gagagcttcc acttcaagaa tggagaagat gccccagatc 850

ttctgaaagt catcactaag cccttcacaa agctcattgt ccagctggat 900

aaaaaagtta tcagccaaat tgcaatgaac gatgaaaaag cgaaaaacaa 950

gagtcttgtc aagatttggt gcaaaacttt taccaacaag acccagatta 1000

acgtcactgt tccctcgact gctaactgca cctccccttc cctctgttgg 1050

acggatggca tccaaaactg gaccatgaag aatgtgacct acaaggagaa 1100

catcgccaaa tgccagcata tctttgtgaa tttccacctc ccggatcttg 1150

ctgtgggcac catcttgctc atactctccc tgctggtcct ctgtggttgc 1200

ctgatcatga ttgtcaagat cctgggctct gtgctcaagg ggcaggtcgc 1250

cactgtcatc aagaagacca tcaacactga tttccccttt cccttgcat 1300

ggttgactgg ctacctggcc atcctcgtcg gggcaggcat gaccttcatc 1350

gtacagagca gctctgtgtt cacgtcggcc ttgacccccc tgattggaat 1400

cggcgtgata accattgaga gggcttatcc actcacgctg ggctccaaca 1450

tcggcaccac caccaccgcc atcctggccg cctagccag ccctggcaat 1500

gcattgagga gttcactcca gatcgccctg tgccactttt tcttcaacat 1550

ctccggcatc ttgctgtggt acccgatccc gttcactcgc ctgcccatcc 1600

gcatggccaa ggggctgggc aacatctctg ccaagtatcg ctggttcgcc 1650

gtcttctacc tgatcatctt cttcttcctg atcccgctga cggtgtttgg 1700

cctctcgctg gccggctggc gggtgctggt tggtgtcggg gttcccgtcg 1750

tcttcatcat catcctggta ctgtgcctcc gactcctgca gtctcgctgc 1800

ccacgcgtcc tgccgaagaa actccagaac tggaacttcc tgccgctgtg 1850

gatgcgctcg ctgaagccct gggatgccgt cgtctccaag ttcaccggct 1900

gcttccagat gcgctgctgc tactgctgcc gcgtgtgctg ccgcgcgtgc 1950

tgcttgctgt gtggctgccc caagtgctgc cgctgcagca agtgctgcga 2000

ggacttggag gaggcgcagg aggggcagga tgtccctgtc aaggctcctg 2050

agacctttga taacataacc attagcagag aggctcaggg tgaggtccct 2100

gcctcggact caaagaccga atgcacggcc ttgtagggga cgccccagat 2150

tgtcagggat ggggggatgg tccttgagtt ttgcatgctc tcctccctcc 2200

cacttctgca ccctttcacc acctcgagga gatttgctcc ccattagcga 2250

atgaaattga tgcagtccta aaaaaaaaaa 2280

<210> 29
<211> 2366
<212> DNA
<213> Homo sapiens

<400> 29
cagcccagca cctgcggagg gagcgctgac catggctccc tggcctgaat 50

tgggagatgc ccagcccaac cccgataagt acctcgaagg ggccgcaggt 100

cagcagccca ctgcccctga taaaagcaaa gagaccaaca aaacagataa 150

cactgaggca cctgtaacca agattgaact tctgccgtcc tactccacgg 200

ctacactgat agatgagccc actgaggtgg atgacccctg gaacctaccc 250

actcttcagg actcggggat caagtggtca gagagagaca ccaaagggaa 300

gattctctgt ttcttccaag ggattgggag attgatttta cttctcggat 350

ttctctactt tttcgtgtgc tccctggata ttcttagtag cgccttccag 400

ctggttggag gaaaaatggc aggacagttc ttcagcaaca gctctattat 450

gtccaaccct ttgttggggc tggtgatcgg ggtgctggtg accgtcttgg 500

tgcagagctc cagcacctca acgtccatcg ttgtcagcat ggtgtcctct 550

tcattgctca ctgttcgggc tgccatcccc attatcatgg gggccaacat 600

tggaacgtca atcaccaaca ctattgttgc gctcatgcag gtgggagatc 650

ggagtgagtt cagaagagct tttgcaggag ccactgtcca tgacttcttc 700

aactggctgt ccgtgttggt gctcttgccc gtggaggtgg ccacccatta 750

cctcgagatc ataacccagc ttatagtgga gagcttccac ttcaagaatg 800

gagaagatgc cccagatctt ctgaaagtca tcactaagcc cttcacaaag 850

ctcattgtcc agctggataa aaaagttatc agccaaattg caatgaacga 900

tgaaaaagcg aaaaacaaga gtcttgtcaa gatttggtgc aaaactttta 950

ccaacaagac ccagattaac gtcactgttc cctcgactgc taactgcacc 1000

tccccttccc tctgttggac ggatggcatc caaaactgga ccatgaagaa 1050

tgtgacctac aaggagaaca tcgccaaatg ccagcatatc tttgtgaatt 1100

tccacctccc ggatcttgct gtgggcacca tcttgctcat actctccctg 1150

ctggtcctct gtggttgcct gatcatgatt gtcaagatcc tgggctctgt 1200

gctcaagggg caggtcgcca ctgtcatcaa gaagaccatc aacactgatt 1250

tcccctttcc ctttgcatgg ttgactggct acctggccat cctcgtcggg 1300

gcaggcatga ccttcatcgt acagagcagc tctgtgttca cgtcggcctt 1350

gacccccctg attggaatcg gcgtgataac cattgagagg gcttatccac 1400

tcacgctggg ctccaacatc ggcaccacca ccaccgccat cctggccgcc 1450

ttagccagcc ctggcaatgc attgaggagt tcactccaga tcgccctgtg 1500

ccacttttte ttcaacatct ccggcatctt gctgtggtac ccgatcccgt 1550

tcactcgcct gcccatccgc atggccaagg ggctgggcaa catctctgcc 1600

aagtatcgct ggttcgccgt cttctacctg atcatcttct tcttcctgat 1650

cccgctgacg gtgtttggcc tctcgctggc cggctggcgg gtgctggttg 1700

gtgtcggggt tcccgtcgtc ttcatcatca tcctggtact gtgcctccga 1750

ctcctgcagt ctcgctgccc acgcgtcctg ccgaagaaac tccagaactg 1800

gaacttcctg ccgctgtgga tgcgctcgct gaagccctgg gatgccgtcg 1850

tctccaagtt caccggctgc ttccagatgc gctgctgctg ctgctgccgc 1900

gtgtgctgcc gcgcgtgctg cttgctgtgt ggctgcccca agtgctgccg 1950

ctgcagcaag tgctgcgagg acttggagga ggcgcaggag gggcaggatg 2000

tccctgtcaa ggctcctgag acctttgata acataaccat tagcagagag 2050

gctcagggtg aggtccctgc ctcggactca aagaccgaat gcacggcctt 2100

gtaggggacg ccccagattg tcagggatgg ggggatggtc cttgagtttt 2150

gcatgctctc ctccctccca cttctgcacc ctttcaccac ctcgaggaga 2200

tttgctcccc attagcgaat gaaattgatg cagtcctacc taactcgatt 2250

ccctttggct tggtgggtag gcctgcaggg cacttttatt ccaacccatg 2300

gcctccatga ctttttcaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2350

aaaaaaaaaa aaaaaa 2366

<210> 30
<211> 1633
<212> DNA
<213> Homo sapiens

<400> 30
ggcgtgggag ctgcctaggc cgggccctgc cagggagcaa gtctgcatcg 50

cagaggccag ggcagagtgg ctcctcacag cctgaagctc atccttctgc 100

acgggccagc caggccagca cagaggcacc agggcagcag tgcacacagg 150

tccccgggga ccccaccatg tggagcggat ggtggctgtg gcccccttgtg 200

gccgtctgca ctgcagactt ctttcgggac gaggcagaga ggatcatgag 250

ggactcccct gtcattgatg ggcacaatga cctcccctgg cagctgctgg 300

atatgttcaa caaccggctg caggacgaga gggccaacct gaccaccttg 350

gccggcacac acaccaacat ccccaagctg agggccggct ttgtgggagg 400

ccagttctgg tccgtgtaca cgccctgcga cacccagaac aaagacgccg 450

tgcggaggac gctggagcag atggacgtgg tccaccgcat gtgccggatg 500

tacccggaga ccttcctgta tgtcaccagc agtgcaggca ttcggcaggc 550

cttccgggaa gggaaggtgg ccagcctgat cggcgtggag ggcggccact 600

ccattgacag cagtttgggc gtcctgcggg cactctatca gctgggcatg 650

cggtacctga ccctcaccca cagctgcaac acgccctggg ctgacaactg 700

gctggtggac acgggagaca gcgagcccca gagccaaggc ttgtcaccct 750

ttgggcagcg tgtggtgaag gagctgaacc gtctgggggt cctcatcgac 800

ttggctcacg tgtctgtggc caccatgaag gccaccctgc agctgtccag 850

agccccggtc atcttcagcc actcctcggc ctacagcgtg tgcgcaagcc 900

ggcgcaacgt gcctgacgac gtcctgaggc tggtgaaaca gacagacagc 950

ctggtgatgg tgaacttcta caacaattac atttcctgca ccaacaaggc 1000

caacctgtcc caagtggccg accatctgga tcacatcaag gaggtggcag 1050

gagccagagc cgtgggtttt ggtggggact ttgatggtgt tccaagggtc 1100

cctgaggggc tggaggacgt ctccaagtat ccagacctga tcgctgagct 1150

gctcaggagg aactggacgg aggcggaggt caagggcgca ctggctgaca 1200

acctgctgag ggtcttccag gctgtggaac aggccagcaa cctcacacag 1250

gctcccgagg aggagcccat cccgctggac cagctgggtg gctcctgcag 1300

gacccattac ggctactcct ctggggcttc cagcctccat cgccactggg 1350

ggctcctgct ggcctccctc gctcccctgg tcctctgtct gtctctcctg 1400

tgaaacctgg gagaccagag tccccttcag ggttcccgga gctccgggaa 1450

gacccgccca tcccaggact ccagatgcca ggagccctgc tgcccacatg 1500

caaggaccag catctcctga gaggacgcct gggcttacct gggggcagg 1550

atgcctgggg acagttcagg acacacacac agtaggcccg caataaaagc 1600

aacacccctt caaaaaaaaa aaaaaaaaaa aaa 1633

<210> 31
<211> 1589
<212> DNA
<213> Homo sapiens

<400> 31
catctaaagc ctcctcagcc ttctgagtca gcctgaaagg aacaggccga 50

actgctgtat gggctctact gccagtgtga cctcaccctc tccagtcacc 100

cctcctcagt tccagctatg agttcctgca acttcacaca tgccaccttt 150

gtgcttattg gtatcccagg attagagaaa gcccatttct gggttggctt 200

ccccctcctt tccatgtatg tagtggcaat gtttggaaac tgcatcgtgg 250

tcttcatcgt aaggacggaa cgcagcctgc acgctccgat gtacctcttt 300

ctctgcatgc ttgcagccat tgacctggcc ttatccacat ccaccatgcc 350

taagatcctt gccctttttct ggtttgattc ccgagagatt agctttgagg 400

cctgtcttac ccagatgttc tttattcatg ccctctcagc cattgaatcc 450

accatcctgc tggccatggc ctttgaccgt tatgtggcca tctgccaccc 500

actgcgccat gctgcagtgc tcaacaatac agtaacagcc cagattggca 550

tcgtggctgt ggtccgcgga tccctctttt ttttcccact gcctctgctg 600

atcaagcggc tggccttctg ccactccaat gtcctctcgc actcctattg 650

tgtccaccag gatgtaatga agttggccta tgcagacact ttgcccaatg 700

tggtatatgg tcttactgcc attctgctgg tcatgggcgt ggacgtaatg 750

ttcatctcct tgtcctattt tctgataata cgaacggttc tgcaactgcc 800

ttccaagtca gagcgggcca aggcctttgg aacctgtgtg tcacacattg 850

gtgtggtact cgccttctat gtgccactta ttggcctctc agttgtacac 900

cgctttggaa acagccttca tcccattgtg cgtgttgtca tgggtgacat 950

ctacctgctg ctgcctcctg tcatcaatcc catcatctat ggtgccaaaa 1000

ccaaacagat cagaacacgg gtgctggcta tgttcaagat cagctgtgac 1050

aaggacttgc aggctgtggg aggcaagtga cccttaacac tacacttctc 1100

cttatcttta ttggcttgat aaacataatt atttctaaca ctagcttatt 1150

tccagttgcc cataagcaca tcagtacttt tctctggctg gaatagtaaa 1200

ctaaagtatg gtacatctac ctaaaggact attatgtgga ataatacata 1250

ctaatgaagt attacatgat ttaaagacta caataaaacc aaacatgctt 1300

ataacattaa gaaaaacaat aaagatacat gattgaaacc aagttgaaaa 1350

atagcatatg ccttggagga aatgtgctca aattactaat gatttagtgt 1400

tgtccctact ttctctctct ttttttcttc ttttttttt attatggtta 1450

gctgtcacat acaacttttt ttttttttga gatggggtct ccagcctggg 1500

caacagagca agaccctgtc tcaaagcata aaatggaata acatatcaaa 1550

tgaaacaggg aaaatgaagc tgacaattta tgggagcca 1589

<210> 32
<211> 7941
<212> DNA
<213> Homo sapiens

<400> 32
cacacatacg cacgcacgat ctcacttcga tctatacact ggaggattaa 50

aacaaacaaa caaaaaaaac atttccttcg ctccccctcc ctctccactc 100

tgagaagcag aggagccgca cggcgagggg ccgcagaccg tctggaaatg 150

cgaatcctaa agcgtttcct cgcttgcatt cagctcctct gtgtttgccg 200

cctggattgg gctaatggat actacagaca acagagaaaa cttgttgaag 250

agattggctg gtcctataca ggagcactga atcaaaaaaa ttggggaaag 300

aaatatccaa catgtaatag cccaaaacaa tctcctatca atattgatga 350

agatcttaca caagtaaatg tgaatcttaa gaaacttaaa tttcagggtt 400

gggataaaac atcattggaa aacacattca ttcataacac tgggaaaaca 450

gtggaaatta atctcactaa tgactaccgt gtcagcggag gagtttcaga 500

aatggtgttt aaagcaagca agataacttt tcactgggga aaatgcaata 550

tgtcatctga tggatcagag catagtttag aaggacaaaa atttccactt 600

gagatgcaaa tctactgctt tgatgcggac cgattttcaa gttttgagga 650

agcagtcaaa ggaaaaggga agttaagagc tttatccatt ttgtttgagg 700

ttgggacaga agaaaatttg gatttcaaag cgattattga tggagtcgaa 750

agtgttagtc gttttgggaa gcaggctgct ttagatccat tcatactgtt 800

gaaccttctg ccaaactcaa ctgacaagta ttacatttac aatggctcat 850

tgacatctcc tccctgcaca gacacagttg actggattgt ttttaaagat 900

acagttagca tctctgaaag ccagttggct gtttttgtg aagttcttac 950

aatgcaacaa tctggttatg tcatgctgat ggactactta caaaacaatt 1000

ttcgagagca acagtacaag ttctctagac aggtgttttc ctcatacact 1050

ggaaaggaag agattcatga agcagtttgt agttcagaac cagaaaatgt 1100

tcaggctgac ccagagaatt ataccagcct tcttgttaca tgggaaagac 1150

ctcgagtcgt ttatgatacc atgattgaga agtttgcagt tttgtaccag 1200

cagttggatg gagaggacca aaccaagcat gaatttttga cagatggcta 1250

tcaagacttg ggtgctattc tcaataattt gctacccaat atgagttatg 1300

ttcttcagat agtagccata tgcactaatg gcttatatgg aaaatacagc 1350

gaccaactga ttgtcgacat gcctactgat aatcctgaac ttgatctttt 1400

ccctgaatta attggaactg aagaaataat caaggaggag gaagagggaa 1450

aagacattga agaaggcgct attgtgaatc ctggtagaga cagtgctaca 1500

aaccaaatca ggaaaaagga accccagatt tctaccacaa cacactacaa 1550

tcgcataggg acgaaataca atgaagccaa gactaaccga tccccaacaa 1600

gaggaagtga attctctgga aagggtgatg ttcccaatac atctttaaat 1650

tccacttccc aaccagtcac taaattagcc acagaaaaag atatttcctt 1700

gacttctcag actgtgactg aactgccacc tcacactgtg gaaggtactt 1750

cagcctcttt aaatgatggc tctaaaactg ttcttagatc tccacatatg 1800

aacttgtcgg ggactgcaga atccttaaat acagtttcta taacagaata 1850

tgaggaggag agtttattga ccagtttcaa gcttgatact ggagctgaag 1900

attcttcagg ctccagtccc gcaacttctg ctatcccatt catctctgag 1950

aacatatccc aagggtatat attttcctcc gaaaacccag agacaataac 2000

atatgatgtc cttataccag aatctgctag aaatgcttcc gaagattcaa 2050

cttcatcagg ttcagaagaa tcactaaagg atccttctat ggagggaaat 2100

gtgtggtttc ctagctctac agacataaca gcacagcccg atgttggatc 2150

aggcagagag agctttctcc agactaatta cactgagata cgtgttgatg 2200

aatctgagaa gacaaccaag tccttttctg caggcccagt gatgtcacag 2250

ggtccctcag ttacagatct ggaaatgcca cattattcta cctttgccta 2300

cttcccaact gaggtaacac ctcatgcttt taccccatcc tccagacaac 2350

aggatttggt ctccacggtc aacgtggtat actcgcagac aacccaaccg 2400

gtatacaatg gtgagacacc tcttcaacct tcctacagta gtgaagtctt 2450

tcctctagtc acccctttgt tgcttgacaa tcagatcctc aacactaccc 2500

ctgctgcttc aagtagtgat tcggccttgc atgctacgcc tgtatttccc 2550

agtgtcgatg tgtcatttga atccatcctg tcttcctatg atggtgcacc 2600

tttgcttcca ttttcctctg cttccttcag tagtgaattg tttcgccatc 2650

tgcatacagt ttctcaaatc cttccacaag ttacttcagc taccgagagt 2700

gataaggtgc ccttgcatgc ttctctgcca gtggctgggg gtgatttgct 2750

attagagccc agccttgctc agtattctga tgtgctgtcc actactcatg 2800

ctgcttcaga gacgctggaa tttggtagtg aatctggtgt tctttataaa 2850

acgcttatgt tttctcaagt tgaaccaccc agcagtgatg ccatgatgca 2900

tgcacgttct tcagggcctg aaccttctta tgccttgtct gataatgagg 2950

gctcccaaca catcttcact gtttcttaca gttctgcaat acctgtgcat 3000

gattctgtgg gtgtaactta tcagggttcc ttatttagcg gccctagcca 3050

tataccaata cctaagtctt cgttaataac cccaactgca tcattactgc 3100

agcctactca tgccctctct ggtgatgggg aatggtctgg agcctcttct 3150

gatagtgaat ttcttttacc tgacacagat gggctgacag cccttaacat 3200

ttcttcacct gtttctgtag ctgaatttac atatacaaca tctgtgtttg 3250

gtgatgataa taaggcgctt tctaaaagtg aaataatata tggaaatgag 3300

actgaactgc aaattccttc tttcaatgag atggtttacc cttctgaaag 3350

cacagtcatg cccaacatgt atgataatgt aaataagttg aatgcgtctt 3400

tacaagaaac ctctgtttcc atttctagca ccaagggcat gtttccaggg 3450

tcccttgctc ataccaccac taaggttttt gatcatgaga ttagtcaagt 3500

tccagaaaat aacttttcag ttcaacctac acatactgtc tctcaagcat 3550

ctggtgacac ttcgcttaaa cctgtgctta gtgcaaactc agagccagca 3600

tcctctgacc ctgcttctag tgaaatgtta tctccttcaa ctcagctctt 3650

attttatgag acctcagctt cttttagtac tgaagtattg ctacaacctt 3700

cctttcaggc ttctgatgtt gacaccttgc ttaaaactgt tcttccagct 3750

gtgcccagtg atccaatatt ggttgaaacc cccaaagttg ataaaattag 3800

ttctacaatg ttgcatctca ttgtatcaaa ttctgcttca agtgaaaaca 3850

tgctgcactc tacatctgta ccagtttttg atgtgtcgcc tacttctcat 3900

atgcactctg cttcacttca aggtttgacc atttcctatg caagtgagaa 3950

atatgaacca gttttgttaa aaagtgaaag ttcccaccaa gtggtacctt 4000

ctttgtacag taatgatgag ttgttccaaa cggccaattt ggagattaac 4050

caggcccatc ccccaaaagg aaggcatgta tttgctacac ctgttttatc 4100

aattgatgaa ccattaaata cactaataaa taagcttata cattccgatg 4150

aaattttaac ctccaccaaa agttctgtta ctggtaaggt atttgctggt 4200

attccaacag ttgcttctga tacatttgta tctactgatc attctgttcc 4250

tataggaaat gggcatgttg ccattacagc tgtttctccc cacagagatg 4300

gttctgtaac ctcaacaaag ttgctgtttc cttctaaggc aacttctgag 4350

ctgagtcata gtgccaaatc tgatgccggt ttagtgggtg gtggtgaaga 4400

tggtgacact gatgatgatg gtgatgatga tgatgacaga gatagtgatg 4450

gcttatccat tcataagtgt atgtcatgct catcctatag agaatcacag 4500

gaaaaggtaa tgaatgattc agacacccac gaaaacagtc ttatggatca 4550

gaataatcca atctcatact cactatctga gaattctgaa gaagataata 4600

gagtcacaag tgtatcctca gacagtcaaa ctggtatgga cagaagtcct 4650

ggtaaatcac catcagcaaa tgggctatcc caaaagcaca atgatggaaa 4700

agaggaaaat gacattcaga ctggtagtgc tctgcttcct ctcagccctg 4750

aatctaaagc atgggcagtt ctgacaagtg atgaagaaag tggatcaggg 4800

caaggtacct cagatagcct taatgagaat gagacttcca cagatttcag 4850

ttttgcagac actaatgaaa aagatgctga tgggatcctg gcagcaggtg 4900

actcagaaat aactcctgga ttcccacagt ccccaacatc atctgttact 4950

agcgagaact cagaagtgtt ccacgtttca gaggcagagg ccagtaatag 5000

tagccatgag tctcgtattg gtctagctga ggggttggaa tccgagaaga 5050

aggcagttat acccccttgtg atcgtgtcag ccctgacttt tatctgtcta 5100

gtggttcttg tgggtattct catctactgg aggaaatgct tccagactgc 5150

acacttttac ttagaggaca gtacatcccc tagagttata tccacacctc 5200

caacacctat ctttccaatt tcagatgatg tcggagcaat tccaataaag 5250

cactttccaa agcatgttgc agatttacat gcaagtagtg ggtttactga 5300

agaatttgag acactgaaag agttttacca ggaagtgcag agctgtactg 5350

ttgacttagg tattacagca gacagctcca accacccaga caacaagcac 5400

aagaatcgat acataaatat cgttgcctat gatcatagca gggttaagct 5450

agcacagctt gctgaaaagg atggcaaact gactgattat atcaatgcca 5500

attatgttga tggctacaac agaccaaaag cttatattgc tgcccaaggc 5550

ccactgaaat ccacagctga agatttctgg agaatgatat gggaacataa 5600

tgtggaagtt attgtcatga taacaaacct cgtggagaaa ggaaggagaa 5650

aatgtgatca gtactggcct gccgatggga gtgaggagta cgggaacttt 5700

ctggtcactc agaagagtgt gcaagtgctt gcctattata ctgtgaggaa 5750

ttttactcta agaaacacaa aaataaaaaa gggctcccag aaaggaagac 5800

ccagtggacg tgtggtcaca cagtatcact acacgcagtg gcctgacatg 5850

ggagtaccag agtactccct gccagtgctg acctttgtga gaaaggcagc 5900

ctatgccaag cgccatgcag tggggcctgt tgtcgtccac tgcagtgctg 5950

gagttggaag aacaggcaca tatattgtgc tagacagtat gttgcagcag 6000

attcaacacg aaggaactgt caacatattt ggcttcttaa aacacatccg 6050

ttcacaaaga aattatttgg tacaaactga ggagcaatat gtcttcattc 6100

atgatacact ggttgaggcc atacttagta aagaaactga ggtgctggac 6150

agtcatattc atgcctatgt taatgcactc ctcattcctg gaccagcagg 6200

caaaacaaag ctagagaaac aattccagct cctgagccag tcaaatatac 6250

agcagagtga ctattctgca gccctaaagc aatgcaacag ggaaaagaat 6300

cgaacttctt ctatcatccc tgtggaaaga tcaagggttg gcatttcatc 6350

cctgagtgga gaaggcacag actacatcaa tgcctcctat atcatgggct 6400

attaccagag caatgaattc atcattaccc agcaccctct ccttcatacc 6450

atcaaggatt tctggaggat gatatgggac cataatgccc aactggtggt 6500

tatgattcct gatggccaaa acatggcaga agatgaattt gtttactggc 6550

caaataaaga tgagcctata aattgtgaga gctttaaggt cactcttatg 6600

gctgaagaac acaaatgtct atctaatgag gaaaaactta taattcagga 6650

ctttatctta gaagctacac aggatgatta tgtacttgaa gtgaggcact 6700

ttcagtgtcc taaatggcca aatccagata gccccattag taaaactttt 6750

gaacttataa gtgttataaa agaagaagct gccaataggg atgggcctat 6800

gattgttcat gatgagcatg gaggagtgac ggcaggaact ttctgtgctc 6850

tgacaaccct tatgcaccaa ctagaaaaag aaaattccgt ggatgtttac 6900

caggtagcca agatgatcaa tctgatgagg ccaggagtct ttgctgacat 6950

tgagcagtat cagtttctct acaaagtgat cctcagcctt gtgagcacaa 7000

ggcaggaaga gaatccatcc acctctctgg acagtaatgg tgcagcattg 7050

cctgatggaa atatagctga gagcttagag tctttagttt aacacagaaa 7100

ggggtggggg gactcacatc tgagcattgt tttcctcttc ctaaaattag 7150

gcaggaaaat cagtctagtt ctgttatctg ttgatttccc atcacctgac 7200

agtaactttc atgacatagg attctgccgc caaatttata tcattaacaa 7250

tgtgtgcctt tttgcaagac ttgtaattta cttattatgt ttgaactaaa 7300

atgattgaat tttacagtat ttctaagaat ggaattgtgg tatttttttc 7350

tgtattgatt ttaacagaaa atttcaattt atagaggtta ggaattccaa 7400

actacagaaa atgtttgttt ttagtgtcaa atttttagct gtatttgtag 7450

caattatcag gtttgctaga aatataactt ttaatacagt agcctgtaaa 7500

taaaacactc ttccatatga tattcaacat tttacaactg cagtattcac 7550

ctaaagtaga aataatctgt tacttattgt aaatactgcc ctagtgtctc 7600

catggaccaa atttatattt ataattgtag atttttatat tttactactg 7650

agtcaagttt tctagttctg tgtaattgtt tagtttaatg acgtagttca 7700

ttagctggtc ttactctacc agttttctga cattgtattg tgttacctaa 7750

gtcattaact ttgtttcagc atgtaatttt aacttttgtg gaaaatagaa 7800

ataccttcat tttgaaagaa gtttttatga gaataacacc ttaccaaaca 7850

ttgttcaaat ggtttttatc caaggaattg caaaaataaa tataaatatt 7900

gccattaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa a 7941

<210> 33
<211> 2755
<212> DNA
<213> Homo sapiens

<400> 33
attgtctggg aattgcagcc gcggggcggg cggcggcggc ggcggcggcg 50

gccgggaccc agcgggccag gtggggacgg cgcggagcgg gtgcgggaga 100

tgccgtgcgg gactggggcc acctgagccg cccgcctcgt ccccgccttc 150

tgtgggaagg atgtgcgcgc ggatggccgg tcgcacaaga gcggcccctc 200

gggggcccta cggcccctgg ctctgcctcc tggtggccct cgccctggac 250

gtcgtgagag tggactgtgg ccaggctccc ctggaccctg tctacctgca 300

tgtgacagcc gcccgcccag cccagcccac actctggact gccaagctag 350

accgcttcaa gggctccagg caccacacca ccctcatcac ctgccaccgt 400

gctgggctca cagagccaga ttccagcagt ccccttgaac tgtctgagtt 450

cctatgggtg gactttgtgg tggagaatag cactggtggg ggcgtagcgg 500

tcactcgccc cgtcacgtgg cagctggagt acccaggcca ggcccctgaa 550

gcagagaagg acaaaatggt gtgggaaatc ctggtgtctg agcgggacat 600

cagagccctt atcccactgg ccaaggctga ggagctggtg aatacagcac 650

cactgactgg agtgccccag catgtccccg tgcgccttgt cactgtggac 700

ggcgggggcg ccttggtgga ggtgacagag catgtcggct gcgagtctgc 750

caacacacag gtcctgcagg tgtctgaggc ctgtgatgcc gtgttcgtgg 800

ctggcaagga gagccggggc gcccgggggg tgcgagtgga cttctggtgg 850

cgccggctcc gcgcctcgct gcggctgacc gtgtgggccc cgctgctacc 900

gctgcgtatc gagctcaccg acaccaccct cgagcaggtc cgcggctgga 950

gggtacctgg ccctgctgaa gggcctgcgg aacccgctgc agaggcgtca 1000

gatgaggccg agcggcgcgc ccgtggctgc cacctgcagt accagcgggc 1050

cggtgtgcgc ttcctcgccc ccttcgcggc ccacccgctg gacggcggcc 1100

gccgcctcac gcacctgctt ggccccgact ggctgctaga cgtgtcccac 1150

ctcgtggcgc cacacgcccg cgtgctggac tcgcgtgtag cctctctgga 1200

gggtggccgt gtcgtggtgg gccgggagcc cggtgtcacc tccattgagg 1250

tgcgttcccc actgtctgac tccatcctgg gggagcaggc gctggctgtg 1300

acggacgaca aggtctcagt gctggagctg agggtgcagc cagtgatggg 1350

catctcgctg accttgagcc ggggcactgc ccaccccggg gaggtcacag 1400

ctacgtgctg ggcacagtca gcccttcccg ccccaaagca ggaggtggcc 1450

ctctccctat ggctgtcctt ctctgatcac actgtggccc cagctgagct 1500

ctacgaccgc cgtgacctgg gactgtccgt ctcagccgag gagcctggtg 1550

ccatcctgcc agctgaggag cagggtgccc agctcggggt ggtggtgagt 1600

ggggcaggcg ccgaggggct gccgctgcat gtggctctgc acccgcccga 1650

gccctgccgc cggggccgcc accgtgtgcc tctggcctct ggcaccgcct 1700

ggctggggct gccccctgcc tccactccag cccctgctct cccatccagc 1750

cctgcttgga gcccaccagc cacagaagcc accatgggtg gtaaacggca 1800

ggtggcaggc agtgtcgggg gcaacacagg tgtgaggggc aagtttgagc 1850

gggcagagga ggaggccagg aaggaggaga ccgaagccag ggaggaggag 1900

gaggaagagg aggaggagat ggtccctgcc cctcagcatg tcactgagct 1950

agagctgggc atgtacgccc tgctgggagt cttctgcgtg gccatcttca 2000

tcttcttggt caatggtgtg gtcttcgtcc tgcgctatca gcgcaaagaa 2050

cctcccgaca gtgccactga ccccacctcc ccccagcccc acaactgggt 2100

ctggctgggc actgaccagg aggaactgag ccgccagctg gaccggcagt 2150

cccctggccc gcccaagggg gaggggagct gcccctgtga gagtgggggа 2200

ggaggggagg cccctaccct ggcccctggc cctcctgggg gcaccaccag 2250

ctcctcaagc accctggccc gaaaggaggc tgggggggcgg cggaagcgag 2300

tagagtttgt gacatttgtg ccagcccctc cagcccagtc acctgaggag 2350

cctgtagggg cccctgctgt gcagtccatc cttgtggcag gcgaggagga 2400

catccgctgg gtgtgtgagg acatggggct gaaggaccct gaggagcttc 2450

gcaactacat ggagaggatc cggggcagct cctgaccctc cacagccacc 2500

tggtcagcca ccagctgggg caacgagggt ggaggtccca ctgagcctct 2550

cgcctgcccc cgccactcgt ctggtgcttg ttgatccaag tcccctgcct 2600

ggtcccccac aaggactccc atccaggccc cctctgccct gccccttgtc 2650

atggaccatg gtcgtgagga agggctcatg ccccttattt atgggaacca 2700

tctcattcta acagaataaa ccgagaagga aaccagaaaa aaaaaaaaaa 2750

aaaaa 2755

<210> 34
<211> 1795
<212> DNA
<213> Homo sapiens

<400> 34
acgcgtccgc ccacgcgtcc gcccacgcgt ccggtcgggg ccagagcgca 50

ggtgtacctg gcggccgtgc tggagcacct gaccgccgag atcctggagc 100

tggctggcaa cccggcccgc gacaagaaga cccgcatcat cctgcgccac 150

ctgtagctgg ccattcgcaa cggcgaggag cttaacaagc tgctgggcga 200

agtcaccatc gcgcagggcg gtgtcctgcc caacattcag ggcgtgcttc 250

tgccccagaa gaccaagagc caccacaagg ccaagggtga aaaccattca 300

ctaggagagg agaaacacaa tggccaccaa gacagagttg agtcccacag 350

caagggagag caagaacgca caagatatgc aagtggatga gacactgatc 400

cccaggaaag gtccaagttt atgttctgct cgctatggaa tagccctcgt 450

cttacatttc tgcaatttca caacgatagc acaaaatgtc atcatgaaca 500

tcaccatggt agccatggtc aacagcacaa gccctcaatc ccagctcaat 550

gattcctctg aggtgctgcc tgttgactca tttggtggcc taagtaaagc 600

cccaaagagt cttcctgcaa agtcctcaat acttgggggt cagtttgcaa 650

tttgggaaaa gtggggccct ccacaagaac gaagcagact ctgcagcatt 700

gctttatcag gaatgttact gggatgcttt actgccatcc tcataggtgg 750

cttcattagt gaaacccttg ggtggccctt tgtcttctat atctttggag 800

gtgttggctg tgtctgctgc cttctctggt ttgttgtgat ttatgatgac 850

ccctttttcct atccatggat aagcacctca gaaaaagaat acatcatatc 900

ctccttgaaa caacaggtcg ggtcttctaa gcagcctctt cccatcaaag 950

ctatgctcag atctctaccc atttggtcca tatgtttagg ctgtttcagc 1000

catcaatggt tagttagcac aatggttgta tacataccaa cttacatcag 1050

ctctgtgtac catgttaaca tcagagacaa tggacttcta tctgcccttc 1100

cttttattgt tgcctgggtc ataggcatgg tgggaggcta tctggcagat 1150

ttccttctaa ccaaaaagtt tagactcatc actgtgagga aaattgccac 1200

aattttagga agtctccccct cttcagcact cattgtgtct ctgccttacc 1250

tcaattccgg ctatatcaca gcaactgcct tgctgacgct ctcttgcgga 1300

ttaagcacat tgtgtcagtc agggatttat atcaatgtct tagatattgc 1350

tccaaggtat tccagttttc tcatgggagc atcaagagga ttttcgagca 1400

tagcacctgt cattgtaccc actgtcagcg gatttcttct tagtcaggac 1450

cctgagtttg ggtggaggaa tgtcttcttc ttgctgtttg ccgttaacct 1500

gttaggacta ctcttctacc tcatatttgg agaagcagat gtccaagaat 1550

gggctaaaga gagaaaactc actcgtttat gaagttatcc caccttggat 1600

ggaaaagtca ttaggcaccg tattgcataa aatagaaggc ttccgtgatg 1650

aaaataccag tgaaaagatt tttttttcct gtggctcttt tcaattatga 1700

gatcagttca ttatttatt cagacttttt tttgagagaa atgtaagatg 1750

aataaaaatt caaataaaat gataactaag aaaaaaaaaa aaaaa 1795

<210> 35
<211> 1795
<212> DNA
<213> Homo sapiens

<400> 35
acgcgtccgc ccacgcgtcc gcccacgcgt ccggtcgggg ccagagcgca 50

ggtgtacctg gcggccgtgc tggagcacct gaccgccgag atcctggagc 100

tggctggcaa cccggcccgc gacaagaaga cccgcatcat cctgcgccac 150

ctgtagctgg ccattcgcaa cggcgaggag cttaacaagc tgctgggcga 200

agtcaccatc gcgcagggcg gtgtcctgcc caacattcag ggcgtgcttc 250

tgccccagaa gaccaagagc caccacaagg ccaagggtga aaaccattca 300

ctaggagagg agaaacacaa tggccaccaa gacagagttg agtcccacag 350

caagggagag caagaacgca caagatatgc aagtggatga gacactgatc 400

cccaggaaag gtccaagttt atgttctgct cgctatggaa tagccctcgt 450

cttacatttc tgcaatttca caacgatagc acaaaatgtc atcatgaaca 500

tcaccatggt agccatggtc aacagcacaa gccctcaatc ccagctcaat 550

gattcctctg aggtgctgcc tgttgactca tttggtggcc taagtaaagc 600

cccaaagagt cttcctgcaa agtcctcaat acttgggggt cagtttgcaa 650

tttgggaaaa gtggggccct ccacaagaac gaagcagact ctgcagcatt 700

gctttatcag gaatgttact gggatgcttt actgccatcc tcataggtgg 750

cttcattagt gaaacccttg ggtggccctt tgtcttctat atctttggag 800

gtgttggctg tgtctgctgc cttctctggt ttgttgtgat ttatgatgac 850

ccctttttcct atccatggat aagcacctca gaaaaagaat acatcatatc 900

ctccttgaaa caacaggtcg ggtcttctaa gcagcctctt cccatcaaag 950

ctatgctcag atctctaccc atttggtcca tatgtttagg ctgtttcagc 1000

catcaatggt tagttagcac aatggttgta tacataccaa cttacatcag 1050

ctctgtgtac catgttaaca tcagagacaa tggacttcta tctgcccttc 1100

cttttattgt tgcctgggtc ataggcatgg tgggaggcta tctggcagat 1150

ttccttctaa ccaaaaagtt tagactcatc actgtgagga aaattgccac 1200

aattttagga agtctccccct cttcagcact cattgtgtct ctgccttacc 1250

tcaattccgg ctatatcaca gcaactgcct tgctgacgct ctcttgcgga 1300

ttaagcacat tgtgtcagtc agggatttat atcaatgtct tagatattgc 1350

tccaaggtat tccagttttc tcatgggagc atcaagagga ttttcgagca 1400

tagcacctgt cattgtaccc actgtcagcg gatttcttct tagtcaggac 1450

cctgagtttg ggtggaggaa tgtcttcttc ttgctgtttg ccgttaacct 1500

gttaggacta ctcttctacc tcatatttgg agaagcagat gtccaagaat 1550

gggctaaaga gagaaaactc actcgtttat gaagttatcc caccttggat 1600

ggaaaagtca ttaggcaccg tattgcataa aatagaaggc ttccgtgatg 1650

aaaataccag tgaaaagatt ttttttcct gtggctcttt tcaattatga 1700

gatcagttca ttattttatt cagacttttt tttgagagaa atgtaagatg 1750

aataaaaatt caaataaaat gataactaag aaaaaaaaaa aaaaa 1795

<210> 36
<211> 4524
<212> DNA
<213> Homo sapiens

<400> 36
gctcgctggg ccgctgctcc cgggtgtccc aggcccggcc ggtgcgcaga 50

gcatggcggg tgcgggcccg aagcggcgcg cgctagcggc gccggcggcc 100

gaggagaagg aagaggcgcg ggagaagatg ctggccgcca agagcgcgga 150

cggctcggcg ccggcaggcg agggcgaggg cgtgaccctg cagcggaaca 200

tcacgctgct caacggcgtg gccatcatcg tggggaccat tatcggctcg 250

ggcatcttcg tgacgcccac gggcgtgctc aaggaggcag gctcgccggg 300

gctggcgctg gtggtgtggg ccgcgtgcgg cgtcttctcc atcgtgggcg 350

cgctctgcta cgcggagctc ggcaccacca tctccaaatc gggcggcgac 400

tacgcctaca tgctggaggt ctacggctcg ctgcccgcct tcctcaagct 450

ctggatcgag ctgctcatca tccggccttc atcgcagtac atcgtggccc 500

tggtcttcgc cacctacctg ctcaagccgc tcttccccac ctgcccggtg 550

cccgaggagg cagccaagct cgtggcctgc ctgtgcgtgc tgctgctcac 600

ggccgtgaac tgctacagcg tgaaggccgc cacccgggtc caggatgcct 650

ttgccgccgc caagctcctg gccctggccc tgatcatcct gctgggcttc 700

gtccagatcg gaaagggtga tgtgtccaat ctagatccca agttctcatt 750

tgaaggcacc aaactggatg tggggaacat tgtgctggca ttatacagcg 800

gcctctttgc ctatggagga tggaattact tgaatttcgt cacagaggaa 850

atgatcaacc cctacagaaa cctgcccctg gccatcatca tctccctgcc 900

catcgtgacg ctggtgtacg tgctgaccaa cctggcctac ttcaccaccc 950

tgtccaccga gcagatgctg tcgtccgagg ccgtggccgt ggacttcggg 1000

aactatcacc tgggcgtcat gtcctggatc atccccgtct tcgtgggcct 1050

gtcctgcttt ggctccgtca atgggtcccc gttcacatcc tccaggctct 1100

tcttcgtggg gtcccgggaa ggccacctgc cctccatcct ctccatgatc 1150

cacccacagc tcctcacccc cgtgccgtcc ctcgtgttca cgtgtgtgat 1200

gacgctgctc tacgccttct ccaaggacat cttctccgtc atcaacttct 1250

tcagcttctt caactggctc tgcgtggccc tggccatcat cggcatgatc 1300

tggctgcgcc acagaaagcc tgagcttgag cggcccatca aggtgaacct 1350

ggccctgcct gtgttcttca tcctggcctg cctcttcctg atcgccgtct 1400

ccttctggaa gacacccgtg gagtgtggca tcggcttcac catcatcctc 1450

agcgggctgc ccgtctactt cttcggggtc tggtggaaaa acaagcccaa 1500

gtggctcctc cagggcatct tctccacgac cgtcctgtgt cagaagctca 1550

tgcaggtggt cccccaggag acatagccag gaggccgagt ggctgccgga 1600

ggagcatgcg cagaggccag ttaaagtaga tcacctcctc gaacccactc 1650

cggttccccg caacccacag ctcagctgcc catcccagtc ctcgccgtcc 1700

ctcccaggtc gggcagtgga ggctgctgtg aaaactctgg tacgaatctc 1750

atccctcaac tgagggccag ggacccaggt gtgcctgtgc tcctgcccag 1800

gagcagcttt tggtctcctt gggccctttt tcccttccct cctttgttta 1850

cttatatata tatttttttt aaacttaaat tttgggtcaa cttgacacca 1900

ctaagatgat tttttaagga gctgggggaa ggcaggagcc ttcctttctc 1950

ctgccccaag ggcccagacc ctgggcaaac agagctactg agacttggaa 2000

cctcattgct accacagact tgcactgaag ccagacagct gcccagacac 2050

atgggcttgt gacattcgtg aaaaccaacc ctgtgggctt atgtctctgc 2100

cttagggttt gcagagtgga aactcagccg tagggtggca ctgggagggg 2150

gtgggggatc tgggcaaggt gggtgattcc tcccaggagg tgcttgaggc 2200

cccgatggac tcctgaccat aatcctagcc ccgagacacc atcctgagcc 2250

agggaacagc cccagggttg gggggtgccg gcatctcccc tagctcacca 2300

ggcctggcct ctgggcagtg tggcctcttg gctatttctg ttccagtttt 2350

ggaggctgag ttctggttca tgcagacaaa gccctgtcct tcagtcttct 2400

agaaacagag acaagaaagg cagacacacc gcggccaggc acccatgtgg 2450

gcgcccaccc tgggctccac acagcagtgt cccctgcccc agaggtcgca 2500

gctaccctca gcctccaatg cattggcctc tgtaccgccc ggcagcccct 2550

tctggccggt gctgggttcc cactcccggc ctaggcacct ccccgctctc 2600

cctgtcacgc tcatgtcctg tcctggtcct gatgcccgtt gtctaggaga 2650

cagagccaag cactgctcac gtctctgccg cctgcgtttg gaggcccctg 2700

ggctctcacc cagtccccac ccgcctgcag agagggaact agggcacccc 2750

ttgtttctgt tgttcccgtg aatttttttc gctatgggag gcagccgagg 2800

cctggccaat gcggcccact ttcctgagct gtcgctgcct ccatggcagc 2850

agccaaggac ccccagaaca agaagacccc cccgcaggat ccctcctgag 2900

ctcggggggc tctgccttct caggccccgg gcttcccttc tccccagcca 2950

gaggtggagc caagtggtcc agcgtcactc cagtgctcag ctgtggctgg 3000

aggagctggc ctgtggcaca gccctgagtg tcccaagccg ggagccaacg 3050

aagccggaca cggcttcact gaccagcggc tgctcaagcc gcaagctctc 3100

agcaagtgcc cagtggagcc tgccgccccc acctgggcac cgggaccccc 3150

tcaccatcca gtgggcccgg agaaacctga tgaacagttt ggggactcag 3200

gaccagatgt ccgtctctct tgcttgagga atgaagacct ttattcaccc 3250

ctgccccgtt gcttcccgct gcacatggac agacttcaca gcgtctgctc 3300

ataggacctg catccttcct ggggacgaat tccactcgtc caagggacag 3350

cccacggtct ggaggccgag gaccaccagc aggcaggtgg actgactgtg 3400

ttgggcaaga cctcttccct ctgggcctgt tctcttggct gcaaataagg 3450

acagcagctg gtgccccacc tgcctggtgc attgctgtgt gaatccagga 3500

ggcagtggac atcgtaggca gccacggccc caggtccagg agaagtgctc 3550

cctggaggca cggaccactg cttcccactg gggccggcgg ggcccacgca 3600

cgacgtcagc ctcttacctt cccgcctcgg ctaggggtcc tcgggatgcc 3650

gttctgttcc aacctcctgt tctgggaggt ggacatgcct caaggataca 3700

gggagccggc ggcctctcga cggcacgcac ttcctgttgg ctgctgcggc 3750

tgtgggcgag catgggggct gccagcgtct gttgtggaaa gtagctgcta 3800

gtgaaatggc tggggccgct ggggtccgtc ttcacactgc gcaggtctct 3850

tctgggcgtc tgagctgggg tgggagctcc tccgcagaag gttggtgggg 3900

ggtccagtct gtgatccttg gtgctgtgtg ccccactcca gcctggggac 3950

cccacttcag aaggtagggg ccgtgtcccg cggtgctgac tgaggcctgc 4000

ttccccctcc ccctcctgct gtgctggaat tccacaggga ccagggccac 4050

cgcaggggac tgtctcagaa gacttgattt ttccgtccct ttttctccac 4100

actccactga caaacgtccc cagcggtttc cacttgtggg cttcaggtgt 4150

tttcaagcac aacccaccac aacaagcaag tgcattttca gtcgttgtgc 4200

ttttttgttt tgtgctaacg tcttactaat ttaaagatgc tgtcggcacc 4250

atgtttattt atttccagtg gtcatgctca gccttgctgc tctgcgtggc 4300

gcaggtgcca tgcctgctcc ctgtctgtgt cccagccacg cagggccatc 4350

cactgtgacg tcggccgacc aggctggaca ccctctgccg agtaatgacg 4400

tgtgtggctg ggaccttctt tattctgtgt taatggctaa cctgttacac 4450

tgggctgggt tgggtagggt gttctggctt ttttgtgggg tttttatttt 4500

taaagaaaca ctcaatcatc ctag 4524

<210> 37
<211> 1590
<212> DNA
<213> Homo sapiens

<400> 37
attcgccctt tgctcccggg tgtcccaggc ccggccggtg cgcagagcat 50

ggcgggtgcg ggcccgaagc ggcgcgcgct agcggcgccg gcggccgagg 100

agaaggaaga ggcgcgggag aagatgctgg ccgccaagag cgcggacggc 150

tcggcgccgg caggcgaggg cgagggcgtg accctgcagc ggaacatcac 200

gctgctcaac ggcgtggcca tcatcgtggg gaccattatc ggctcgggca 250

tcttcgtgac gcccacgggc gtgctcaagg aggcaggctc gccggggctg 300

gcgctggtgg tgtgggccgc gtgcggcgtc ttctccatcg tgggcgcgct 350

ctgctacgcg gagctcggca ccaccatctc caaatcgggc ggcgactacg 400

cctacatgct ggaggtctac ggctcgctgc ccgccttcct caagctctgg 450

atcgagctgc tcatcatccg gccttcatcg cagtacatcg tggccctggt 500

cttcgccacc tacctgctca agccgctctt ccccacctgc ccggtgcccg 550

aggaggcagc caagctcgtg gcctgcctct gcgtgctgct gctcacggcc 600

gtgaactgct acagcgtgaa ggccgccacc cgggtccagg atgcctttgc 650

cgccgccaag ctcctggccc tggccctgat catcctgctg ggcttcgtcc 700

agatcgggaa gggtgatgtg tccaatctag atcccaactt ctcatttgaa 750

ggcaccaaac tggatgtggg gaacattgtg ctggcattat acagcggcct 800

ctttgcctat ggaggatgga attacttgaa tttcgtcaca gaggaaatga 850

tcaacccctа cagaaacctg cccctggcca tcatcatctc cctgcccatc 900

gtgacgctgg tgtacgtgct gaccaacctg gcctacttca ccaccctgtc 950

caccgagcag atgctgtcgt ccgaggccgt ggccgtggac ttcgggaact 1000

atcacctggg cgtcatgtcc tggatcatcc ccgtcttcgt gggcctgtcc 1050

tgcttcggct ccgtcaatgg gtccctgttc acatcctcca ggctcttctt 1100

cgtggggtcc cgggaaggcc acctgccctc catcctctcc atgatccacc 1150

cacagctcct cacccccgtg ccgtccctcg tgttcacgtg tgtgatgacg 1200

ctgctctacg ccttctccaa ggacatcttc tccgtcatca acttcttcag 1250

cttcttcaac tggctctgcg tggccctggc catcatcggc atgatctggc 1300

tgcgccacag aaagcctgag cttgagcggc ccatcaaggt gaacctggcc 1350

ctgcctgtgt tcttcatcct ggcctgcctc ttcctgatcg ccgtctcctt 1400

ctggaagaca cccgtggagt gtggcatcgg cttcaccatc atcctcagcg 1450

ggctgcccgt ctacttcttc ggggtctggt ggaaaaacaa gcccaagtgg 1500

ctcctccagg gcatcttctc cacgaccgtc ctgtgtcaga agctcatgca 1550

ggtggtcccc caggagacat agccaggagg ccgaagggcg 1590

<210> 38
<211> 2436
<212> DNA
<213> Homo sapiens

<400> 38
ccaggctctc cacccccact tcccaattga ggaaaccgag gcagaggagg 50

ctcaggtgtg gccaatcacc ctgcacatca gagttaccct gggcagggcc 100

cactgagacc tgggaggggc cactcgggac ctggagggct gggggctgcc 150

cgggcgttag gggtaaagct ccctacccaa ctgcgcagaa ggcctcagag 200

gcctgggggc tgggcttccc ctttcacatc gcccctttaga ggcccacgtg 250

tgggcattgg cccgcgatct gaaaggggct gtcctgttcc tcatgggcgc 300

tgccagcgcc acgcactcct ctttctgcct ggccggccac tcccgtctgc 350

tgtgacgcgc ggacagagag ctaccggtgg acccacggtg cctccctccc 400

tgggatctac acagaccatg gccttgccaa cggctcgacc cctgttgggg 450

tcctgtggga cccccgccct cggcagcctc ctgttcctgc tcttcagcct 500

cggatgggtg cagccctcga ggaccctggc tggagagaca gggcaggagg 550

ctgcacccct ggacggagtc ctggccaacc cacctaacat ttccagcctc 600

tcccctcgcc aactccttgg cttcccgtgt gcggaggtgt ccggcctgag 650

cacggagcgt gtccgggagc tggctgtggc cttggcacag aagaatgtca 700

agctctcaac agagcagctg cgctgtctgg ctcaccggct ctctgagccc 750

cccgaggacc tggacgccct cccattggac ctgctgctat tcctcaaccc 800

agatgcgttc tcggggcccc aggcctgcac ccgtttcttc tcccgcatca 850

cgaaggccaa tgtggacctg ctcccgaggg gggctcccga gcgacagcgg 900

ctgctgcctg cggctctggc ctgctggggt gtgcgggggt ctctgctgag 950

cgaggctgat gtgcgggctc tgggaggcct ggcttgcgac ctgcctgggc 1000

gctttgtggc cgagtcggcc gaagtgctgc taccccggct ggtgagctgc 1050

ccgggacccc tggaccagga ccagcaggag gcagccaggg cggctctgca 1100

gggcgggggа ccccctacg gccccccgtc gacatggtct gtctccacga 1150

tggacgctct gcggggcctg ctgcccgtgc tgggccagcc catcatccgc 1200

agcatcccgc agggcatcgt ggccgcgtgg cggcaacgct cctctcggga 1250

cccatcctgg cggcagcctg aacggaccat cctccggccg cggttccggc 1300

gggaagtgga gaagacagcc tgtccttcag gcaagaaggc ccgcgagata 1350

gacgagagcc tcatcttcta caagaagtgg gagctggaag cctgcgtgga 1400

tgcggccctg ctggccaccc agatggaccg cgtgaacgcc atccccttca 1450

cctacgagca gctggacgtc ctaaagcata aactggatga gctctaccca 1500

caaggttacc ccgagtctgt gatccagcac ctgggctacc tcttcctcaa 1550

gatgagccct gaggacattc gcaagtggaa tgtgacgtcc ctggagaccc 1600

tgaaggcttt gcttgaagtc aacaaagggc acgaaatgag tcctcaggtg 1650

gccaccctga tcgaccgctt tgtgaaggga aggggccagc tagacaaaga 1700

caccctagac accctgaccg ccttctaccc tgggtacctg tgctccctca 1750

gccccgagga gctgagctcc gtgcccccca gcagcatctg ggcggtcagg 1800

ccccaggacc tggacacgtg tgacccaagg cagctggacg tcctctatcc 1850

caaggcccgc cttgctttcc agaacatgaa cgggtccgaa tacttcgtga 1900

agatccagtc cttcctgggt ggggcccccа cggaggattt gaaggcgctc 1950

agtcagcaga atgtgagcat ggacttggcc acgttcatga agctgcggac 2000

ggatgcggtg ctgccgttga ctgtggctga ggtgcagaaa cttctgggac 2050

cccacgtgga gggcctgaag gcggaggagc ggcaccgccc ggtgcgggac 2100

tggatcctac ggcagcggca ggacgacctg gacacgctgg ggctggggct 2150

acagggcggc atccccaacg gctacctggt cctagacctc agcgtgcaag 2200

aggccctctc ggggacgccc tgcctcctag gacctggacc tgttctcacc 2250

gtcctggcac tgctcctagc ctccaccctg gcctgagggc cccactccct 2300

tgctggcccc agccctgctg gggatccccg cctggccagg agcaggcacg 2350

ggtgatcccc gttccacccc aagagaactc gcgctcagta aacgggaaca 2400

tgccccctgc agacacgtaa aaaaaaaaaa aaaaaa 2436

<210> 39
<211> 1979
<212> DNA
<213> Homo sapiens

<400> 39
tgggatctac acagaccatg gccttgccaa cggctcgacc cctgttgggg 50

tcctgtggga cccccgccct cggcagcctc ctgttcctgc tcttcagcct 100

cggatgggtg cagccctcga ggaccctggc tggagagaca gggcaggagg 150

ctgcgccccc ggacggagtc ctggccaacc cacctaacat ttccagcctc 200

tcccctcgcc aactccttgg cttcccgtgt gcggaggtgt ccggcctgag 250

cacggagcgt gtccgggagc tggctgtggc cttggcacag aagaatgtca 300

agctctcaac agagcagctg cgctgtctgg ctcaccggct ctctgagccc 350

cccgaggacc tggacgccct cccattggac ctgctgctat tcctcaaccc 400

agatgcgttc tcggggcccc aggcctgcac ccgtttcttc tcccgcatca 450

cgaaggccaa tgtggacctg ctcccgaggg gggctcccga gcgacagcgg 500

ctgctgcctg cggctctggc ctgctggggt gtgcgggggt ctctgctgag 550

cgaggctgat gtgcgggctc tgggaggcct ggcttgcgac ctgcctgggc 600

gctttgtggc cgagtcggcc gaagtgctgc taccccggct ggtgagctgc 650

ccgggacccc tggaccagga ccagcaggag gcagccaggg cggctctgca 700

gggcgggggа ccccctacg gccccccgtc gacatggtct gtctccacga 750

tggacgctct gcggggcctg ctgcccgtgc tgggccagcc catcatccgc 800

agcatcccgc agggcatcgt ggccgcgtgg cggcaacgct cctctcggga 850

cccatcctgg cggcagcctg aacggaccat cctccggccg cggttccggc 900

gggaagtgga gaagacagcc tgtccttcag gcaagaaggc ccgcgagata 950

gacgagagcc tcatcttcta caagaagtgg gagctggaag cctgcgtgga 1000

tgcggccctg ctggccaccc agatggaccg cgtgaacgcc atccccttca 1050

cctacgagca gctggacgtc ctaaagcata aactggatga gctctaccca 1100

caaggttacc ccgagtctgt gatccagcac ctgggctacc tcttcctcaa 1150

gatgagccct gaggacattc gcaagtggaa tgtgacgtcc ctggagaccc 1200

tgaaggcttt gcttgaagtc aacaaagggc acgaaatgag tcctcaggtg 1250

gccaccctga tcgaccgctt tgtgaaggga aggggccagc tagacaaaga 1300

caccctagac accctgaccg ccttctaccc tgggtacctg tgctccctca 1350

gccccgagga gctgagctcc gtgccccccа gcagcatctg ggcggtcagg 1400

ccccaggacc tggacacgtg tgacccaagg cagctggacg tcctctatcc 1450

caaggcccgc cttgctttcc agaacatgaa cgggtccgaa tacttcgtga 1500

agatccagtc cttcctgggt ggggcccccа cggaggattt gaaggcgctc 1550

agtcagcaga atgtgagcat ggacttggcc acgttcatga agctgcggac 1600

ggatgcggtg ctgccgttga ctgtggctga ggtgcagaaa cttctgggac 1650

cccacgtgga gggcctgaag gcggaggagc ggcaccgccc ggtgcgggac 1700

tggatcctac ggcagcggca ggacgacctg gacacgctgg ggctggggct 1750

acagggcggc atccccaacg gctacctggt cctagacctc agcgtgcaag 1800

aggccctctc ggggacgccc tgcctcctag gacctggacc tgttctcacc 1850

gtcctggcac tgctcctagc ctccaccctg gcctgagggc cccactccct 1900

tgctggcccc agccctgctg gggatccccg cctggccagg agcaggcacg 1950

ggtgatcccc gttccacccc aagagaact 1979

<210> 40
<211> 2456
<212> DNA
<213> Homo sapiens

<400> 40
ataacagcat gaagtgccgt ggaactggaa taggcgtgtc ctctccctcg 50

accctccccc tccttgtccc tctgctcacc cctcgctcgt tccctccctc 100

cggcgagggc cgcctttata acaactgctc agagtgcgag ggcgggatag 150

ctgtccaagg tctcccccag cactgaggag ctcgcctgct gccctcttgc 200

gcgcgggaag cagcaccaag ttcacggcca acgccttggc actagggtcc 250

agaatggcta caacagtccc tgatggttgc cgcaatggcc tgaaatccaa 300

gtactacaga ctttgtgata aggctgaagc ttggggcatc gtcctagaaa 350

cggtggccac agccggggtt gtgacctcgg tggccttcat gctcactctc 400

ccgatcctcg tctgcaaggt gcaggactcc aacaggcgaa aaatgctgcc 450

tactcagttt ctcttcctcc tgggtgtgtt gggcatcttt ggcctcacct 500

tcgccttcat catcggactg gacgggagca cagggcccac acgcttcttc 550

ctctttggga tcctcttttc catctgcttc tcctgcctgc tggctcatgc 600

tgtcagtctg accaagctcg tccgggggag gaagcccctt tccctgttgg 650

tgattctggg tctggccgtg ggcttcagcc tagtccagga tgttatcgct 700

attgaatata ttgtcctgac catgaatagg accaacgtca atgtcttttc 750

tgagctttcc gctcctcgtc gcaatgaaga ctttgtcctc ctgctcacct 800

acgtcctctt cttgatggcg ctgaccttcc tcatgtcctc cttcaccttc 850

tgtggttcct tcacgggctg gaagagacat ggggcccaca tctacctcac 900

gatgctcctc tccattgcca tctgggtggc ctggatcacc ctgctcatgc 950

ttcctgactt tgaccgcagg tgggatgaca ccatcctcag ctccgccttg 1000

gctgccaatg gctgggtgtt cctgttggct tatgttagtc ccgagttttg 1050

gctgctcaca aagcaacgaa accccatgga ttatcctgtt gaggatgctt 1100

tctgtaaacc tcaactcgtg aagaagagct atggtgtgga gaacagagcc 1150

tactctcaag aggaaatcac tcaaggtttt gaagagacag gggacacgct 1200

ctatgccccc tattccacac attttcagct gcagaaccag cctccccaaa 1250

aggaattctc catcccacgg gcccacgctt ggccgagccc ttacaaagac 1300

tatgaagtaa agaaagaggg cagctaactc tgtcctgaag agtgggacaa 1350

atgcagccgg gcggcagatc tagcgggagc tcaaagggat gtgggcgaaa 1400

tcttgagtct tctgagaaaa ctgtacaaga cactacggga acagtttgcc 1450

tccctcccag cctcaaccac aattcttcca tgctggggct gatgtgggct 1500

agtaagactc cagttcttag aggcgctgta gtatttttt ttttttgtct 1550

catcctttgg atacttcttt taagtgggag tctcaggcaa ctcaagttta 1600

gacccttact ctttttgttt gttttttgaa acaggatctt gctctgtcac 1650

ccaggcttga gtgcagtggt gcgatcacag cccagtgcag cctcgaccac 1700

ctgtgctcaa gcaatcctcc catctccatc tcccaaagtg ctgggatgac 1750

aggcgtgagc cacagctccc agcctaggcc cttaatcttg ctgttatttt 1800

ccatggacta aaggtctggt catctgagct cacgctggct cacacagctc 1850

taggggcctg ctcctctaac tcacagtggg ttttgtgagg ctctgtggcc 1900

cagagcagac ctgcatatct gagcaaaaat agcaaaagcc tctctcagcc 1950

cactggcctg aatctacact ggaagccaac ttgctggcac ccccgctccc 2000

caacccttct tgcctgggta ggagaggcta aagatcaccc taaatttact 2050

catctctcta gtgctgcctc acattgggcc tcagcagctc cccagcacca 2100

attcacaggt cacccctctc ttcttgcact gtccccaaac ttgctgtcaa 2150

ttccgagatc taatctcccc ctacgctctg ccaggaattc tttcagacct 2200

cactagcaca agcccggttg ctccttgtca ggagaatttg tagatcattc 2250

tcacttcaaa ttcctggggc tgatacttct ctcatcttgc accccaacct 2300

ctgtaaatag atttaccgca tttacggctg cattctgtaa gtgggcatgg 2350

tctcctaatg gaggagtgtt cattgtataa taagttattc acctgagtat 2400

gcaataaaga tgtggtggcc actctttcat ggtggtggca gcaaaaaaaa 2450

aaaaaa 2456

<210> 41
<211> 3213
<212> DNA
<213> Homo sapiens

<400> 41
agagactcaa gatgattccc tttttaccca tgttttctct actattgctg 50

cttattgtta accctataaa cgccaacaat cattatgaca agatcttggc 100

tcatagtcgt atcaggggtc gggaccaagg cccaaatgtc tgtgcccttc 150

aacagatttt gggcaccaaa aagaaatact tcagcacttg taagaactgg 200

tataaaaagt ccatctgtgg acagaaaacg actgttttat atgaatgttg 250

ccctggttat atgagaatgg aaggaatgaa aggctgccca gcagttttgc 300

ccattgacca tgtttatggc actctgggca tcgtgggagc caccacaacg 350

cagcgctatt ctgacgcctc aaaactgagg gaggagatcg agggaaaggg 400

atccttcact tactttgcac cgagtaatga ggcttgggac aacttggatt 450

ctgatatccg tagaggtttg gagagcaacg tgaatgttga attactgaat 500

gctttacata gtcacatgat taataagaga atgttgacca aggacttaaa 550

aaatggcatg attattcctt caatgtataa caatttgggg cttttcatta 600

accattatcc taatggggtt gtcactgtta attgtgctcg aatcatccat 650

gggaaccaga ttgcaacaaa tggtgttgtc catgtcattg accgtgtgct 700

tacacaaatt ggtacctcaa ttcaagactt cattgaagca gaagatgacc 750

tttcatcttt tagagcagct gccatcacat cggacatatt ggaggccctt 800

ggaagagacg gtcacttcac actctttgct cccaccaatg aggcttttga 850

gaaacttcca cgaggtgtcc tagaaaggtt catgggagac aaagtggctt 900

ccgaagctct tatgaagtac cacatcttaa atactctcca gtgttctgag 950

tctattatgg gaggagcagt ctttgagacg ctggaaggaa atacaattga 1000

gataggatgt gacggtgaca gtataacagt aaatggaatc aaaatggtga 1050

acaaaaagga tattgtgaca aataatggtg tgatccattt gattgatcag 1100

gtcctaattc ctgattctgc caaacaagtt attgagctgg ctggaaaaca 1150

gcaaaccacc ttcacggatc ttgtggccca attaggcttg gcatctgctc 1200

tgaggccaga tggagaatac actttgctgg cacctgtgaa taatgcattt 1250

tctgatgata ctctcagcat ggttcagcgc ctccttaaat taattctgca 1300

gaatcacata ttgaaagtaa aagttggcct taatgagctt tacaacgggc 1350

aaatactgga aaccatcgga ggcaaacagc tcagagtctt cgtatatcgt 1400

acagctgtct gcattgaaaa ttcatgcatg gagaaaggga gtaagcaagg 1450

gagaaacggt gcgattcaca tattccgcga gatcatcaag ccagcagaga 1500

aatccctcca tgaaaagtta aaacaagata agcgctttag caccttcctc 1550

agcctacttg aagctgcaga cttgaaagag ctcctgacac aacctggaga 1600

ctggacatta tttgtgccaa ccaatgatgc ttttaaggga atgactagtg 1650

aagaaaaaga aattctgata cgggacaaaa atgctcttca aaacatcatt 1700

ctttatcacc tgacaccagg agttttcatt ggaaaaggat ttgaacctgg 1750

tgttactaac attttaaaga ccacacaagg aagcaaaatc tttctgaaag 1800

aagtaaatga tacacttctg gtgaatgaat tgaaatcaaa agaatctgac 1850

atcatgacaa caaatggtgt aattcatgtt gtagataaac tcctctatcc 1900

agcagacaca cctgttggaa atgatcaact gctggaaata cttaataaat 1950

taatcaaata catccaaatt aagtttgttc gtggtagcac cttcaaagaa 2000

atccccgtga ctgtctatac aactaaaatt ataaccaaag ttgtggaacc 2050

aaaaattaaa gtgattgaag gcagtcttca gcctattatc aaaactgaag 2100

gacccacact aacaaaagtc aaaattgaag gtgaacctga attcagactg 2150

attaaagaag gtgaaacaat aactgaagtg atccatggag agccaattat 2200

taaaaaatac accaaaatca ttgatggagt gcctgtggaa ataactgaaa 2250

aagagacacg agaagaacga atcattacag gtcctgaaat aaaatacact 2300

aggatttcta ctggaggtgg agaaacagaa gaaactctga agaaattgtt 2350

acaagaagag gtcaccaagg tcaccaaatt cattgaaggt ggtgatggtc 2400

atttatttga agatgaagaa attaaaagac tgcttcaggg agacacaccc 2450

gtgaggaagt tgcaagccaa caaaaaagtt caaggttcta gaagacgatt 2500

aagggaaggt cgttctcagt gaaaatccaa aaaccagaaa aaaatgttta 2550

tacaacccta agtcaataac ctgaccttag aaaattgtga gagccaagtt 2600

gacttcagga actgaaacat cagcacaaag aagcaatcat caaataattc 2650

tgaacacaaa tttaatattt ttttttctga atgagaaaca tgagggaaat 2700

tgtggagtta gcctcctgtg gtaaaggaat tgaagaaaat ataacacctt 2750

acacccttt tcatcttgac attaaaagtt ctggctaact ttggaatcca 2800

ttagagaaaa atccttgtca ccagattcat tacaattcaa atcgaagagt 2850

tgtgaactgt tatcccattg aaaagaccga gccttgtatg tatgttatgg 2900

atacataaaa tgcacgcaag ccattatctc tccatgggaa gctaagttat 2950

aaaaataggt gcttggtgta caaaactttt tatatcaaaa ggctttgcac 3000

atttctatat gagtgggttt actggtaaat tatgttattt tttacaacta 3050

attttgtact ctcagaatgt ttgtcatatg cttcttgcaa tgcatatttt 3100

ttaatctcaa acgtttcaat aaaaccattt ttcagatata aagagaatta 3150

cttcaaattg agtaattcag aaaaactcaa gatttaagtt aaaaagtggt 3200

ttggacttgg gaa 3213

<210> 42
<211> 2450
<212> DNA
<213> Homo sapiens

<400> 42
ccettgagac tcaagatgat tccctttta cccatgtttt ctctactatt 50

gctgcttatt gttaacccta taaacgccaa caatcattat gacaagatct 100

tggctcatag tcgtatcagg ggtcgggacc aaggcccaaa tgtctgtgcc 150

cttcaacaga ttttgggcac caaaaagaaa tacttcagca cttgtaagaa 200

ctggtataaa aagtccatct gtggacagaa aacgactgtg ttatatgaat 250

gttgccctgg ttatatgaga atggaaggaa tgaaaggctg cccagcagtt 300

ttgcccattg accatgttta tggcactctg ggcatcgtgg gagccaccac 350

aacgcagcgc tattctgacg cctcaaaact gagggaggag atcgagggaa 400

agggatcctt cacttacttt gcaccgagta atgaggcttg ggacaacttg 450

gattctgata tccgtagagg tttggagagc aacgtgaatg ttgaattact 500

gaatgcttta catagtcaca tgattaataa gagaatgttg accaaggact 550

taaaaaatgg catgattatt ccttcaatgt ataacaattt ggggcttttc 600

attaaccatt atcctaatgg ggttgtcact gttaattgtg ctcgaatcat 650

ccatgggaac cagattgcaa caaatggtgt tgtccatgtc attgaccgtg 700

tgcttacaca aattggtacc tcaattcaag acttcattga agcagaagat 750

gacctttcat cttttagagc agctgccatc acatcggaca tattggaggc 800

ccttggaaga gacggtcact tcacactctt tgctcccacc aatgaggctt 850

ttgagaaact tccacgaggt gtcctagaaa ggatcatggg agacaaagtg 900

gcttccgaag ctcttatgaa gtaccacatc ttaaatactc tccagtgttc 950

tgagtctatt atgggaggag cagtctttga gacgctggaa ggaaatacaa 1000

ttgagatagg atgtgacggt gacagtataa cagtaaatgg aatcaaaatg 1050

gtgaacaaaa aggatattgt gacaaataat ggtgtgatcc atttgattga 1100

tcaggtccta attcctgatt ctgccaaaca agttattgag ctggctggaa 1150

aacagcaaac caccttcacg gatcttgtgg cccaattagg cttggcatct 1200

gctctgaggc cagatggaga atacactttg ctggcacctg tgaataatgc 1250

attttctgat gatactctca gcatggatca gcgcctcctt aaattaattc 1300

tgcagaatca catattgaaa gtaaaagttg gccttaatga gctttacaac 1350

gggcaaatac tggaaaccat cggaggcaaa cagctcagag tcttcgtata 1400

tcgtacagct gtctgcattg aaaattcatg catggagaaa gggagtaagc 1450

aagggagaaa cggtgcgatt cacatattcc gcgagatcat caagccagca 1500

gagaaatccc tccatgaaaa gttaaaacaa gataagcgct ttagcacctt 1550

cctcagccta cttgaagctg cagacttgaa agagctcctg acacaacctg 1600

gagactggac attatttgtg ccaaccaatg atgcttttaa gggaatgact 1650

agtgaagaaa aagaaattct gatacgggac aaaaatgctc ttcaaaacat 1700

cattctttat cacctgacac caggagtttt cattggaaaa ggatttgaac 1750

ctggtgttac taacatttta aagaccacac aaggaagcaa aatctttctg 1800

aaagaagtaa atgatacact tctggtgaat gaattgaaat caaaagaatc 1850

tgacatcatg acaacaaatg gtgtaattca tgttgtagat aaactcctct 1900

atccagcaga cacacctgtt ggaaatgatc aactgctgga aatacttaat 1950

aaattaatca aatacatcca aattaagttt gttcgtggta gcaccttcaa 2000

agaaatcccc gtgactgtct atagacccac actaacaaaa gtcaaaattg 2050

aaggtgaacc tgaattcaga ctgattaaag aaggtgaaac aataactgaa 2100

gtgatccatg gagagccaat tattaaaaaa tacaccaaaa tcattgatgg 2150

agtgcctgtg gaaataactg aaaaagagac acgagaagaa cgaatcatta 2200

caggtcctga aataaaatac actaggattt ctactggagg tggagaaaca 2250

gaagaaactc tgaagaaatt gttacaagaa gaggtcacca aggtcaccaa 2300

attcattgaa ggtggtgatg gtcatttatt tgaagatgaa gaaattaaaa 2350

gactgcttca gggagacaca cccgtgagga agttgcaagc caacaaaaaa 2400

gttcaaggat ctagaagacg attaagggaa ggtcgttctc agtgaagggc 2450

<210> 43
<211> 2276
<212> DNA
<213> Homo sapiens

<400> 43
cccttgagac tcaagatgat tccctttta cccatgtttt ctctactatt 50

gctgcttatt gttaacccta taaacgccaa caatcattat gacaagatct 100

tggctcatag tcgtatcagg ggtcgggacc aaggcccaaa tgtctgtgcc 150

cttcaacaga ttttgggcac caaaaagaaa tacttcagca cttgtaagaa 200

ctggtataaa aagtccatct gtggacagaa aacgactgtg ttatatgaat 250

gttgccctgg ttatatgaga atggaaggaa tgaaaggctg cccagcagtt 300

ttgcccattg accatgttta tggcactctg ggcatcgtgg gagccaccac 350

aacgcagcgc tattctgacg cctcaaaact gagggaggag atcgagggaa 400

agggatcctt cacttacttt gcaccgagta atgaggcttg ggacaacttg 450

gattctgata tccgtagagg tttggagagc aacgtgaatg ttgaattact 500

gaatgcttta catagtcaca tgattaataa gagaatgttg accaaggact 550

taaaaaatgg catgattatt ccttcaatgt ataacaattt ggggcttttc 600

attaaccatt atcctaatgg ggttgtcact gttaattgtg ctcgaatcat 650

ccatgggaac cagattgcaa caaatggtgt tgtccatgtc attgaccgtg 700

tgcttacaca aattggtacc tcaattcaag acttcattga agcagaagat 750

gacctttcat cttttagagc agctgccatc acatcggaca tattggaggc 800

ccttggaaga gacggtcact tcacactctt tgctcccacc aatgaggctt 850

ttgagaaact tccacgaggt gtcctagaaa ggatcatggg agacaaagtg 900

gcttccgaag ctcttatgaa gtaccacatc ttaaatactc tccagtgttc 950

tgagtctatt atgggaggag cagtctttga gacgctggaa ggaaatacaa 1000

ttgagatagg atgtgacggt gacagtataa cagtaaatgg aatcaaaatg 1050

gtgaacaaaa aggatattgt gacaaataat ggtgtgatcc atttgattga 1100

tcaggtccta attcctgatt ctgccaaaca agttattgag ctggctggaa 1150

aacagcaaac caccttcacg gatcttgtgg cccaattagg cttggcatct 1200

gctctgaggc cagatggaga atacactttg ctggcacctg tgaataatgc 1250

attttctgat gatactctca gcatggatca gcgcctcctt aaattaattc 1300

tgcagaatca catattgaaa gtaaaagttg gccttaatga gctttacaac 1350

gggcaaatac tggaaaccat cggaggcaaa cagctcagag tcttcgtata 1400

tcgtacagct gtctgcattg aaaattcatg catggagaaa gggagtaagc 1450

aagggagaaa cggtgcgatt cacatattcc gcgagatcat caagccagca 1500

gagaaatccc tccatgaaaa gttaaaacaa gataagcgct ttagcacctt 1550

cctcagccta cttgaagctg cagacttgaa agagctcctg acacaacctg 1600

gagactggac attatttgtg ccaaccaatg atgcttttaa gggaatgact 1650

agtgaagaaa aagaaattct gatacgggac aaaaatgctc ttcaaaacat 1700

cattctttat cacctgacac caggagtttt cattggaaaa ggatttgaac 1750

ctggtgttac taacatttta aagaccacac aaggaagcaa aatctttctg 1800

aaagaagtaa atgatacact tctggtgaat gaattgaaat caaaagaatc 1850

tgacatcatg acaacaaatg gtgtaattca tgttgtagat aaactcctct 1900

atccagcaga cacacctgtt ggaaatgatc aactgctgga aatacttaat 1950

aaattaatca aatacatcca aattaagttt gttcgtggta gcaccttcaa 2000

agaaatcccc gtgactgtct ataagccaat tattaaaaaa tacaccaaaa 2050

tcattgatgg agtgcctgtg gaaataactg aaaaagagac acgagaagaa 2100

cgaatcatta caggtcctga aataaaatac actaggattt ctactggagg 2150

tggagaaaca gaagaaactc tgaagaaatt gttacaagaa gacacacccg 2200

tgaggaagtt gcaagccaac aaaaaagttc aaggatctag aagacgatta 2250

agggaaggtc gttctcagtg aagggc 2276

<210> 44
<211> 2366
<212> DNA
<213> Homo sapiens

<400> 44
ccccttgagac tcaagatgat tccctttta cccatgtttt ctctactatt 50

gctgcttatt gttaacccta taaacgccaa caatcattat gacaagatct 100

tggctcatag tcgtatcagg ggtcgggacc aaggcccaaa tgtctgtgcc 150

cttcaacaga ttttgggcac caaaaagaaa tacttcagca cttgtaagaa 200

ctggtataaa aagtccatct gtggacagaa aacgactgtg ttatatgaat 250

gttgccctgg ttatatgaga atggaaggaa tgaaaggctg cccagcagtt 300

ttgcccattg accatgttta tggcactctg ggcatcgtgg gagccaccac 350

aacgcagcgc tattctgacg cctcaaaact gagggaggag atcgagggaa 400

agggatcctt cacttacttt gcaccgagta atgaggcttg ggacaacttg 450

gattctgata tccgtagagg tttggagagc aacgtgaatg ttgaattact 500

gaatgcttta catagtcaca tgattaataa gagaatgttg accaaggact 550

taaaaaatgg catgattatt ccttcaatgt ataacaattt ggggcttttc 600

attaaccatt atcctaatgg ggttgtcact gttaattgtg ctcgaatcat 650

ccatgggaac cagattgcaa caaatggtgt tgtccatgtc attgaccgtg 700

tgcttacaca aattggtacc tcaattcaag acttcattga agcagaagat 750

gacctttcat cttttagagc agctgccatc acatcggaca tattggaggc 800

ccttggaaga gacggtcact tcacactctt tgctcccacc aatgaggctt 850

ttgagaaact tccacgaggt gtcctagaaa ggatcatggg agacaaagtg 900

gcttccgaag ctcttatgaa gtaccacatc ttaaatactc tccagtgttc 950

tgagtctatt atgggaggag cagtctttga gacgctggaa ggaaatacaa 1000

ttgagatagg atgtgacggt gacagtataa cagtaaatgg aatcaaaatg 1050

gtgaacaaaa aggatattgt gacaaataat ggtgtgatcc atttgattga 1100

tcaggtccta attcctgatt ctgccaaaca agttattgag ctggctggaa 1150

aacagcaaac caccttcacg gatcttgtgg cccaattagg cttggcatct 1200

gctctgaggc cagatggaga atacactttg ctggcacctg tgaataatgc 1250

attttctgat gatactctca gcatggatca gcgcctcctt aaattaattc 1300

tgcagaatca catattgaaa gtaaaagttg gccttaatga gctttacaac 1350

gggcaaatac tggaaaccat cggaggcaaa cagctcagag tcttcgtata 1400

tcgtacagct gtctgcattg aaaattcatg catggagaaa gggagtaagc 1450

aagggagaaa cggtgcgatt cacatattcc gcgagatcat caagccagca 1500

gagaaatccc tccatgaaaa gttaaaacaa gataagcgct ttagcacctt 1550

cctcagccta cttgaagctg cagacttgaa agagctcctg acacaacctg 1600

gagactggac attatttgtg ccaaccaatg atgcttttaa gggaatgact 1650

agtgaagaaa aagaaattct gatacgggac aaaaatgctc ttcaaaacat 1700

cattctttat cacctgacac caggagtttt cattggaaaa ggatttgaac 1750

ctggtgttac taacatttta aagaccacac aaggaagcaa aatctttctg 1800

aaagaagtaa atgatacact tctggtgaat gaattgaaat caaaagaatc 1850

tgacatcatg acaacaaatg gtgtaattca tgttgtagat aaactcctct 1900

atccagcaga cacacctgtt ggaaatgatc aactgctgga aatacttaat 1950

aaattaatca aatacatcca aattaagttt gttcgtggta gcaccttcaa 2000

agaaatcccc gtgactgtct atagacccac actaacaaaa gtcaaaattg 2050

aaggtgaacc tgaattcaga ctgattaaag aaggtgaaac aataactgaa 2100

gtgatccatg gagagccaat tattaaaaaa tacaccaaaa tcattgatgg 2150

agtgcctgtg gaaataactg aaaaagagac acgagaagaa cgaatcatta 2200

caggtcctga aataaaatac actaggattt ctactggagg tggagaaaca 2250

gaagaaactc tgaagaaatt gttacaagaa gacacacccg tgaggaagtt 2300

gcaagccaac aaaaaagttc aaggatctag aagacgatta agggaaggtc 2350

gttctcagtg aagggc 2366

<210> 45
<211> 2360
<212> DNA
<213> Homo sapiens

<400> 45
ccctttgagac tcaagatgat tccctttttа cccatgtttt ctctactatt 50

gctgcttatt gttaacccta taaacgccaa caatcattat gacaagatct 100

tggctcatag tcgtatcagg ggtcgggacc aaggcccaaa tgtctgtgcc 150

cttcaacaga ttttgggcac caaaaagaaa tacttcagca cttgtaagaa 200

ctggtataaa aagtccatct gtggacagaa aacgactgtg ttatatgaat 250

gttgccctgg ttatatgaga atggaaggaa tgaaaggctg cccagcagtt 300

ttgcccattg accatgttta tggcactctg ggcatcgtgg gagccaccac 350

aacgcagcgc tattctgacg cctcaaaact gagggaggag atcgagggaa 400

agggatcctt cacttacttt gcaccgagta atgaggcttg ggacaacttg 450

gattctgata tccgtagagg tttggagagc aacgtgaatg ttgaattact 500

gaatgcttta catagtcaca tgattaataa gagaatgttg accaaggact 550

taaaaaatgg catgattatt ccttcaatgt ataacaattt ggggcttttc 600

attaaccatt atcctaatgg ggttgtcact gttaattgtg ctcgaatcat 650

ccatgggaac cagattgcaa caaatggtgt tgtccatgtc attgaccgtg 700

tgcttacaca aattggtacc tcaattcaag acttcattga agcagaagat 750

gacctttcat cttttagagc agctgccatc acatcggaca tattggaggc 800

ccttggaaga gacggtcact tcacactctt tgctcccacc aatgaggctt 850

ttgagaaact tccacgaggt gtcctagaaa ggatcatggg agacaaagtg 900

gcttccgaag ctcttatgaa gtaccacatc ttaaatactc tccagtgttc 950

tgagtctatt atgggaggag cagtctttga gacgctggaa ggaaatacaa 1000

ttgagatagg atgtgacggt gacagtataa cagtaaatgg aatcaaaatg 1050

gtgaacaaaa aggatattgt gacaaataat ggtgtgatcc atttgattga 1100

tcaggtccta attcctgatt ctgccaaaca agttattgag ctggctggaa 1150

aacagcaaac caccttcacg gatcttgtgg cccaattagg cttggcatct 1200

gctctgaggc cagatggaga atacactttg ctggcacctg tgaataatgc 1250

attttctgat gatactctca gcatggatca gcgcctcctt aaattaattc 1300

tgcagaatca catattgaaa gtaaaagttg gccttaatga gctttacaac 1350

gggcaaatac tggaaaccat cggaggcaaa cagctcagag tcttcgtata 1400

tcgtacagct gtctgcattg aaaattcatg catggagaaa gggagtaagc 1450

aagggagaaa cggtgcgatt cacatattcc gcgagatcat caagccagca 1500

gagaaatccc tccatgaaaa gttaaaacaa gataagcgct ttagcacctt 1550

cctcagccta cttgaagctg cagacttgaa agagctcctg acacaacctg 1600

gagactggac attatttgtg ccaaccaatg atgcttttaa gggaatgact 1650

agtgaagaaa aagaaattct gatacgggac aaaaatgctc ttcaaaacat 1700

cattctttat cacctgacac caggagtttt cattggaaaa ggatttgaac 1750

ctggtgttac taacatttta aagaccacac aaggaagcaa aatctttctg 1800

aaagaagtaa atgatacact tctggtgaat gaattgaaat caaaagaatc 1850

tgacatcatg acaacaaatg gtgtaattca tgttgtagat aaactcctct 1900

atccagcaga cacacctgtt ggaaatgatc aactgctgga aatacttaat 1950

aaattaatca aatacatcca aattaagttt gttcgtggta gcaccttcaa 2000

agaaatcccc gtgactgtct ataagccaat tattaaaaaa tacaccaaaa 2050

tcattgatgg agtgcctgtg gaaataactg aaaaagagac acgagaagaa 2100

cgaatcatta caggtcctga aataaaatac actaggattt ctactggagg 2150

tggagaaaca gaagaaactc tgaagaaatt gttacaagaa gaggtcacca 2200

aggtcaccaa attcattgaa ggtggtgatg gtcatttatt tgaagatgaa 2250

gaaattaaaa gactgcttca gggagacaca cccgtgagga agttgcaagc 2300

caacaaaaaa gttcaaggat ctagaagacg attaagggaa ggtcgttctc 2350

agtgaagggc 2360

<210> 46
<211> 2186
<212> DNA
<213> Homo sapiens

<400> 46
acacggacca aggagtctaa cacgtgcgcg agtcgggggc tcgcacgaaa 50

gccgccgtgg cgcaatgaag gtgaaggccg gcgcgctcgc cggccgaggt 100

gggatcccga ggcctctcca gtccgccgag ggcgcaccac cggcccgtct 150

cgcccgccgc gccggggagg tggagcacga gcgcacgtgt taggacccga 200

aagatggtga actatgcctg ggcagggcga agccagagga aactctggtg 250

gaggtccgta gcggtcctga cgtgcaaatc ggtcgtccga cctgggtata 300

ggggcgggct ccaggcgagg cggtcgacgc tcctgaaaac ttgcgcgcgc 350

gctcgcgcca ctgcgcccgg agcgatgaag atggtcgcgc cctggacgcg 400

gttctactcc aacagctgct gcttgtgctg ccatgtccgc accggcacca 450

tcctgctcgg cgtctggtat ctgatcatca atgctgtggt actgttgatt 500

ttattgagtg ccctggctga tccggatcag tataactttt caagttctga 550

actgggaggt gactttgagt tcatggatga tgccaacatg tgcattgcca 600

ttgcgatttc tcttctcatg atcctgatat gtgctatggc tacttacgga 650

gcgtacaagc aacgcgcagc ctggatcatc ccattcttct gttaccagat 700

ctttgacttt gccctgaaca tgttggttgc aatcactgtg cttatttatc 750

caaactccat tcaggaatac atacggcaac tgcctcctaa ttttccctac 800

agagatgatg tcatgtcagt gaatcctacc tgtttggtcc ttattattct 850

tctgtttatt agcattatct tgacttttaa gggttacttg attagctgtg 900

tttggaactg ctaccgatac atcaatggta ggaactcctc tgatgtcctg 950

gtttatgtta ccagcaatga cactacggtg ctgctacccc cgtatgatga 1000

tgccactgtg aatggtgctg ccaaggagcc accgccacct tacgtgtctg 1050

cctaagcctt caagtgggcg gagctgaggg cagcagcttg actttgcaga 1100

catctgagca atagttctgt tatttcactt ttgccatgag cctctctgag 1150

cttgtttgtt gctgaaatgc tacttttttaa aatttagatg ttagattgaa 1200

aactgtagtt ttcaacatat gctttgctag aacactgtga tagattaact 1250

gtagaattct tcctgtacga ttggggatat aatgggcttc actaaccttc 1300

cctaggcatt gaaacttccc ccaaatctga tggacctaga agtctgcttt 1350

tgtacctgct gggccccaaa gttgggcatt tttctctctg ttccctctct 1400

tttgaaaatg taaaataaaa ccaaaaatag acaacttttt cttcagccat 1450

tccagcatag agaacaaaac cttatggaaa caggaatgtc aattgtgtaa 1500

tcattgttct aattaggtaa atagaagtcc ttatgtatgt gttacaagaa 1550

tttcccccac aacatccttt atgactgaag ttcaatgaca gtttgtgttt 1600

gggtggtaaa ggattttctc catggcctga attaagacca ttagaaagca 1650

ccaggccgtg ggagcagtga ccatctgctg actgttcttg tggatcttgt 1700

gtccagggac atggggtgac atgcctcgta tgtgttagag ggtggaatgg 1750

atgtgtttgg cgctgcatgg gatctggtgc ccctcttctc ctggattcac 1800

atccccaccc agggcccgct tttactaagt gttctgccct agattggttc 1850

aaggaggtca tccaactgac tttatcaagt ggaattggga tatatttgat 1900

atacttctgc ctaacaacat ggaaaagggt tttcttttcc ctgcaagcta 1950

catcctactg ctttgaactt ccaagtatgt ctagtcacct tttaaaatgt 2000

aaacattttc agaaaaatga ggattgcctt ccttgtatgc gctttttacc 2050

ttgactacct gaattgcaag ggatttttat atattcatat gttacaaagt 2100

cagcaactct cctgttggtt cattattgaa tgtgctgtaa attaagttgt 2150

ttgcaattaa aacaaggttt gcccacaaaa aaaaaa 2186

<210> 47
<211> 851
<212> DNA
<213> Homo sapiens

<400> 47
attcgccctt ccgaaagatg gtgaactatg cctgggcagg gcgggctcca 50

ggcgaggcgg tcgacgctcc tgaaaacttg cgcgcgcgct cgcgccactg 100

cgcccggagc gatgaagatg gtcgcgccct ggactcggtt ctactccaac 150

agctgctgct tgtgctgcca tgtccgcacc ggcaccatcc tgctcggcgt 200

ctggtatctg atcatcaatg ctgtggtact gttgatttta ttgagtgccc 250

tggctgatcc ggatcagtat aacttttcaa gttctgaact gggaggtgac 300

tttgagttca tggatgatgc caacatgtgc attgccattg cgatttctct 350

tctcatgatc ctgatatgtg ctatggctac ttacggagcg tacaagcaac 400

gcgcagcctg gatcatccca ttcttctgtt accagatctt tgactttgcc 450

ctgaacatgt tggttgcaat cactgtgctt atttatccaa actccattca 500

ggaatacata cggcaactgc ctcctaattt tccctacaga gatgatgtca 550

tgtcagtgaa tcctacctgt ttggtcctta ttattcttct gtttattagc 600

attatcttga cttttaaggg ttacttgatt agctgtgttt ggaactgcta 650

ccgatacatc aatggtagga actcctctga tgtcctggtt tatgttacca 700

gcaatgacac tacggtgctg ctacccccgt atgatgatgc cactgtgaat 750

ggtgctgcca aggagccacc gccaccttac gtgtctgcct aagccttcaa 800

gtgggcggag ctgagggcag cagcttgact ttgcagacat ctgagaaggg 850

c 851

<210> 48
<211> 1728
<212> DNA
<213> Homo sapiens

<400> 48
agcgccatgg ggcgagccgg gggcgggggc ccgggccggg ggccgccgcc 50

actgctgctg tttctggggg ccgcgctggt cctggcctct ggggccgtgc 100

cggcgcgtga ggcgggcagc gcggttgagg ccgaagagct ggtgaagggc 150

agcccggcgt gggagccgcc tgccaacgac acgcgggaag aagccggccc 200

accagcggct ggggaagatg aggcgtcgtg gacggcgccc ggtggcgagc 250

tggccgggcc agaagaggtg ctgcaggagt cggctgcggt gaccggcacc 300

gcctggctgg aagctgacag cccaggcctg ggaggagtga ccgcagaggc 350

gggcagcggc gatgcccagg cccttccagc tacgctccag gctccccacg 400

aggtcctcgg gcagtcaatc atgccccctg ccattcctga ggctacagag 450

gccagcgggc caccctcccc caccccggc gacaagctga gcccagcttc 500

tgaactcccc aaggagagcc ccttggaggt ttggctgaac ctgggggggca 550

gcacacccga ccctcaagtg ccagagctga cttacccatt tcagggcacc 600

ctggagcccc aaccggcatc agatatcatt gacatcgact acttcgaagg 650

actggatggt gagggtcgtg gcgcagatct ggggagcttc ccagggtcac 700

caggaacctc agagaaccac cctgatactg agggagagac cccttcctgg 750

agcctgcttg acttatacga tgatttcacc cccttcgatg aatctgattt 800

ctaccccacc acatcctttt atgatgactt ggatgaagag gaggaggaag 850

aggaggatga caaagatgca gtaggaggtg gagacctaga agatgaaaat 900

gagcttctag tgcccactgg gaagcctggt ctggggcccg ggacaggcca 950

gcccaccagt cggtggcatg ctgtccctcc acagcacact ctggggtcgg 1000

tccccggcag cagcatcgcc ctcaggcccc gcccaggaga gccaggcagg 1050

gacttggcct ccagtgaaaa tggcactgag tgccgcagtg gctttgtgcg 1100

gcataacggc tcctgccggt cagtgtgcga cctcttccca agttactgtc 1150

acaatggcgg ccagtgctac ctggtggaga acataggggc cttctgcagg 1200

tgcaacacgc aggactacat ctggcacaag gggatgcgct gcgagtccat 1250

catcaccgac ttccaggtga tgtgcgtggc cgtgggctcg gctgccctcg 1300

tcctgctcct gctcttcatg atgacggtgt tctttgccaa gaagctctac 1350

ctgctcaaga cggagaatac caagctgcgt aggaccaaca aattccggac 1400

cccatctgag ctccacaatg ataacttctc cctctccacc attgccgagg 1450

gctctcaccc aaatgatgat cctagtgctc cccacaaaat ccaggaggtt 1500

ctcaagtcct gcctgaaaga ggaggagtca tttaacatcc agaactccat 1550

gtcgcccaaa cttgagggtg gcaaaggtga ccaggctgac ttggatgtga 1600

actgtcttca gaataattta acctaaagca gagcaagaag agaggaagcg 1650

gggtagtggg tgggggtagg gaagaaacat tatctcctct tgtacagagt 1700

ctatttcttg taaccatttg ttaaactc 1728

<210> 49
<211> 4231
<212> DNA
<213> Homo sapiens

<400> 49
ggacaggcgt ggcggccgga gccccagcat ccctgcttga ggtccaggag 50

cggagcccgc ggccaccgcc gcctgatcag cgcgaccccg gcccgcgccc 100

gccccgcccg gcaagatgct gcccgtgtac caggaggtga agcccaaccc 150

gctgcaggac gcgaacatct gctcacgcgt gttcttctgg tggctcaatc 200

ccttgtttaa aattggccat aaacggagat tagaggaaga tgatatgtat 250

tcagtgctgc cagaagaccg ctcacagcac cttggagagg agttgcaagg 300

gttctgggat aaagaagttt taagagctga gaatgacgca cagaagcctt 350

ctttaacaag agcaatcata aagtgttact ggaaatctta tttagttttg 400

ggaatttttа cgttaattga ggaaagtgcc aaagtaatcc agcccatatt 450

tttgggaaaa attattaatt attttgaaaa ttatgatccc atggattctg 500

tggctttgaa cacagcgtac gcctatgcca cggtgctgac tttttgcacg 550

ctcattttgg ctatactgca tcacttatat ttttatcacg ttcagtgtgc 600

tgggatgagg ttacgagtag ccatgtgcca tatgatttat cggaaggcac 650

ttcgtcttag taacatggcc atggggaaga caaccacagg ccagatagtc 700

aatctgctgt ccaatgatgt.gaacaagttt gatcaggtga cagtgttctt 750

acacttcctg tgggcaggac cactgcaggc gatcgcagtg actgccctac 800

tctggatgga gataggaata tcgtgccttg ctgggatggc agttctaatc 850

attctcctgc ccttgcaaag ctgttttggg aagttgttct catcactgag 900

gagtaaaact gcaactttca cggatgccag gatcaggacc atgaatgaag 950

ttataactgg tataaggata ataaaaatgt acgcctggga aaagtcattt 1000

tcaaatctta ttaccaattt gagaaagaag gagatttcca agattctgag 1050

aagttcctgc ctcaggggga tgaatttggc ttcgtttttc agtgcaagca 1100

aaatcatcgt gtttgtgacc ttcaccacct acgtgctcct cggcagtgtg 1150

atcacagcca gccgcgtgtt cgtggcagtg acgctgtatg gggctgtgcg 1200

gctgacggtt accctcttct tcccctcagc cattgagagg gtgtcagagg 1250

caatcgtcag catccgaaga atccagacct ttttgctact tgatgagata 1300

tcacagcgca accgtcagct gccgtcagat ggtaaaaaga tggtgcatgt 1350

gcaggatttt actgcttttt gggataaggc atcagagacc ccaactctac 1400

aaggcctttc ctttactgtc agacctggcg aattgttagc tgtggtcggc 1450

cccgtgggag cagggaagtc atcactgtta agtgccgtgc tcggggaatt 1500

ggccccaagt cacgggctgg tcagcgtgca tggaagaatt gcctatgtgt 1550

ctcagcagcc ctgggtgttc tcgggaactc tgaggagtaa tattttattt 1600

gggaagaaat atgaaaagga acgatatgaa aaagtcataa aggcttgtgc 1650

tctgaaaaag gatttacagc tgttggagga tggtgatctg actgtgatag 1700

gagatcgggg aaccacgctg agtggagggc agaaagcacg ggtaaacctt 1750

gcaagagcag tgtatcaaga tgctgacatc tatctcctgg acgatcctct 1800

cagtgcagta gatgcggaag ttagcagaca cttgttcgaa ctgtgtattt 1850

gtcaaatttt gcatgagaag atcacaattt tagtgactca tcagttgcag 1900

tacctcaaag ctgcaagtca gattctgata ttgaaagatg gtaaaatggt 1950

gcagaagggg acttacactg agttcctaaa atctggtata gattttggct 2000

ccctttaaa gaaggataat gaggaaagtg aacaacctcc agttccagga 2050

actcccacac taaggaatcg taccttctca gagtcttcgg tttggtctca 2100

acaatcttct agaccctcct tgaaagatgg tgctctggag agccaagata 2150

cagagaatgt cccagttaca ctatcagagg agaaccgttc tgaaggaaaa 2200

gttggttttc aggcctataa gaattacttc agagctggtg ctcactggat 2250

tgtcttcatt ttccttattc tcctaaacac tgcagctcag gttgcctatg 2300

tgcttcaaga ttggtggctt tcatactggg caaacaaaca aagtatgcta 2350

aatgtcactg taaatggagg aggaaatgta accgagaagc tagatcttaa 2400

ctggtactta ggaatttatt caggtttaac tgtagctacc gttctttttg 2450

gcatagcaag atctctattg gtattctacg tccttgttaa ctcttcacaa 2500

actttgcaca acaaaatgtt tgagtcaatt ctgaaagctc cggtattatt 2550

ctttgataga aatccaatag gaagaatttt aaatcgtttc tccaaagaca 2600

ttggacactt ggatgatttg ctgccgctga cgtttttaga tttcatccag 2650

acattgctac aagtggttgg tgtggtctct gtggctgtgg ccgtgattcc 2700

ttggatcgca ataccccttgg ttccccttgg aatcattttc atttttcttc 2750

ggcgatattt tttggaaacg tcaagagatg tgaagcgcct ggaatctaca 2800

actcggagtc cagtgttttc ccacttgtca tcttctctcc aggggctctg 2850

gaccatccgg gcatacaaag cagaagagag gtgtcaggaa ctgtttgatg 2900

cacaccagga tttacattca gaggcttggt tcttgttttt gacaacgtcc 2950

cgctggttcg ccgtccgtct ggatgccatc tgtgccatgt ttgtcatcat 3000

cgttgccttt gggtccctga ttctggcaaa aactctggat gccgggcagg 3050

ttggtttggc actgtcctat gccctcacgc tcatggggat gtttcagtgg 3100

tgtgttcgac aaagtgctga agttgagaat atgatgatct cagtagaaag 3150

ggtcattgaa tacacagacc ttgaaaaaga agcaccttgg gaatatcaga 3200

aacgcccacc accagcctgg ccccatgaag gagtgataat ctttgacaat 3250

gtgaacttca tgtacagtcc aggtgggcct ctggtactga agcatctgac 3300

agcactcatt aaatcacaag aaaaggttgg cattgtggga agaaccggag 3350

ctggaaaaag ttccctcatc tcagcccttt ttagattgtc agaacccgaa 3400

1140
1139

ggtaaaattt ggattgataa gatcttgaca actgaaattg gacttcacga 3450

tttaaggaag aaaatgtcaa tcatacctca ggaacctgtt ttgttcactg 3500

gaacaatgag gaaaaacctg gatcccttta aggagcacac ggatgaggaa 3550

ctgtggaatg ccttacaaga ggtacaactt aaagaaacca ttgaagatct 3600

tcctggtaaa atggatactg aattagcaga atcaggatcc aattttagtg 3650

ttggacaaag acaactggtg tgccttgcca gggcaattct caggaaaaat 3700

cagatattga ttattgatga agcgacggca aatgtggatc caagaactga 3750

tgagttaata caaaaaaaaa tccgggagaa atttgcccac tgcaccgtgc 3800

taaccattgc acacagattg aacaccatta ttgacagcga caagataatg 3850

gttttagatt caggaagact gaaagaatat gatgagccgt atgttttgct 3900

gcaaaataaa gagagcctat tttacaagat ggtgcaacaa ctgggcaagg 3950

cagaagccgc tgccctcact gaaacagcaa aacaggtata cttcaaaaga 4000

aattatccac atattggtca cactgaccac atggttacaa acacttccaa 4050

tggacagccc tcgaccttaa ctattttcga gacagcactg tgaatccaac 4100

caaaatgtca agtccgttcc gaaggcattt tccactagtt tttggactat 4150

gtaaaccaca ttgtactttt ttttactttg gcaacaaata tttatacata 4200

caagatgcta gttcatttga atatttctcc c 4231

<210> 50
<211> 2861
<212> DNA
<213> Homo sapiens

<400> 50
ctccgtccca ggggagaagg agaggcgtct gcaggggggca gagaccgcag 50

ctacctgccg ggtgcgcccc ccacccagga gcgctcgctt cgcccccttt 100

cctccccccg cccccacctc cttattggtg ctagtttgca gcgcccagct 150

cctgcgcctt cgcttcgcgt ttgaatctgg ctcgcccctt cgtattatgt 200

ctgcactccg aaggaaattt ggggacgatt atcaggtagt gaccacatcg 250

tccagcggct cgggcttgca gccccagggg ccaggccagg accctcagca 300

gcagcttgtg cccaagaaga agcggcagcg gttcgtggac aagaacggcc 350

ggtgcaatgt acagcacggc aacctgggca gcgagacaag ccgctacctc 400

tcggacctct tcaccacgct ggtggacctc aagtggcgct ggaacctctt 450

catcttcatt ctcacctaca ccgtggcctg gcttttcatg gcgtccatgt 500

ggtgggtgat cgcctacact cggggcgacc tgaacaaagc ccacgtcggt 550

aactacacgc cttgcgtggc caatgtctat aacttccctt ctgccttcct 600

cttcttcatc gagacggagg ccaccatcgg ctatggctac cgatacatca 650

cagacaagtg ccccgagggc atcatcctct tcctcttcca gtccatcctg 700

ggctccatcg tggacgcctt cctcatcggc tgcatgttca tcaagatgtc 750

ccagcccaag aagcgcgccg agaccctcat gttcagcgag cacgcggtga 800

tctccatgag ggacggaaaa ctcacgctta tgttccgggt gggcaacctg 850

cgcaacagcc acatggtctc cgcgcagatt cgctgcaagc tgctcaaatc 900

tcggcagaca cctgagggtg agttccttcc ccttgaccaa cttgaactgg 950

atgtaggttt tagtacaggg gcagatcaac tttttcttgt gtccccctc 1000

acaatttgcc acgtgatcga tgccaaaagc ccctttttatg acctatccca 1050

gcgaagcatg caaactgaac agttcgagat tgtcgtcatc ctagaaggca 1100

ttgtggaaac aactgggatg acttgtcaag ctcgaacatc atatactgaa 1150

gatgaagttc tttggggtca tcgtttttt cctgtaattt ccttagaaga 1200

gggattcttt aaagttgatt actcccagtt ccacgcaaca tttgaagtcc 1250

ccaccccacc ttacagtgtg aaagagcagg aggaaatgct tctcatgtcg 1300

tcccctttaa tagcaccagc cataactaac agcaaagaaa gacataattc 1350

tgtggaatgc ttagatggac tagatgatat tactacaaaa ctaccatcta 1400

agctgcagaa aattactgga agagaagact ttcccaaaaa actcttgagg 1450

atgagttcta caacttcaga aaaagcctac agcttgggag acttgcccat 1500

gaaacttcaa cgaataagtt cagttccggg caactcagaa gaaaaactgg 1550

tatctaaaac caccaagatg ttatctgatc ccatgagcca gtctgtggct 1600

gatttgccac caaagcttca aaagatggct ggaggagcag ctaggatgga 1650

agggaacctt ccagccaaat taagaaaaat gaactctgat cgcttcacat 1700

aacaaagcac tcccttaggc attatttaat gtttgattta gtaatagtcc 1750

aatatttggc gatgaggtaa ttctccctaa ggaatctgaa agtatatttt 1800

cctcccagtt ctacagcata tttcgagaac ccttcctttc ccaagtattg 1850

cgaatgtgca gaaagcaaca gttacggagg gaggacatca taaggaagtt 1900

attaacgggc atgtattatc acatcaagca tgcaataatg tgcaaatttt 1950

gcatttagtt ttatggcatg atttatatat ggcatattta tattgtatat 2000

tctggaaaaa aaaatatata tatatattta aaggggagat actctccctg 2050

acatttctaa catatgtatt aagccaaaca tgagtgaata gctttcaggg 2100

cgataaaact aaatatatgt ctgtgtgtgt gtgtgtatgt atacacacat 2150

atacatatat atatacacat acatacacat acatacatac atacatatat 2200

atctgataaa attgtgatgt tttgttcaaa gttgtagttc ttgtgcatgt 2250

ttactttatt agagtaggaa ggctactggc attaattatt aataccaaat 2300

attttagcct taaatttttg tcattttaaa atctgattta atgttttctg 2350

ctgtttaagg tcttgggagg ctttcaattg tattttatat gagagaatca 2400

cacaagtttg tgctatctat ggccctgcaa aaatataacc attacatgtt 2450

taaattgtaa attttagagc ataccagtac tcagtatagc attgaacatt 2500

tcttatgatt tttaaaagtt gctagtactg gggagaaata attgttgatt 2550

aatttgagaa ttattccttt cctagactaa ttaaaatctg gaaatctgtt 2600

ttgtatatga tctaatacaa agatgagctc tgaacaaaca ctgaatcatg 2650

ttaatagaca gtagccaagt tatattgaat atatcagaat ctgtgtgaag 2700

ttacacaatt aattgtccct gtttcaaact gagtaaattg gaaacatttt 2750

ctttcttttt ctggaaattt tgtccatttt aaaaaccaat cattttaaga 2800

agacatgaca atgcaatgaa acagatgata aatatttatg cttaaaataa 2850

aaaaaaaaaa a 2861

<210> 51
<211> 3557
<212> DNA
<213> Homo sapiens

<400> 51
gagagggtcc ttcagggtct gcttatgccc ttgttcaaga acaccagtgt 50

cagctctctg tactctggtt gcagactgac cttgctcagg cctgagaagg 100

atggggcagc caccagagtg gatgctgtct gcacccatcg tcctgacccc 150

aaaagccctg gactggacag agagcggctg tactggaagc tgagccagct 200

gacccacggc atcactgagc tgggccccta caccctggac aggcacagtc 250

tctatgtcaa tggtttcacc catcagagct ctatgacgac caccagaact 300

cctgatacct ccacaatgca cctggcaacc tcgagaactc cagcctccct 350

gtctggacct acgaccgcca gccctctcct ggtgctattc acaattaact 400

tcaccatcac taacctgcgg tatgaggaga acatgcatca ccctggctct 450

agaaagttta acaccacgga gagagtcctt cagggtctgc tcaggcctgt 500

gttcaagaac accagtgttg gccctctgta ctctggctgc agactgacct 550

tgctcaggcc caagaaggat ggggcagcca ccaaagtgga tgccatctgc 600

acctaccgcc ctgatcccaa aagccctgga ctggacagag agcagctata 650

ctgggagctg agccagctaa cccacagcat cactgagctg ggcccctaca 700

ccctggacag ggacagtctc tatgtcaatg gtttcacaca gcggagctct 750

gtgcccacca ctagcattcc tgggaccccc acagtggacc tgggaacatc 800

tgggactcca gtttctaaac ctggtccctc ggctgccagc cctctcctgg 850

tgctattcac tctcaacttc accatcacca acctgcggta tgaggagaac 900

atgcagcacc ctggctccag gaagttcaac accacggaga gggtccttca 950

gggcctgctc aggtccctgt tcaagagcac cagtgttggc cctctgtact 1000

ctggctgcag actgactttg ctcaggcctg aaaaggatgg gacagccact 1050

ggagtggatg ccatctgcac ccaccaccct gaccccaaaa gccctaggct 1100

ggacagagag cagctgtatt gggagctgag ccagctgacc cacaatatca 1150

ctgagctggg ccactatgcc ctggacaacg acagcctctt tgtcaatggt 1200

ttcactcatc ggagctctgt gtccaccacc agcactcctg ggacccccac 1250

agtgtatctg ggagcatcta agactccagc ctcgatattt ggcccttcag 1300

ctgccagcca tctcctgata ctattcaccc tcaacttcac catcactaac 1350

ctgcggtatg aggagaacat gtggcctggc tccaggaagt tcaacactac 1400

agagagggtc cttcagggcc tgctaaggcc cttgttcaag aacaccagtg 1450

ttggccctct gtactctggc tccaggctga ccttgctcag gccagagaaa 1500

gatggggaag ccaccggagt ggatgccatc tgcacccacc gccctgaccc 1550

cacaggccct gggctggaca gagagcagct gtatttggag ctgagccagc 1600

tgacccacag catcactgag ctgggccccc acacactgga cagggacagt 1650

ctctatgtca atggtttcac ccatcggagc tctgtaccca ccaccagcac 1700

cggggtggtc agcgaggagc cattcacact gaacttcacc atcaacaacc 1750

tgcgctacat ggcggacatg ggccaacccg gctccctcaa gttcaacatc 1800

acagacaacg tcatgaagca cctgctcagt cctttgttcc agaggagcag 1850

cctgggtgca cggtacacag gctgcagggt catcgcacta aggtctgtga 1900

agaacggtgc tgagacacgg gtggacctcc tctgcaccta cctgcagccc 1950

ctcagcggcc caggtctgcc tatcaagcag gtgttccatg agctgagcca 2000

gcagacccat ggcatcaccc ggctgggccc ctactctctg gacaaagaca 2050

gcctctacct taacggttac aatgaacctg gtctagatga gcctcctaca 2100

actcccaagc cagccaccac attcctgcct cctctgtcag aagccacaac 2150

agccatgggg taccacctga agaccctcac actcaacttc accatctcca 2200

atctccagta ttcaccagat atgggcaagg gctcagctac attcaactcc 2250

accgaggggg tccttcagca cctgctcaga cccttgttcc agaagagcag 2300

catgggcccc ttctacttgg gttgccaact gatctccctc aggcctgaga 2350

aggatggggc agccactggt gtggacacca cctgcaccta ccaccctgac 2400

cctgtgggcc ccgggctgga catacagcag ctttactggg agctgagtca 2450

gctgacccat ggtgtcaccc aactgggctt ctatgtcctg gacagggata 2500

gcctcttcat caatggctat gcaccccaga atttatcaat ccggggcgag 2550

taccagataa atttccacat tgtcaactgg aacctcagta atccagaccc 2600

cacatcctca gagtacatca ccctgctgag ggacatccag gacaaggtca 2650

ccacactcta caaaggcagt caactacatg acacattccg cttctgcctg 2700

gtcaccaact tgacgatgga ctccgtgttg gtcactgtca aggcattgtt 2750

ctcctccaat ttggacccca gcctggtgga gcaagtcttt ctagataaga 2800

ccctgaatgc ctcattccat tggctgggct ccacctacca gttggtggac 2850

atccatgtga cagaaatgga gtcatcagtt tatcaaccaa caagcagctc 2900

cagcacccag cacttctacc cgaatttcac catcaccaac ctaccatatt 2950

cccaggacaa agcccagcca ggcaccacca attaccagag gaacaaaagg 3000

aatattgagg atgcgctcaa ccaactcttc cgaaacagca gcatcaagag 3050

ttatttttct gactgtcaag tttcaacatt caggtctgtc cccaacaggc 3100

accacaccgg ggtggactcc ctgtgtaact tctcgccact ggctcggaga 3150

gtagacagag ttgccatcta tgaggaattt ctgcggatga cccggaatgg 3200

tacccagctg cagaacttca ccctggacag gagcagtgtc cttgtggatg 3250

ggtattctcc caacagaaat gagcccttaa ctgggaattc tgaccttccc 3300

ttctgggctg tcatcttcat cggcttggca ggactcctgg gactcatcac 3350

atgcctgatc tgcggtgtcc tggtgaccac ccgccggcgg aagaaggaag 3400

gagaatacaa cgtccagcaa cagtgcccag gctactacca gtcacaccta 3450

gacctggagg atctgcaatg actggaactt gccggtgcct ggggtgcctt 3500

tcccccagcc agggtccaaa gaagcttggc tggggcagaa ataaaccata 3550

ttggtcg 3557

<210> 52
<211> 3866
<212> DNA
<213> Homo sapiens

<400> 52
gcccgggcag gtgctgatag cacagttctg tccagagaag gaaggcggaa 50

taaacttatt cattcccagg aactcttggg gtaggtgtgt gtttttcaca 100

tcttaaaggc tcacagaccc tgcgctggac aaatgttcca ttcctgaagg 150

acctctccag aatccggatt gctgaatctt ccctgttgcc tagaagggct 200

ccaaaccacc tcttgacaat gggaaactgg gtggttaacc actggttttc 250

agttttgttt ctggttgttt ggttagggct gaatgttttc ctgtttgtgg 300

atgccttcct gaaatatgag aaggccgaca aatactacta cacaagaaaa 350

atccttgggt caacattggc ctgtgcccga gcgtctgctc tctgcttgaa 400

ttttaacagc acgctgatcc tgcttcctgt gtgtcgcaat ctgctgtcct 450

tcctgagggg cacctgctca ttttgcagcc gcacactgag aaagcaattg 500

gatcacaacc tcaccttcca caagctggtg gcctatatga tctgcctaca 550

tacagctatt cacatcattg cacacctgtt taactttgac tgctatagca 600

gaagccgaca ggccacagat ggctcccttg cctccattct ctccagccta 650

tctcatgatg agaaaaaggg gggttcttgg ctaaatccca tccagtcccg 700

aaacacgaca gtggagtatg tgacattcac cagcgttgct ggtctcactg 750

gagtgatcat gacaatagcc ttgattctca tggtaacttc agctactgag 800

ttcatccgga ggagttattt tgaagtcttc tggtatactc accacctttt 850

tatcttctat atccttggct tagggattca cggcattggt ggaattgtcc 900

ggggtcaaac agaggagagc atgaatgaga gtcatcctcg caagtgtgca 950

gagtcttttg agatgtggga tgatcgtgac tcccactgta ggcgccctaa 1000

gtttgaaggg catccccctg agtcttggaa gtggatcctt gcaccggtca 1050

ttctttatat ctgtgaaagg atcctccggt tttaccgctc ccagcagaag 1100

gttgtgatta ccaaggttgt tatgcaccca tccaaagttt tggaattgca 1150

gatgaacaag cgtggcttca gcatggaagt ggggcagtat atctttgtta 1200

attgcccctc aatctctctc ctggaatggc atcctttttac tttgacctct 1250

gctccagagg aagatttctt ctccattcat atccgagcag caggggactg 1300

gacagaaaat ctcataaggg ctttcgaaca acaatattca ccaattccca 1350

ggattgaagt ggatggtccc tttggcacag ccagtgagga tgttttccag 1400

tatgaagtgg ctgtgctggt tggagcagga attggggtca ccccctttgc 1450

ttctatcttg aaatccatct ggtacaaatt ccagtgtgca gaccacaacc 1500

tcaaaacaaa aaagatctat ttctactgga tctgcaggga gacaggtgcc 1550

ttttcctggt tcaacaacct gttgacttcc ctggaacagg agatggagga 1600

attaggcaaa gtgggttttc taaactaccg tctcttcctc accggatggg 1650

acagcaatat tgttggtcat gcagcattaa actttgacaa ggccactgac 1700

atcgtgacag gtctgaaaca gaaaacctcc tttgggagac caatgtggga 1750

caatgagttt tctacaatag ctacctccca ccccaagtct gtagtgggag 1800

ttttcttatg tggccctcgg actttggcaa agagcctgcg caaatgctgt 1850

caccgatatt ccagtctgga tcctagaaag gttcaattct acttcaacaa 1900

agaaaatttt tgagttatag gaataaggac ggtaatctgc attttgtctc 1950

tttgtatctt cagtaattga gttataggaa taaggacggt aatctgcatt 2000

ttgtctcttt gtatcttcag taatttactt ggtctcgtca ggtttgagca 2050

gtcactttag gataagaatg tgcctctcaa gccttgactc cctggtattc 2100

ttttttgat tgcattcaac ttcgttactt gagcttcagc aacttaagaa 2150

cttctgaagt tcttaaagtt ctgaagttct taaagcccat ggatcctttc 2200

tcagaaaaat aactgtaaat ctttctggac agccatgact gtagcaaggc 2250

ttgatagcag aggtttggtg gttcagagtt atacaactaa tcccaggtga 2300

ttttatcaat tccagtgtta ccatctcctg agttttggtt tgtaatcttt 2350

tgtccctccc acccccacag aagatttcta agtagggtga ctttttaaat 2400

aaaaatttat tgaataatta atgataaaac ataataataa acataaataa 2450

taaacaaaat taccgagaac cccatcccca tataacacca acagtgtaca 2500

tgtttactgt cacttttgat atggtcttat ccagtgtgaa cagcaattta 2550

ttatttttgc tcatcaaaaa ataaaggatt ttcttcttca cttgatgaat 2600

gagtcttttg ttacttcttt ttgggctctc catttgggaa agaaaatcta 2650

gaaagtatgc tcattgtaat tgaaattgta ttataattgc aagtcagcta 2700

tactgtcatc ttgtattatt ctgttaaaag cgattgtgaa aacttgaact 2750

ccatgttagg atgacttatt tagtcaagag gagctcttta tcactattaa 2800

gggagtttat taatgtctag tttaaaaaat aacaatttat aattttcata 2850

gatttttgtg gactcctacc tatatcgcta aggaatcatt cagctaataa 2900

cataagcaaa atgccttcat tctctatcag aatgcttttt atttctgcca 2950

tctacccaac ctcggcctac aaggaccaca ctggctctgg gatgttagtc 3000

gtcacagtga atccctagaa ggagagaagc ccagggctct taagagagaa 3050

acctcacact tcattgtgcc cagacaattc aactcattat agattagaag 3100

aatggttgta ctgaattagc attatattat ctatttttc taccattaag 3150

tgaccacgaa tgaacaaatg gttgtttcta atagtttagg aacctctacg 3200

gcttgaggga taaaactgtt tttggttttt ttttgtcact agtctgagca 3250

caataaaggt aggtcatgaa ctcatggtgg ctaatcatgt ttcatttatt 3300

acaacagtaa aacaatactg atggttgttt acatagggag ggataccact 3350

gctgttgcca atttgttcac ccggatttca gctaattcag gtttaataag 3400

gctttgggaa aatgcacaat gaaaccttga agacatgatt cttattagac 3450

aactactgac caaatgacaa cttcatcttc aatgaccttg atttggaatc 3500

ttgaaactta caacgagttc ctaatgagaa aggcatcact cttggtctcc 3550

atgtctaaat atattgacac actcacttca agatatgatg gctgccaaaa 3600

tcaatatccc aaggcttgaa tatttatggg agtaggcata aattcgtgat 3650

tatattcatg gtaatgaata taataaatcc tatgcatagc caacatccaa 3700

gttgaactgg ggatgagcag gatacacaga aagctcctca cctgctttga 3750

gcaagttttt cttgcaaagt tgtaatatga aataccttaa tcaacaaatt 3800

tcttataaat gtttatatga gtagaaaatc cctcaatttc attgtcactc 3850

aaaaagttga ttatct 3866

<210> 53
<211> 1817
<212> DNA
<213> Homo sapiens

<400> 53
cccttgggct ccaaaccacc tcttgacaat gggaaactgg gtggttaacc 50

actggttttc agttttgttt ctggttgttt ggttagggct gaatgttttc 100

ctgtttgtgg atgccttcct gaaatatgag aaggccgaca aatactacta 150

cacaagaaaa atccttgggt caacattggc ctgtgcccga gcgtctgctc 200

tctgcttgaa ttttaacagc acgctgatcc tgcttcctgt gtgtcgcaat 250

ctgctgtcct tcctgagggg cacctgctca ttttgcagcc gcacactgag 300

aaagcaattg gatcacaacc tcaccttcca caagctggtg gcctatatga 350

tctgcctaca tacagctatt cacatcattg cacacctgtt taactttgac 400

tgctatagca gaagccgaca ggccacagat ggctcccttg cctccattct 450

ctccagccta tctcatgatg agaaaaaggg gggttcttgg ctaaatccca 500

tccagtcccg aaacacgaca gtggagtatg tgacattcac cagcattgct 550

ggtctcactg gagtgatcat gacaatagcc ttgattctca tggtaacttc 600

agctactgag ttcatccgga ggagttattt tgaagtcttc tggtatactc 650

accacctttt tatcttctat atccttggct tagggattca cggcattggt 700

ggaattgtcc ggggtcaaac agaggagagc atgaatgaga gtcatcctcg 750

caagtgtgca gagtcttttg agatgtggga tgatcgtgac tcccactgta 800

ggcgccctaa gtttgaaggg catccccctg agtcttggaa gtggatcctt 850

gcaccggtca ttctttatat ctgtgaaagg atcctccggt tttaccgctc 900

ccagcagaag gttgtgatta ccaaggttgt tatgcaccca tccaaagttt 950

tggaattgca gatgaacaag cgtggcttca gcatggaagt ggggcagtat 1000

atctttgtta attgcccctc aatctctctc ctggaatggc atccttttac 1050

tttgacctct gctccagagg aagatttctt ctccattcat atccgagcag 1100

caggggactg gacagaaaat ctcataaggg ctttcgaaca acaatattca 1150

ccaattccca ggattgaagt ggatggtccc tttggcacag ccagtgagga 1200

tgttttccag tatgaagtgg ctgtgctggt tggagcagga attggggtca 1250

ccccctttgc ttctatcttg aaatccatct ggtacaaatt ccagtgtgca 1300

gaccacaacc tcaaaacaaa aaagatctat ttctactgga tctgcaggga 1350

gacaggtgcc ttttcctggt tcaacaacct gttgacttcc ctggaacagg 1400

agatggagga attaggcaaa gtgggttttc taaactaccg tctcttcctc 1450

accggatggg acagcaatat tgttggtcat gcagcattaa actttgacaa 1500

ggccactgac atcgtgacag gtctgaaaca gaaaacctcc tttgggagac 1550

caatgtggga caatgagttt tctacaatag ctacctccca ccccaagtct 1600

gtagtgggag ttttcttatg tggccctcgg actttggcaa agagcctgcg 1650

caaatgctgt caccgatatt ccagtctgga tcctagaaag gttcaattct 1700

acttcaacaa agaaaatttt tgagttatag gaataaggac ggtaatctgc 1750

attttgtctc tttgtatctt cagtaattta cttggtctcg tcaggtttga 1800

gcagtcactt taggaag 1817

<210> 54
<211> 1000
<212> DNA
<213> Homo sapiens

<400> 54
gggctccaaa ccacctcttg acaatgggaa actgggtggt taaccactgg 50

ttttcagttt tgtttctggt tgtttggtta gggctgaatg ttttcctgtt 100

tgtggatgcc ttcctgaaat atgagaaggc cgacaaatac tactacacaa 150

gaaaaatcct tgggtcttgg aagtggatcc ttgcaccggt cattctttat 200

atctgtgaaa ggatcctccg gttttaccgc tcccagcaga aggttgtgat 250

taccaaggtt gttatgcacc catccaaagt tttggaattg cagatgaaca 300

agcgtggctt cagcatggaa gtggggcagt atatctttgt taattgcccc 350

tcaatctctc tcctggaatg gcatcctttt actttgacct ctgctccaga 400

ggaagatttc ttctccattc atatccgagc agcaggggac tggacagaaa 450

atctcataag ggctttcgaa caacaatatt caccaattcc caggattgaa 500

gtggatggtc cctttggcac agccagtgag gatgttttcc agtatgaagt 550

ggctgtgctg gttggagcag gaattggggt cacccccttt gcttctatct 600

tgaaatccat ctggtacaaa ttccagtgtg cagaccacaa cctcaaaaca 650

aaaaaggttg gtcatgcagc attaaacttt gacaaggcca ctgacatcgt 700

gacaggtctg aaacagaaaa cctcctttgg gagaccaatg tgggacaatg 750

agttttctac aatagctacc tcccacccca agtctgtagt gggagttttc 800

ttatgtggcc ctcggacttt ggcaaagagc ctgcgcaaat gctgtcaccg 850

atattccagt ctggatccta gaaaggttca attctacttc aacaaagaaa 900

atttttgagt tataggaata aggacggtaa tctgcatttt gtctctttgt 950

atcttcagta atttacttgg tctcgtcagg tttgagcagt cactttagga 1000

<210> 55
<211> 1665
<212> DNA
<213> Homo sapiens

<400> 55
gggctccaaa ccacctcttg acaatgggaa actgggtggt taaccactgg 50

ttttcagttt tgtttctggt tgtttggtta gggctgaatg ttttcctgtt 100

tgtggatgcc ttcctgaaat atgagaaggc cgacaaatac tactacacaa 150

gaaaaatcct tgggtcaaca ttggcctgtg cccgagcgtc tgctctctgc 200

ttgaatttta acagcacgct gatcctgctt cctgtgtgtc gcaatctgct 250

gtccttcctg aggggcacct gctcattttg cagccgcaca ctgagaaagc 300

aattggatca caacctcacc ttccacaagc tggtggccta tatgatctgc 350

ctacatacag ctattcacat cattgcacac ctgtttaact ttgactgcta 400

tagcagaagc cgacaggcca cagatggctc ccttgcctcc attctctcca 450

gcctatctca tgatgagaaa aaggggggtt cttggctaaa tcccatccag 500

tcccgaaaca cgacagtgga gtatgtgaca ttcaccagca ttgctggtct 550

cactggagtg atcatgacaa tagccttgat tctcatggta acttcagcta 600

ctgagttcat ccggaggagt tattttgaag tcttctggta tactcaccac 650

ctttttatct tctatatcct tggcttaggg attcacggca ttggtggaat 700

tgtccggggt caaacagagg agagcatgaa tgagagtcat cctcgcaagt 750

gtgcagagtc ttttgagatg tgggatgatc gtgactccca ctgtaggcgc 800

cctaagtttg aagggcatcc ccctgagtct tggaagtgga tccttgcacc 850

ggtcattctt tatatctgtg aaaggatcct ccggttttac cgctcccagc 900

agaaggttgt gattaccaag gttgttatgc acccatccaa agttttggaa 950

ttgcagatga acaagcgtgg cttcagcatg gaagtggggc agtatatctt 1000

tgttaattgc ccctcaatct ctctcctgga atggcatcct tttactttga 1050

cctctgctcc agaggaagat ttcttctcca ttcatatccg agcagcaggg 1100

gactggacag aaaatctcat aagggctttc gaacaacaat attcaccaat 1150

tcccaggatt gaagtggatg gtccctttgg cacagccagt gaggatgttt 1200

tccagtatga agtggctgtg ctggttggag caggaattgg ggtcaccccc 1250

tttgcttcta tcttgaaatc catctggtac aaattccagt gtgcagacca 1300

caacctcaaa acaaaaaagg ttggtcatgc agcattaaac tttgacaagg 1350

ccactgacat cgtgacaggt ctgaaacaga aaacctcctt tgggagacca 1400

atgtgggaca atgagttttc tacaatagct acctcccacc ccaagtctgt 1450

agtgggagtt ttcttatgtg gccctcggac tttggcaaag agcctgcgca 1500

aatgctgtca ccgatattcc agtctggatc ctagaaaggt tcaattctac 1550

ttcaacaaag aaaatttttg agttatagga ataaggacgg taatctgcat 1600

tttgtctctt tgtatcttca gtaatttact tggtctcgtc aggtttgagc 1650

agtcacttta ggaag 1665

<210> 56
<211> 2400
<212> DNA
<213> Homo sapiens

<400> 56
ggactgcttt gtaactgcta agattgcaga cagaaatagc acacaaccac 50

tgtgagctgt atgcgattca gaaaccaaga ccaaattttg ctcactttca 100

ttaatcagtt gctcagatag aaggaaatga catctggttc tgtcttcttc 150

tacatcttaa tttttggaaa atatttttct catgggggtg gacaggatgt 200

caagtgctcc cttggctatt tcccctgtgg gaacatcaca aagtgcttgc 250

ctcagctcct gcactgtaac ggtgtggacg actgcgggaa tcaggccgat 300

gaggacaact gtggagacaa caatggatgg tccatgcaat ttgacaaata 350

ttttgccagt tactacaaaa tgacttccca atatcctttt gaggcagaaa 400

cacctgaatg tttggtcggt tctgtgccag tgcaatgtct ttgccaaggt 450

ctggagcttg actgtgatga aaccaattta cgagctgttc catcggtttc 500

ttcaaatgtg actgcaatgt cacttcagtg gaacttaata agaaagcttc 550

ctcctgattg cttcaagaat tatcatgatc ttcagaagct gtacctgcaa 600

aacaataaga ttacatccat ctccatctat gctttcagag gactgaatag 650

ccttactaaa ctgtatctca gtcataacag aataaccttc ctgaagccgg 700

gtgtttttga agatcttcac agactagaat ggctgataat tgaagataat 750

cacctcagtc gaatttcccc accaacattt tatggactaa attctcttat 800

tctcttagtc ctgatgaata acgtcctcac ccgtttacct gataaacctc 850

tctgtcaaca catgccaaga ctacattggc tggaccttga aggcaaccat 900

atccataatt taagaaattt gacttttatt tcctgcagta atttaactgt 950

tttagtgatg aggaaaaaca aaattaatca cttaaatgaa aatacttttg 1000

cacctctcca gaaactggat gaattggatt taggaagtaa taagattgaa 1050

aatcttccac cgcttatatt caaggacctg aaggagctgt cacaattgaa 1100

tctttcctat aatccaatcc agaaaattca agcaaaccaa tttgattatc 1150

ttgtcaaact caagtctctc agcctagaag ggattgaaat ttcaaatatc 1200

caacaaagga tgtttagacc tcttatgaat ctctctcaca tatattttaa 1250

gaaattccag tactgtgggt atgcaccaca tgttcgcagc tgtaaaccaa 1300

acactgatgg aatttcatct ctagagaatc tcttggcaag cattattcag 1350

agagtatttg tctgggttgt atctgcagtt acctgctttg gaaacatttt 1400

tgtcatttgc atgcgacctt atatcaggtc tgagaacaag ctgtatgcca 1450

tgtcaatcat ttctctctgc tgtgccgact gcttaatggg aatatattta 1500

ttcgtgatcg gaggctttga cctaaagttt cgtggagaat acaataagca 1550

tgcgcagctg tggatggaga gtactcattg tcagcttgta ggatctttgg 1600

ccattctgtc cacagaagta tcagttttac tgttaacatt tctgacattg 1650

gaaaaataca tctgcattgt ctatcctttt agatgtgtga gacctggaaa 1700

atgcagaaca attacagttc tgattctcat ttggattact ggttttatag 1750

tggctttcat tccattgagc aataaggaat ttttcaaaaa ctactatggc 1800

accaatggag tatgcttccc tcttcattca gaagatacag aaagtattgg 1850

agcccagatt tattcagtgg caatttttct tggtattaat ttggccgcat 1900

ttatcatcat agttttttcc tatggaagca tgttttatag tgttcatcaa 1950

agtgccataa cagcaactga aatacggaat caagttaaaa aagagatgat 2000

ccttgccaaa cgtttttttct ttatagtatt tactgatgca ttatgctgga 2050

tacccatttt tgtagtgaaa tttctttcac tgcttcaggt agaaatacca 2100

ggtaccataa cctcttgggt agtgattttt attctgccca ttaacagtgc 2150

tttgaaccca attctctata ctctgaccac aagaccattt aaagaaatga 2200

ttcatcggtt ttggtataac tacagacaaa gaaaatctat ggacagcaaa 2250

ggtcagaaaa catatgctcc atcattcatc tgggtggaaa tgtggccact 2300

gcaggagatg ccacctgagt taatgaagcc ggacctttтc acataccccт 2350

gtgaaatgtc actgatttct caatcaacga gactcaattc ctattcatga 2400

<210> 57
<211> 2967
<212> DNA
<213> Homo sapiens

<400> 57
tagagatccc tcgacctcga cccacgcgtc cgaggaaaga aaaaaagagg 50

aatggaaaga gacagagaaa ggaaatggga gtggaaggag ggaggactgc 100

tttgtaactg ctaagattgc agacagaaat agcacacaac cactgtgagc 150

tgtatgcgat tcagaaacca agaccaaatt ttgctcactt tcattaatca 200

gttgctcaga tagaaggaaa tgacatctgg ttctgtcttc ttctacatct 250

taatttttgg aaaatatttt tctcatgggg gtggacagga tgtcaagtgc 300

tcccttggct atttcccctg tgggaacatc acaaagtgct tgcctcagct 350

cctgcactgt aacggtgtgg acgactgcgg gaatcaggcc gatgaggaca 400

actgtggaga caacaatgga tggtctctgc aatttgacaa atattttgcc 450

agttactaca aaatgacttc ccaatatcct tttgaggcag aaacacctga 500

atgtttggtc ggttctgtgc cagtgcaatg tctttgccaa ggtctggagc 550

ttgactgtga tgaaaccaat ttacgagctg ttccatcggt ttcttcaaat 600

gtgactgcaa tgtcacttca gtggaactta ataagaaagc ttcctcctga 650

ttgcttcaag aattatcatg atcttcagaa gctgtacctg caaaacaata 700

agattacatc catctccatc tatgctttca gaggactgaa tagccttact 750

aaactgtatc tcagtcataa cagaataacc ttcctgaagc cgggtgtttt 800

tgaagatctt cacagactag aatggctgat aattgaagat aatcacctca 850

gtcgaatttc cccaccaaca ttttatggac taaattctct tattctctta 900

gtcctgatga ataacgtcct cacccgttta cctgataaac ctctctgtca 950

acacatgcca agactacatt ggctggacct tgaaggcaac catatccata 1000

atttaagaaa tttgactttt atttcctgca gtaatttaac tgttttagtg 1050

atgaggaaaa acaaaattaa tcacttaaat gaaaatactt ttgcacctct 1100

ccagaaactg gatgaattgg atttaggaag taataagatt gaaaatcttc 1150

caccgcttat attcaaggac ctgaaggagc tgtcacaatt gaatctttcc 1200

tataatccaa tccagaaaat tcaagcaaac caatttgatt atcttgtcaa 1250

actcaagtct ctcagcctag aagggattga aatttcaaat atccaacaaa 1300

ggatgtttag acctcttatg aatctctctc acatatattt taagaaattc 1350

cagtactgtg ggtatgcacc acatgttcgc agctgtaaac caaacactga 1400

tggaatttca tctctagaga atctcttggc aagcattatt cagagagtat 1450

ttgtctgggt tgtatctgca gttacctgct ttggaaacat ttttgtcatt 1500

tgcatgcgac cttatatcag gtctgagaac aagctgtatg ccatgtcaat 1550

catttctctc tgctgtgccg actgcttaat gggaatatat ttattcgtga 1600

tcggaggctt tgacctaaag tttcgtggag aatacaataa gcatgcgcag 1650

ctgtggatgg agagtactca ttgtcagctt gtaggatctt tggccattct 1700

gtccacagaa gtatcagttt tactgttaac atttctgaca ttggaaaaat 1750

acatctgcat tgtctatcct tttagatgtg tgagacctgg aaaatgcaga 1800

acaattacag ttctgattct catttggatt actggtttta tagtggcttt 1850

cattccattg agcaataagg aatttttcaa aaactactat ggcaccaatg 1900

gagtatgctt ccctcttcat tcagaagata cagaaagtat tggagcccag 1950

atttattcag tggcaatttt tcttggtatt aatttggccg catttatcat 2000

catagttttt tcctatggaa gcatgtttta tagtgttcat caaagtgcca 2050

taacagcaac tgaaatacgg aatcaagtta aaaaagagat gatccttgcc 2100

aaacgttttt tctttatagt atttactgat gcattatgct ggatacccat 2150

ttttgtagtg aaatttcttt cactgcttca ggtagaaata ccaggtacca 2200

taacctcttg ggtagtgatt tttattctgc ccattaacag tgctttgaac 2250

ccaattctct atactctgac cacaagacca tttaaagaaa tgattcatcg 2300

gttttggtat aactacagac aaagaaaatc tatggacagc aaaggtcaga 2350

aaacatatgc tccatcattc atctgggtgg aaatgtggcc actgcaggag 2400

atgccacctg agttaatgaa gccggacctt ttcacatacc cctgtgaaat 2450

gtcactgatt tctcaatcaa cgagactcaa ttcctattca tgactgactc 2500

tgaaattcat ttcttcgcag agaatactgt gggggtgctt catgagggat 2550

ttactggtat gaaatgaata ccacaaaatt aatttataat aatagctaag 2600

ataaatattt tacaaggaca tgaggaaaaa taaaaatgac taatgctctt 2650

acaaagggaa gtaattatat caataatgta tatatattag tagacatttt 2700

gcataagaaa ttaagagaaa tctacttcag taacattcat tcatttttct 2750

aacatgcatt tattgagtac ccactactat gtgcatagca ttgcaatata 2800

gtcctggaag tagacagtgc agaacctttc aatctgtaga tggtgtttaa 2850

tgacaaaaga ctatacaaag tccatctgca gttcctagtt taaagtagag 2900

ctttacctgt catgtgcatc agcaagaatc ataggcactt ttaaataaag 2950

gtttaaagtt ttggaat 2967

<210> 58
<211> 3032
<212> DNA
<213> Homo sapiens

<400> 58
gatccattgt gctggcgcgg attctttatc actggaagga ggaggagggg 50

gagagtcgct cccgccgggc gagcatgggg cgcctggcct cgaggccgct 100

gctgctggcg ctcctgtcgt tggctctttg ccgagggcgt gtggtgagag 150

tccccacagc gaccctggtt cgagtggtgg gcactgagct ggtcatcccc 200

tgcaacgtca gtgactatga tggccccagc gagcaaaact ttgactggag 250

cttctcatct ttggggagca gctttgtgga gcttgcaagc acctgggagg 300

tggggttccc agcccagctg taccaggagc ggctgcagag gggcgagatc 350

ctgttaaggc ggactgccaa cgacgccgtg gagctccaca taaagaacgt 400

ccagccttca gaccaaggcc actacaaatg ttcaaccccc agcacagatg 450

ccactgtcca gggaaactat gaggacacag tgcaggttaa agtgctggcc 500

gactccctgc acgtgggccc cagcgcgcgg cccccgccga gcctgagcct 550

gcgggagggg gagcccttcg agctgcgctg caccgccgcc tccgcctcgc 600

cgctgcacac gcacctggcg ctgctgtggg aggtgcaccg cggcccggcc 650

aggcggagcg tcctcgccct gacccacgag ggcaggttcc acccgggcct 700

ggggtacgag cagcgctacc acagtgggga cgtgcgcctc gacaccgtgg 750

gcagcgacgc ctaccgcctc tcagtgtccc gggctctgtc tgccgaccag 800

ggctcctaca ggtgtatcgt cagcgagtgg atcgccgagc agggcaactg 850

gcaggaaatc caagaaaagg ccgtggaagt tgccaccgtg gtgatccagc 900

cgacagttct gcgagcagct gtgcccaaga atgtgtctgt ggctgaagga 950

aaggaactgg acctgacctg taacatcaca acagaccgag ccgatgacgt 1000

ccggcccgag gtgacgtggt ccttcagcag gatgcctgac agcaccctac 1050

ctggctcccg cgtgttggcg cggcttgacc gtgattccct ggtgcacagc 1100

tcgcctcatg ttgctttgag tcatgtggat gcacgctcct accatttact 1150

ggttcgggat gttagcaaag aaaactctgg ctactattac tgccacgtgt 1200

ccctgtgggc acccggacac aacaggagct ggcacaaagt ggcagaggcc 1250

gtgtcttccc cagctggtgt gggtgtgacc tggctagaac cagactacca 1300

ggtgtacctg aatgcttcca aggtccccgg gtttgcggat gaccccacag 1350

agctggcatg ccgggtggtg gacacgaaga gtggggaggc gaatgtccga 1400

ttcacggttt cgtggtacta caggatgaac cggcgcagcg acaatgtggt 1450

gaccagcgag ctgcttgcag tcatggacgg ggactggacg ctaaaatatg 1500

gagagaggag caagcagcgg gcccaggatg gagactttat tttttctaag 1550

gaacatacag acacgttcaa tttccggatc caaaggacta cagaggaaga 1600

cagaggcaat tattactgtg ttgtgtctgc ctggaccaaa cagcggaaca 1650

acagctgggt gaaaagcaag gatgtcttct ccaagcctgt taacatattt 1700

tgggcattag aagattccgt gcttgtggtg aaggcgaggc agccaaagcc 1750

tttctttgct gccggaaata catttgagat gacttgcaaa gtatcttcca 1800

agaatattaa gtcgccacgc tactctgttc tcatcatggc tgagaagcct 1850

gtcggcgacc tctccagtcc caatgaaacg aagtacatca tctctctgga 1900

ccaggattct gtggtgaagc tggagaattg gacagatgca tcacgggtgg 1950

atggcgttgt tttagaaaaa gtgcaggagg atgagttccg ctatcgaatg 2000

taccagactc aggtctcaga cgcagggctg taccgctgca tggtgacagc 2050

ctggtctcct gtcaggggca gcctttggcg agaagcagca accagtctct 2100

ccaatcctat tgagatagac ttccaaacct caggtcctat atttaatgct 2150

tctgtgcatt cagacacacc atcagtaatt cggggagatc tgatcaaatt 2200

gttctgtatc atcactgtcg agggagcagc actggatcca gatgacatgg 2250

cctttgatgt gtcctggttt gcggtgcact cttttggcct ggacaaggct 2300

cctgtgctcc tgtcttccct ggatcggaag ggcatcgtga ccacctcccg 2350

gagggactgg aagagcgacc tcagcctgga gcgcgtgagt gtgctggaat 2400

tcttgctgca agtgcatggc tccgaggacc aggactttgg caactactac 2450

tgttccgtga ctccatgggt gaagtcacca acaggttcct ggcagaagga 2500

ggcagagatc cactccaagc ccgtttttat aactgtgaag atggatgtgc 2550

tgaacgcctt caagtatccc ttgctgatcg gcatcggtct gtccacggtc 2600

atcgggctcc tgtcctgtct catcgggtac tgcagctccc actggtgttg 2650

taagaaggag gttcaggaga cacggcgcga gcgccgcagg ctcatgtcga 2700

tggagatgga ctaggctggc ccgggagggg agtgacagag ggacgttcta 2750

ggagcaattg gggcaagaag aggacagtga tattttaaaa caaagtgtgt 2800

tacactaaaa accagtcctc tctaatctca ggtgggactt ggcgctctct 2850

cttttctgca tgtcaagttc tgagcgcgga catgtttacc agcacacggc 2900

tcttcttccc acggcacttt ctgatgtaac aatcgagtgt gtgttttccc 2950

aactgccagt gataaagaat ccgcgccagc acaatggatc tcgagggatc 3000

ttccatacct accagttctg cgcctgcagg tc 3032

<210> 59
<211> 2771
<212> DNA
<213> Homo sapiens

<400> 59
cccacgcgtc cgcggtaaat cacttcttgg aggtgccctg cacgccggtc 50

ctgggagcag gcggcctccc gggggtgcgg gagccccact cctccgtggt 100

gtgttccatt tgcttcccac atctggagga gctgacgtgc cagcctcccc 150

cagcaccacc cagggacggg aggcatgagc cggtcaaggc acctgggcaa 200

aatccggaag cgtctggaag atgtcaagag ccagtgggtc cggccagcca 250

gggctgactt tagtgacaac gagagtgccc ggctggccac ggacgccctc 300

ttggatgggg gttctgaagc ctactggcgg gtgctcagcc aggaaggcga 350

ggtggacttc ttgtcctcgg tggaggccca gtacatccag gcccaggcca 400

gggagccccc gtgtccccca gacaccctgg gaggggcgga agcaggccct 450

aagggactgg actccagctc cctacagtcc ggcacctact tccctgtggc 500

ctcagagggc agcgagccgg ccctactgca cagctgggcc tcagctgaga 550

agccctacct gaaggaaaaa tccagcgcca ctgtgtactt ccagaccgtc 600

aagcacaaca acatcagaga cctcgtccgc cgctgcatca cccggactag 650

ccaggtcctg gtcatcctga tggatgtgtt cacggatgtg gagatcttct 700

gtgacattct agaggcagcc aacaagcgtg gggtgttcgt ttgtgtgctc 750

ctggaccagg gaggtgtgaa gctcttccag gagatgtgtg acaaagtcca 800

gatctctgac agtcacctca agaacatttc catccggagt gtggaaggag 850

agatatactg tgccaagtca ggcaggaaat tcgctggcca aatccgggag 900

aagttcatca tctcggactg gagatttgtc ctgtctggat cttacagctt 950

cacctggctc tgcggacacg tgcaccggaa catcctctcc aagttcacag 1000

gccaggcggt ggagctgttt gacgaggagt tccgccacct ctacgcctcc 1050

tccaagcctg tgatgggcct gaagtccccg cggctggtcg cccccgtccc 1100

gcccggagca gccccggcca atggccgcct tagcagcagc agtggctccg 1150

ccagtgaccg cacgtcctcc aacccccttca gcggccgctc ggcaggcagc 1200

caccccggta cccgaagtgt gtccgcgtct tcagggccct gtagccccgc 1250

ggccccacac ccgcctccac cgccccggtt ccagccccac caaggccctt 1300

ggggagcccc gagtccccag gcccacctct ccccgcggcc ccacgacggc 1350

ccgcccgccg ctgtctacag caacctgggg gcctacaggc ccacgcggct 1400

gcagctggag cagctgggcc tggtgccgag gctgactcca acctggaggc 1450

ccttcctgca ggcctccccct cacttctgaa ggtcccatcc cctgctgccc 1500

tccgcaggcc cagggctggg cactccctga gacccaaaga cccacctcaa 1550

cgacgagtgg cgttgagcca cttccctttg aaaagacact caaaatcact 1600

gccatggttc aatgttccca ggccccaggc catccacttg ccggcccccа 1650

ccagttcttg ggttccccgc tctagtttga cctgtgcagc acattccaga 1700

aggttccagg ggaggttgtg gggcagctag aggacaaaat catgaaaaca 1750

gagtccctgt cttccagaga tcatccgggg ctttaatatt aatggccccc 1800

aaaactccgt aagaagcagg aaatgcagcc caagttttac aaatgggtaa 1850

acagaggcac tgagagatag atggtagttt ggtacttctg gttcccagtg 1900

cccaggaatg gtccactccc aagaaattca ggaaagaaag actgaggaga 1950

aggtgtggga acattctgga tgtttcggga gagttgggga aactcctcct 2000

cttaggaaag gctaatacta gggtatcctt gggcccaatg aattaggggt 2050

gaggccccag aacccgttat ctatgagttg tatgggggag ccatctgaag 2100

ctgtagccac cagggatgca gctagctgag gagtttgggg tgttgggttg 2150

gacaaggcag gttagtagac tcagattctt gcttcaaaga gccttgggct 2200

ggcctggagg tccctggagt ctagactgga cctaggagct tgagttgtca 2250

ggggccagga ctggccccac tgcagtgccc aggccagtct tgagcagcag 2300

ggagggctca gctgtcccca gatccaggtg cctctgacca gcctggtcac 2350

ctcctgagga ataaatgctg aacctcacaa gccccatcat tcatttcttc 2400

tcaattcaca gtgccctct ttgtttctgg ggtggaacta ggtcctgagg 2450

gcacagccta gctgagtgca aagaaatata ggatgcttag aaagcataca 2500

ggaggggcca ggcgtggtgg ctcatgcctg taatcccaga actttgggat 2550

gccaaggtgg ttggattacc tgagatcagg tggattacct ggtctcgaga 2600

ccagcctgac caatatggtg aaaccccgtc tctactaaaa atacaaaaat 2650

taggctgaga caggagaatt gcttgaaccc aggaagcaga ggttgcaatg 2700

agctgagatt gcatcactgc actccagcat gggcaacaaa gcaagactcc 2750

gtcacagaaa aaaaaaaaaa a 2771

<210> 60
<211> 4060
<212> DNA
<213> Homo sapiens

<400> 60
atggctctgc cgcagccgcc gccgcccgcc ggggcgcgtg acccggtgac 50

cccggagcac tgggcatccg gacccccctc gcaggcgcag ccgctattaa 100

ggcaggaggc taaagaggag gaggaagggg aggagacagg cgtccagggc 150

gcctggggaa ccggcacggc ggagcagcgg cggcggggct gggggggaggc 200

cgccgagtcg gccgcggccg aggaggggca ggcggaggtc gggggcgccg 250

cggccgcggg gtcaggttcc ccggcgggag gcgcaggtgg cggcctgggg 300

agctggcggc cgctcctcgc ctggctgcag cggcggcagc cccagtgctg 350

cccctgtgcg gcgccccttt cccgctccgc cgcgcactgt tgtcatggag 400

gaaccaagat ggcggctctg gcctacaacc tgggcaagcg ggagatcaac 450

cactacttca gcgtgaggag cgccaaggtg ctggcgctgg tggccgtgct 500

gctgctcgca gcgtgccacc tcgcctcccg ccgctaccga ggcaatgatt 550

cgtgtgaata ccttctctca agtggcagat ttcttggaga gaaagtttgg 600

caacctcaca gttgtatgat gcataaatac aaaatcagtg aagcaaagaa 650

ctgccttgta gataaacata ttgcatttat tggagattcc agaattcgtc 700

aattgtttta ttcttttgta aaaataatta atccccaatt caaagaagaa 750

ggaaataagc atgaaaacat tccttttgaa gacaagactg catcagttaa 800

agtggatttt ctgtggcatc ctgaagttaa tggttctatg aaacagtgta 850

tcaaagtgtg gactgaggat tccattgcaa agccacatgt gattgtagca 900

ggagctgcca catggtccat caagattcac aatggtagca gtgaagcgct 950

ttctcaatat aaaatgaaca tcacctccat agcaccactt ttagaaaaat 1000

tggcaaagac tagtgatgtt tattgggtct tacaagatcc tgtttatgaa 1050

gatctattaa gtgaaaatag gaagatgatc actaatgaga agatagatgc 1100

ttacaatgaa gctgcagtca gtattttgaa tagtagcacc agaaattcta 1150

aatcaaatgt taagatgttc agtgtttcca aattaattgc tcaagaaacc 1200

atcatggaat ctttggatgg cttacatctt cctgaatcga gcagagaaac 1250

tactgcaatg attcttatga atgtgtattg caataagatt ttgaagcctg 1300

tagatgggtc ctgttgtcaa cctcggcctc ctgttactct catacagaag 1350

ctagctgctt gtttttttcac tttatctatt atcggatatt taatttttta 1400

cataattcat cgtaatgctc atcggaagaa taagccgtgt actgatttgg 1450

aaagtggaga ggaaaagaaa aatattatca ataccctgt gtcttcatta 1500

gaaatacttt tacaatcttt ctgcaaactt ggcctgatta tggcatattt 1550

ctatatgtgt gaccgtgcaa atctgttcat gaaggaaaac aaatttttata 1600

cacattcatc tttctttatt ccaattatct acattttggt tttgggagta 1650

ttttataatg aaaatactaa agagactaaa gtattaaata gagaacaaac 1700

agacgaatgg aaaggctgga tgcaacttgt gattttgatt tatcacattt 1750

ctggagcaag tacatttttg cctgtataca tgcacattcg agttctggtt 1800

gctgcatatt tatttcagac agggtatggg catttctcat acttttggat 1850

aaaaggagat tttggaatct atagagtatg tcaggtttta tttcgtctca 1900

atttcctggt agtggtgtta tgtatagtaa tggatcgacc ttatcaattc 1950

tattactttg tcccccttggt cactgtatgg ttcatggtca tatatgttac 2000

tttagcacta tggccacaaa taatccaaaa aaaagcaaac ggaaattgtt 2050

tctggcattt tggcttactg ttgaaactag gctttttgct gttattcata 2100

tgtttttttgg catattctca gggtgcattt gagaagatct tttctctttg 2150

gccattgtcc aagtgttttg aactgaaagg gaatgtatat gaatggtggt 2200

tcagatggag gttagaccgt tatgtagttt tccacggaat gctgtttgct 2250

tttatttatc tggctttgca gaagcgtcaa atactttctg aaggaaaggg 2300

tgaacctctt ttttcaaaca aaatttcaaa ttttctgttg tttatttcag 2350

tagtttcttt cttgacctat tccatctggg ctagcagttg taaaaacaaa 2400

gcagagtgca atgaactcca tccgtctgtt tctgtggtac agattttagc 2450

cttcatccta ataagaaaca tccctggata tgcccgttca gtttacagtt 2500

catttttgc ttggtttgga aaaatttcat tagagctatt tatttgccag 2550

tatcacatat ggctggcagc ggacacaagg ggtatcttgg tactgatacc 2600

tggaaaccct atgctcaaca tcattgtcag cactttcata tttgtttgtg 2650

tggcacatga aatttctcag atcactaatg atcttgcaca gattattatt 2700

cctaaagata actcatctct cttgaaaagg ttggcatgta tagctgcatt 2750

tttttgtgga ctcctcatct tatcatccat tcaagataaa tcaaaacatt 2800

aggttccaaa aattctaaaa aacctaaact cttcaggcta cctttgtgtg 2850

tctctagaag agaaaagcat ctatctggag atataaatgt gtwtgtaaat 2900

ataaacgttt gtggcaagag gacagttctg tgacatctgt tgaacatatg 2950

tggttgtata tattggaaat gtacatatcc aatatgaaat actaaaacaa 3000

acaaacaaac aaaaaaccag aatgcattgt ataggattgc atgtgaagtc 3050

ttttctactg aatctatatt tccatttgta agtgatttta agttaacata 3100

tgaaggcagg gaaatgatta cctttccagt aaaaagtata gataatttaa 3150

ttaacttagt gacaccacca agtgttttga tataactaaa tttgtggtaa 3200

taagactgtc tgcacctgta ttcattgtgg aacttcctct ttcattggaa 3250

actttcttac tcaagaatga cggcagtatt gttttcttat atgtgcaatg 3300

aagtggaatg ataaacagta tgcctttaat ttatatgtgt tcttgttctg 3350

atgttgtttc ctgaaatgat ttttcttcct aactgtggtt ttcgggtatg 3400

caagcctaaa tctttgtaca ctttgtctca cagaatagtt ctgaggctcc 3450

atgacagggt tttgtcattg ttgatgttat tgttgcttcg ttttataaaa 3500

aagccaaaat ttttttttcca atccaaacgt tcacctgttt cctttcctca 3550

agctatacca gtgtaatacc agttaccctg tggatccatt taatatgtta 3600

tccccactaa ttaattttcg tatattattt ccaatatttg gaaagctctt 3650

tatagccatt tggtatttcc tattacccac ctcctatttt aaatatttat 3700

cagtctaaac ttgtgcagtg tagtaaacat gcaagttgtt acgattgagc 3750

tgtattacca taagtagaat tttaagtaaa ctggtgaatt tgggcaataa 3800

atgtttttgc tttttgtttg attttttttt acaagctaac tgttagaggt 3850

atacatttat ttatctgttg tacagatttg attatgattt taatgtttga 3900

aagattgcac ttgtttgctt ttactatatg tggggtaaaa tatattttct 3950

gttcacagta tatgaaaata tggagtaatt taaacagtaa ataaacattc 4000

tgtggatgct tatttttgta ttggcaaagt atcaattaaa ctatatgtgt 4050

tctttttcaa 4060

<210> 61
<211> 5472
<212> DNA
<213> Homo sapiens

<400> 61
gaaagacata cacacttcat gtaatgctac ctgcaagtct ccctagaaaa 50

gcagtttttg taggtgaaaa caatgaagcc aggtaatatt gcaaggaggc 100

tgtaatttta gcagacctac caacaacact gatgtaggaa gctcattatt 150

ttaatttctg gagcctttta atttttttctt tagaaagtgt ataaataatt 200

gcagtgctgc tttgcttcca aaactgggca gtgagttcaa caacaacgac 250

aacaacagcc gcagctcatc ctggccgtca tggagtttct tgaaagaacg 300

tatcttgtga atgataaagc tgccaagatg tatgctttca cactagaaag 350

aaggagctgc aaatgaacac ttcatagcaa tgtggaactc caacagaaac 400

cggtgaataa agatcagtgt cccagagaga gaccagagga gctggagtca 450

ggaggcatgt accactgcca cagtggctcc aagcccacag aaaaggggc 500

gaatgagtac gcctatgcca agtggaaact ctgttctgct tcagcaatat 550

gcttcatttt catgattgca gaggtcgtgg gtgggcacat tgctgggagt 600

cttgctgttg tcacagatgc tgcccacctc ttaattgacc tgaccagttt 650

cctgctcagt ctcttctccc tgtggttgtc atcgaagcct ccctctaagc 700

ggctgacatt tggatggcac cgagcagaga tccttggtgc cctgctctcc 750

atcctgtgca tctgggtggt gactggcgtg ctagtgtacc tggcatgtga 800

gcgcctgctg tatcctgatt accagatcca ggcgactgtg atgatcatcg 850

tttccagctg cgcagtggcg gccaacattg tactaactgt ggttttgcac 900

cagagatgcc ttggccacaa tcacaaggaa gtacaagcca atgccagcgt 950

cagagctgct tttgtgcatg cccttggaga tctatttcag agtatcagtg 1000

tgctaattag tgcacttatt atctacttta agccagagta taaaatagcc 1050

gacccaatct gcacattcat cttttccatc ctggtcttgg ccagcaccat 1100

cactatctta aaggacttct ccatcttact catggaaggt gtgccaaaga 1150

gcctgaatta cagtggtgtg aaagagctta ttttagcagt cgacggggtg 1200

ctgtctgtgc acagcctgca catctggtct ctaacaatga atcaagtaat 1250

tctctcagct catgttgcta cagcagccag ccgggacagc caagtggttc 1300

ggagagaaat tgctaaagcc cttagcaaaa gctttacgat gcactcactc 1350

accattcaga tggaatctcc agttgaccag gaccccgact gccttttctg 1400

tgaagacccc tgtgactagc tcagtcacac cgtcagtttc ccaaatttga 1450

caggccacct tcaaacatgc tgctatgcag tttctgcatc atagaaaata 1500

aggaaccaaa ggaagaaatt catgtcatgg tgcaatgcac attttatcta 1550

tttatttagt tccattcacc atgaaggaag aggcactgag atccatcaat 1600

caattggatt atatactgat cagtagctgt gttcaattgc aggaatgtgt 1650

atatagatta ttcctgagtg gagccgaagt aacagctgtt tgtaactatc 1700

ggcaatacca aattcatctc ccttccaata atgcatcttg agaacacata 1750

ggtaaatttg aactcaggaa agtcttacta gaaatcagtg gaagggacaa 1800

atagtcacaa aattttacca aaacattaga aacaaaaaat aaggagagcc 1850

aagtcaggaa taaaagtgac tctgtatgct aacgccacat tagaacttgg 1900

ttctctcacc aagctgtaat gtgatttttt tttctactct gaattggaaa 1950

tatgtatgaa tatacagaga agtgcttaca actaattttt atttacttgt 2000

cacattttgg caataaatcc ctcttatttc taaattctaa cttgtttatt 2050

tcaaaacttt atataatcac tgttcaaaag gaaatatttt cacctaccag 2100

agtgcttaaa cactggcacc agccaaagaa tgtggttgta gagacccaga 2150

agtcttcaag aacagccgac aaaaacattc gagttgaccc caccaagttg 2200

ttgccacaga taatttagat atttacctgc aagaaggaat aaagcagatg 2250

caaccaattc attcagtcca cgagcatgat gtgagcactg ctttgtgcta 2300

gacattgggc ttagcattga aactataaag aggaatcaga cgcagcaagt 2350

gcttctgtgt tctggtagca actcaacact atctgtggag agtaaactga 2400

agatgtgcag gccaacattc tggaaatcct atgtcactgg gtttggtttg 2450

gaaacctgga cttctgcatt tttaaaagtt acccagagat gcttctaaag 2500

atgagccata gtctagaaga ttgtcaacca caggagttca ttgagtggga 2550

cagctagaya catacattgg cagctacaat agtatcatga attgcaatga 2600

tgtagtgggg tataaaagga aagcgatgga tattgccgga tgggcatggc 2650

cagtgatgtt tcacgtcatt gaggtgacag ctctgctgga ctttgaatta 2700

catatggagg ctctccagga agacgaagaa gagaaggaca ttctaggcaa 2750

aaagaagact aggcacaagg cacacttatg tttgtctgtt agcttttagt 2800

tgaaaaagca aaatacatga tgcaaagaaa cctctccacg ctgtgatttt 2850

taaaactaca tacttttgc aactttatgg ttatgagtat tgtagagaac 2900

aggagatagg tcttagatga tttttatgtt gttgtcagac tctagcaagg 2950

tactagaaac ctagcaggca ttaataattg ttgaggcaat gactctgagg 3000

ctatatctgg gccttgtcat tatttatcat ttatatttgt atttttttct 3050

gaaatttgag ggccaagaaa acattgactt tgactgagga ggtcacatct 3100

gtgccatctc tgcaaatcaa tcagcaccac tgaaataact acttagcatt 3150

ctgctgagct ttccctgctc agtagagaca aatatactca tcccccacct 3200

cagtgagctt gtttaggcaa ccaggattag agctgctcag gttcccaacg 3250

tctcctgcca catcgggttc tcaaaatgga aagaatggtt tatgccaaat 3300

cacttttcct gtctgaagga ccactgaatg gttttgtttt tccatatttt 3350

gcataggacg ccctaaagac taggtgactt ggcaaacaca caagtgttag 3400

tataattctt tgcttctgct tctttttgaa aatcatgttt agatttgatt 3450

ttaagtcaga aattcactga atgtcaggta atcattatgg agggagattt 3500

gtgtgtcaac caaagtaatt gtcccatggc cccagggtat ttctgttgtt 3550

tccctgaaat tctgcttttt tagtcagcta gattgaaaac tctgaacagt 3600

agatgtttat atggcaaaat gcaagacaat ctataaggga gattttaagg 3650

attttgagat gaaaaaacag atgctactca ggggctttat ggaccatcca 3700

tcaattctga agttctgact ctcccattac cctttccctg gtgtggtcag 3750

aactccaggt cactggaagt tagtggaatc atgtagttga attctttact 3800

tcaagacatt gtattctctc cagctatcaa aacattaatg atcttttatg 3850

tctttttttt gttattgtta tactttaagt tctggggtac atgtgcggaa 3900

catgtaggtt tgttacatag gtatacatgt gccatggtgg tttgctgcac 3950

tcatcaacct gtcatctaca ttcttttatg tctgtctttc aaagcaacac 4000

tctgttcttc tgagtagtga aatcaggtca actttaccac cagcctccat 4050

ttttaatatg cttcaccatc atccagcacc tacttaagat ttatctaggg 4100

ctctgtggtg atgttaggac ccataaaaga aatttatgcc ttccatatgt 4150

ttggttacag atgggaaatg ggaatgttga aggacatgaa agaaaggatg 4200

tttacacatt aagcatcagt tctgaagcta gattgtctga gtttgaatct 4250

tagctcttcc ctttattagc tctgtgacct cgagctagtt acttaaatgc 4300

tctgatcctc tatttcctga tcagtgaaac ctcccctattc aaatgtgtga 4350

gagtttaata aattaggaca cttaaaaatg ttggagcagt gcatagcatg 4400

tagtgttcag tacatgttaa atgttgtttt ttattatgta caaacatgtg 4450

tgggcacaga attttaaatc atctcaactt ttgagaaatt ttgagttatc 4500

aacaccgttc ccacaagaca gtggcaaaat tattggtgag aattaaacag 4550

ctgtttctca gaggaagcaa tggaggcttg ctgggataaa ggcatttact 4600

gagaggctgt tacctagtga gagtgatgaa ttaattaaaa tagtcgaatc 4650

cctttctgac tgtctctgaa agcttccgct tttatctttg aagagcagaa 4700

ttgtcacccc aaggacattt attaataaaa agaacaactg tccagtgcaa 4750

tgaaggcaaa gtcataggtc tcccaagtct taccccattc ctgtgaaata 4800

tcaagttctt ggcttttctc tgtcatgtag cctcaacttt ctccgaccgg 4850

gtgcatttct ttctctggtt tctaaattgc cagtggcaaa tttggatcac 4900

ttacttaata tctgttaaat tttgtgaccc aacaaagtct tttagcactg 4950

tggtgtcaaa aagaaaaaca cctcccaggc atatacattt tatagattcc 5000

tggagaatgt tgctctccag ctccatcccc acccaatgaa atatgatcca 5050

gagagtcttg caaagagaca agcctcattt tccacaatta gctctaaagt 5100

gcctccagga aatgattttc tcagctcatc tctctgtatt ccctgttttg 5150

gatcacaggg caatctgttt aaatgactaa ttacagaaat cattaaaggc 5200

accaagcaaa tgtcatctct gaatacacac atcccaagct ttacaaatcc 5250

tgcctggctt gacagtgatg aggccactta acagtccagc gcaggcggat 5300

gttaaaaaaa ataaaaaggt gaccatctgc ggtttagttt tttaactttc 5350

tgatttcaca cttaacgtct gtcattctgt tactgggcac ctgtttaaat 5400

tctattttaa aatgttaatg agtgttgttt aaaataaaat caggaaagag 5450

agaaaaaaaa aaaaaaaaaa aa 5472

<210> 62
<211> 5170
<212> DNA
<213> Homo sapiens

<400> 62
gcggcgggag agctgctggc tcgcccggat cccgggagct gcctggaggc 50

gggcccggcc cggggaaggt gagcggctgc gggacccagc ccctcgccgg 100

gagcgggcac catggtgctg tcggtgcctg tgatcgcgct gggcgccacg 150

ctgggcacag ccaccagcat cctcgcgttg tgcggggtca cctgcctgtg 200

tcggcacatg caccccaaga aggggctgct gccgcgggac caggaccccg 250

acctggagaa ggcgaagccc agcttgctcg ggtctgcaca acagttcaat 300

gttaaaaagt ccacggaacc tgttcagccc cgtgccctcc tcaagttccc 350

agacatctat ggacccaggc cagctgtgac ggctccagag gtcatcaact 400

atgcagacta ttcactgagg tctacggagg agcccactgc acctgccagc 450

ccccaacccc cgaatgacag tcgcctcaag aggcaggtca cagaggagct 500

gttcatcctc cctcagaatg gtgtggtgga ggatgtctgt gtcatggaga 550

cctggaaccc agagaaggct gccagttgga accaggcccc caaactccac 600

tactgcctgg actatgactg tcagaaggca gaattgtttg tgactcgcct 650

ggaagctgtg accagcaacc acgacggagg ctgtgactgc tacgtccaag 700

ggagtgtggc caataggacc ggctctgtgg aggctcagac agccctaaag 750

aagcggcagc tgcacaccac ctgggaggag ggcctggtgc tccccctggc 800

ggaggaggag ctccccacag ccaccctgac gctgaccttg aggacctgcg 850

accgcttctc ccgtcacagc gtggccgggg agctccgcct gggcctggac 900

gggacatctg tgcctctagg ggctgcccag tggggcgagc tgaagacttc 950

agcgaaggag ccatctgcag gagctggaga ggtcctacta tccatcagct 1000

acctcccggc tgccaaccgc ctcctggtgg tgctgattaa agccaagaac 1050

ctccactcta accagtccaa ggagctcctg gggaaggatg tctctgtcaa 1100

ggtgaccttg aagcaccagg ctcggaagct gaagaagaag cagactaaac 1150

gagctaagca caagatcaac cccgtgtgga acgagatgat catgtttgag 1200

ctgcctgacg acctgctgca ggcctccagt gtggagctgg aagtgctggg 1250

ccaggacgat tcagggcaga gctgtgcgct tggccactgc agcctgggcc 1300

tgcacacctc gggctctgag cgcagccact gggaggagat gctcaaaaac 1350

cctcgccggc agattgccat gtggcaccag ctgcacctgt aaccagctgc 1400

ccagctgcct cccttcttgg acagccctga cccgtcctct gcaacctcct 1450

ttctgtgccc ctttcctcat tctgacaccc agaagacagt gacagatgtg 1500

tttgcaaggc tgggatggct ctctcatcat actcttgttt cttagaaata 1550

agcaagacag agcaggaaat ggaatatgcg ggtcacactg aggaatgcat 1600

tttgctcatc tgtgttattg aaggaggtgc ttattaaata cagttcctat 1650

gcctgtttta taggtggggt taggccagat gcagagaaag ctaaatgtgg 1700

gaatcatgga tgcaaagaag aatttggctt tttgaaaaac aagcatttca 1750

aaaatgatga aggaagtgaa agtatcctgg atcaactcct agagttagag 1800

attgcccagg tggaaagaaa ccttagccag cgttcaatca agctcaccat 1850

gcagggcagt cacccggcag ttctcaaact ttagcatgtg aagagtcacc 1900

agcagattcc tgggctcgcc tggagacatt cctagtcggt attcctggtc 1950

gaagcccagg agccttcctt tttaacaagc tgatgtagag ggtggagcac 2000

tgtatgtgga gaaattcctt ctacaatatt ccacacaggt ttttggccac 2050

agtccttgat ggagtcccaa aaccatggtg cagccagttc caatgctgga 2100

cacctcaacc atcagggtga aatctggggc ctcagctttt taatttaatt 2150

attttaattc ttaatacttt aatttgtgca tttcataagc cccctgctct 2200

tggactgaat tttgtgcttt ttattgaaga attttattgt ttttatctta 2250

aaatcagttt ctattatcct tggggagacc atccctaaca aagtacaggt 2300

gggatctcct gtgagtcatt ggctgggttc tgattgctag atgtcacacc 2350

caccagcatc accaaagtga ctctgagata gaccggtccc ttctcagcgt 2400

tccagtcact tcaggaggaa tttagttatt gacttagtct atgacatctg 2450

gctacatgta ggtagagaag aaagacaatt ttaaaaagga aatcaggtct 2500

tttgcaactg tgcctccctc tgtctgtttt cacttgaatg ggtaaataac 2550

cagcagctag gttttgaatt cctaccttgt tattctaaac agatgtccac 2600

attgttaatt aaatctaaat tatgagcctt gctgagtgga tacggtactt 2650

acacctgaac caggattcct gggttctgtt gttgacattg cccttcagca 2700

cctgtttggc cagctgtata agataggact aatgactagg aagcctaccc 2750

caatgaatga tatactagat gaaatagtgt tcaaaacctg taggcactct 2800

ctggctaaaa acaaactctg aggccaccag cagatcatct ttaagctaag 2850

ttactatttt tcacctttct ttttagacgg agttttgctc tttgttgccc 2900

aggctggagt gcagtggcac gatctcggct cactgcaacc tccgcctccc 2950

aagttcaagc gattctcctg tctcagcctc ctgggtagct gggattacag 3000

gtgcccacca acatgcctgg ctaatttttg tacttttagt agagatgggg 3050

tttcaccatg ttggccaggc tggtcttcaa ctccagatct caggtgatct 3100

accctcctcg gcctcccaaa gtactgggat tacaggcctg agccaccgcg 3150

cccggcctat ttttcacttt aatttggcag ctgagaatgc ccaaaaagtg 3200

ccagaagcat cgtggcattt ccagaaccat ggattctgcc tttggacccc 3250

tctctattaa tattaaaact ctgggccttc agatgtcacc ctaatccact 3300

gccctaagac agaatttctg gacaagatgg gtaagggctt cattccttca 3350

acaagtcaag tcatacttgg cctctccctg agaatctgag caggagcctt 3400

ataacctgtg gtcattattt tttctttctg tacagaaata gaaaagcatt 3450

agaaataact tctaaccatc ctctgaaaaa acagaaaaaa tatcgaatcc 3500

ctctttcatg agaagtcttt tggataattg gaaaccttca tcactgaggt 3550

tggccagccc ctgccaagtg ttgtgtaggc aaagcacttg ttagtggctt 3600

cctatgaaat gttttagaga tctcttcacc atactggttt cttctctttg 3650

gttggtgtgg gtaaaagaaa acaaaacatt tcctataagc tgaaagctga 3700

ccagcattct cttcttggta acatctacta ctccaaccta gaaaatttgg 3750

attctagacc aaaaatcagg aaacatggct ccttataaat ctgtgcagct 3800

gccttatagt accatcaaag gaatttcagg tgggctgggc ggggccccga 3850

tcccagaatt atcaactcca cccatcatca tttggtcatg aagcatcctt 3900

tcattcttct tcttcttttt tttgggggggg gcggggcggg ggagggatct 3950

caaagtttta gtcttccaga atccaaatta aaggttgccc ctgatggggg 4000

ccaggttccg ccacagaaca tcttagatgt cagccttgac ctcacttagc 4050

agggattaca gaaatgagat acattttgaa ggagagttgt ctgttatgtt 4100

cactgtattc taagtgcctg ggataaagct gtctcatggg tgctccatat 4150

atattcatat atatttgttg agtgaattaa tgaattaaga gtggctggca 4200

gagtaggcag aaaaagacac tgcaaatggc ataaaaatta aagtcctagc 4250

tgagttctca atggtaaagg catcagatgt cttagcagtc aagctagaaa 4300

ttcatgacaa tgagtattac tatttgccta atgacaactc attgctctcc 4350

atgtaaatgt aatcaacaga tgaagagaat ataattgctc tgcttttcca 4400

ctaaaactcc atcttagtga attttaaatt atccagagat gtcaaactgc 4450

caaataaaaa tatttcagta gtctttgcat cagcttacct tgtaccagaa 4500

acatttccaa tttactatca aattatagta actgagcctg tgtgaagtat 4550

ctcatcattt tcgaaaggaa caccttgtgt gatgccagtg agcatttcta 4600

aaaagggtgt gaggtagagg taaaaataag gtgagagacc atttcagaat 4650

gcactgttgc tcaaaaaggt gatctggttc tttcttcaga gatttctacg 4700

gggatagaaa atcgggagtc tgccctcatt aatctgtgac tccacctctt 4750

gcatcaaatc aatatctatt tgttgagcac ttattgatta agaccttgca 4800

tatgtctgtc cattttgatt tgagatacaa ctttttgtgt gggttgaatg 4850

acaaatcact ccaaacaaaa ctgggcacag agaatcagct aggagaccag 4900

ttattcaggg tccatttctc ttggatgtaa aggagtcctg ggtaaaatgt 4950

ggctgtaacc taaaccaact agtccttgtg atttgtttct gccctctgtg 5000

tttcctgttg tcaaatgcta agtgtgtgtt ttgcagtcat gaactaaagc 5050

acaaaaagat gcatgagaca ttgtagtcat atgtctggtg tgacactttg 5100

1163
1162

gagcaaaaac cttgcagtgg taaataaaaa atttccaaca gggaaaaaaa 5150

aaaaaaaaaa aaaaaaaaaa 5170

<210> 63
<211> 5563
<212> DNA
<213> Homo sapiens

<400> 63
agcggcagct ccgctgaaaa ctgcattcag ccagtcctcc ggacttctgg 50

agcggggaca gggcgcaggg catcagcagc caccagcagg acctgggaaa 100

tagggattct tctgcctcca cttcaggttt tagcagcttg gtgctaaatt 150

gctgtctcaa aatgcagagg atctaatttg cagaggaaaa cagccaaaga 200

aggaagagga ggaaaaggaa aaaaaaaggg gtatattgtg gatgctctac 250

ttttcttgga aatgcaaaag attatgcata tttctgtcct cctttctcct 300

gttttatggg gactgatttt tggtgtctct tctaacagca tacagatagg 350

ggggctattt cctaggggcg ccgatcaaga atacagtgca tttcgagtag 400

ggatggttca gttttccact tcggagttca gactgacacc ccacatcgac 450

aatttggagg tggcaaacag cttcgcagtc actaatgctt tctgctccca 500

gttttcgaga ggagtctatg ctatttttgg attttatgac aagaagtctg 550

taaataccat cacatcattt tgcggaacac tccacgtctc cttcatcact 600

cccagcttcc caacagatgg cacacatcca tttgtcattc agatgagacc 650

cgacctcaaa ggagctctcc ttagcttgat tgaatactat caatgggaca 700

agtttgcata cctctatgac agtgacagag gcttatcaac actgcaagct 750

gtgctggatt ctgctgctga aaagaaatgg caagtgactg ctatcaatgt 800

gggaaacatt aacaatgaca agaaagatga gatgtaccga tcacttttc 850

aagatctgga gttaaaaaag gaacggcgtg taattctgga ctgtgaaagg 900

gataaagtaa acgacattgt agaccaggtt attaccattg gaaaacacgt 950

taaagggtac cactacatca ttgcaaatct ggaatttact gatggagacc 1000

tattaaaaat ccagtttgga ggtgcaaatg tctctggatt tcagatagtg 1050

gactatgatg attcgttggt atctaaattt atagaaagat ggtcaacact 1100

ggaagaaaaa gaataccctg gagctcacac aacaacaatt aagtatactt 1150

ctgctctgac ctatgatgcc gttcaagtga tgactgaagc cttccgcaac 1200

ctaaggaagc aaagaattga aatctcccga agggggaatg caggagactg 1250

tctggcaaac ccagcagtgc cctggggaca aggtgtagaa atagaaaggg 1300

ccctcaaaca ggttcaggtt gaaggtctct caggaaatat aaagtttgac 1350

cagaatggaa aaagaataaa ctatacaatt aacatcatgg agctcaaaac 1400

taatgggccc cggaagattg gctactggag tgaagtggac aaaatggttg 1450

ttacccttac tgagctccct tctggaaatg acacctctgg gcttgagaat 1500

aagactgttg ttgtcaccac aattttggaa tctccgtatg ttatgatgaa 1550

gaaaaatcat gaaatgcttg aaggcaatga gcgctatgag ggctactgtg 1600

ttgacctggc tgcagaaatc gccaaacatt gtgggttcaa gtacaagttg 1650

acaattgttg gtgatggcaa gtatggggcc agggatgcag acacgaaaat 1700

ttggaatggg atggttggag aacttgtata tgggaaagct gatattgcaa 1750

ttgctccatt aactattacc cttgtgagag aagaggtgat tgacttctca 1800

aagcccttca tgagcctcgg gatatctatc atgatcaaga agcctcagaa 1850

gtccaaacca ggagtgtttt cctttcttga tcctttagcc tatgagatct 1900

ggatgtgcat tgtttttgcc tacattgggg tcagtgtagt tttattcctg 1950

gtcagcagat ttagccccta cgagtggcac actgaggagt ttgaagatgg 2000

aagagaaaca caaagtagtg aatcaactaa tgaatttggg atttttaata 2050

gtctctggtt ttccttgggt gcctttatgc ggcaaggatg cgatatttcg 2100

ccaagatccc tctctgggcg cattgttgga ggtgtgtggt ggttctttac 2150

cctgatcata atctcctcct acacggctaa cttagctgcc ttcctgactg 2200

tagagaggat ggtgtctccc atcgaaagtg ctgaggatct ttctaagcaa 2250

acagaaattg cttatggaac attagactct ggctccacta aagagttttt 2300

caggagatct aaaattgcag tgtttgataa aatgtggacc tacatgcgga 2350

gtgcggagcc ctctgtgttt gtgaggacta cggccgaagg ggtggctaga 2400

gtgcggaagt ccaaagggaa atatgcctac ttgttggagt ccacgatgaa 2450

cgagtacatt gagcaaagga agccttgcga caccatgaaa gttggtggaa 2500

acctggattc caaaggctat ggcatcgcaa cacctaaagg atcctcatta 2550

ggaaccccag taaatcttgc agtattgaaa ctcagtgagc aaggcgtctt 2600

agacaagctg aaaaacaaat ggtggtacga taaaggtgaa tgtggagcca 2650

aggactctgg aagtaaggaa aagaccagtg ccctcagtct gagcaacgtt 2700

gctggagtat tctacatcct tgtcgggggc cttggtttgg caatgctggt 2750

ggctttgatt gagttctgtt acaagtcaag ggccgaggcg aaacgaatga 2800

aggtggcaaa gaatgcacag aatattaacc catcttcctc gcagaattca 2850

cagaattttg caacttataa ggaaggttac aacgtatatg gcatcgaaag 2900

tgttaaaatt taggggatga ccttgaatga tgccatgagg aacaaggcaa 2950

ggctgtcaat tacaggaagt actggagaaa atggacgtgt tatgactcca 3000

gaatttccca aagcagtgca tgctgtccct tacgtgagtc ctggcatggg 3050

aatgaatgtc agtgtgactg atctctcgtg attgataaga acctttgag 3100

tgccttacac aatggttttc ttgtgtgttt attgtcaaag tggtgagagg 3150

catccagtat cttgaagact tttctttcag ccaagaattc ttaaatatgt 3200

ggagttcatc ttgaattgta aggaatgatt aattaaaaca caacatcttt 3250

ttctactcga gttacagaca aagcgtggtg gacatgcaca gctaacatgg 3300

aagtactata atttacctga agtctttgta cagacaacaa acctgtttct 3350

gcagccacta ttgttagtct cttgattcat aatgacttaa gcacacttga 3400

catcaactgc atcaagatgt gacatgtttt ataaaaaaag gaaaaaaaac 3450

atttaaaact aaaaaatatt tttaggtatt ttcacaaaca aactggcttt 3500

taaataaatt tgcttccata ttggttgaat aagacaaaaa caattaaact 3550

gagtgggaag tgaataaaaa aaggctttag gtatcgattc catatttttc 3600

aaagccaaat atgtaaatgc taaggaaagt aaacaaagag gagattccaa 3650

tcttgtaatt taatattgtt attaaaactt taatgtatcc tattctttaa 3700

catttggtgt taatataaaa ttacttggca atgcttgaca tttgaaataa 3750

acatttttct attgttttat tgcaagtggt ccaattaatt ttgcttagct 3800

acagtttggt cataaatcaa gtgagtttaa agacactacc aagttgttag 3850

gtgcccagag aaaatttctc ccttttaaaa aggccaggtg atttttcaaa 3900

tgtaatcttg cccccaaagt aatatccgaa tatctttttg acatgtctaa 3950

atatatatat atataaagaa atatttgtta acacaaaagc atttgatcta 4000

tgtagataaa tgctaataga tttaaaaagc taatattaac aaataccaga 4050

atacgtgaag ttccattttt aaagtgtttg agcttacaga agagaaacat 4100

tcattttaaa tgaagtaaaa aatgccttga aagtaattct ttagatagtt 4150

gcccattgat taaattccaa aaactaaata tgtttttagc tttaaaatta 4200

taaaagctgt cataaacttt atatattatg aattttaaaa tatgtttgag 4250

tctcctgcaa tatagtttca tcccattgac atcaattaaa aataacccta 4300

atatattatt tttatattta ttcctcaggt ggaatggcta ttttaatatg 4350

cccagtgtgg ataaaatgtc acatttctgt aacttttgac taaagagcct 4400

atatttatct agttaatgaa tttaaaggat ctatctttcc cttcataaaa 4450

tacctcttat ttccattaaa gccccccaag tttaattaat ttaggatttt 4500

gaatgattat tgacatccaa tagttatttt taatatttgt attcttgtta 4550

tttctggaag aaagcctttg tgtagcactt ggtattttgc aaagtgcttt 4600

taaaacattc ttacttaccg tatttcatag aagggaagga aaaatgtaag 4650

gtttaacagt aagcacttgc attgaacatg gaggcatgtg gtatcatgat 4700

attcttcact aaatttagct gtccctaatc acagatccta aggtaatata 4750

atataatttt agtgcatttc tcctcatcag gaatgctgga ggtgcatttt 4800

aagttttaat aataagtgct agaatgacca aattgcagac taattgtttc 4850

catattgtac ttaaaatgag tttttaaaag tgaaaaagaa atgactatat 4900

acaatcaatg ctatttattg tacctctggg cctactcttc taaaaattgt 4950

agcttatcga ttttttctctg tcaagcttga actaatgtaa ataattgaaa 5000

taatgtaaag ttatattttc atgtttttat agatacaaca tgacaagaat 5050

acataatgta agagtatttc aactatggat aatgttgatt ggataatgca 5100

catctcagtt acaagcagta ctcatagttt aatatccatg taacggtgca 5150

tcaatatatt gctatataaa tatgtctgtg tgcatataag tgaaaagtgg 5200

tcaaacaaga gtgatgacag ctgtctaaag gttttttttat tcatttata 5250

taaaaactgt tatggaaaga ccaaaatgtt tatgaactat tcttatgtaa 5300

atttacaatt gtcctttact gtacttttt gtttacagta tagtaccttа 5350

ttttctgctg tgttaagtgg gtgtcaaact ccaagaagac atacactttc 5400

tataacttct attgaagata ttggaatttc caatttttca tgtgtactat 5450

gtcagaaaat gctttcgatt ttatttttaa atctaacatc ggatggcttt 5500

tccggagtgt tgtaaaaact tcaatcatac ataaaacatg ttcttacaaa 5550

aggcaaagat ctt 5563

<210> 64
<211> 1460
<212> DNA
<213> Homo sapiens

<400> 64
gccaacactg gccaaaaggg atagctgtcc aaggtctccc ccagcactga 50

ggagctcgcc tgctgccctc ttgcgcgcgg gaagcagcac caagttcacg 100

gccaacgcct tggcactagg gtccagaatg gctacaacag tccctgatgg 150

ttgccgcaat ggcctgaaat ccaagtacta cagactttgt gataaggctg 200

aagcttgggg catcgtccta gaaacggtgg ccacagccgg ggttgtgacc 250

tcggtggcct tcatgctcac tctcccgatc ctcgtctgca aggtgcagga 300

ctccaacagg cgaaaaatgc tgcctactca gtttctcttc ctcctgggtg 350

tgttgggcat ctttggcctc accttcgcct tcatcatcgg actggacggg 400

agcacagggc ccacacgctt cttcctcttt gggatcctct tttccatctg 450

cttctcctgc ctgctggctc atgctgtcag tctgaccaag ctcgtccggg 500

ggaggaagcc cctttccctg ttggtgattc tgggtctggc cgtgggcttc 550

agcctagtcc aggatgttat cgctattgaa tatattgtcc tgaccatgaa 600

taggaccaac gtcaatgtct tttctgagct ttccgctcct cgtcgcaatg 650

aagactttgt cctcctgctc acctacgtcc tcttcttgat ggcgctgacc 700

ttcctcatgt cctccttcac cttctgtggt tccttcacgg gctggaagag 750

acatggggcc cacatctacc tcacgatgct cctctccatt gccatctggg 800

tggcctggat caccctgctc atgcttcctg actttgaccg caggtgggat 850

gacaccatcc tcagctccgc cttggctgcc aatggctggg tgttcctgtt 900

ggcttatgtt agtcccgagt tttggctgct cacaaagcaa cgaaacccca 950

tggattatcc tgttgaggat gctttctgta aacctcaact cgtgaagaag 1000

agctatggtg tggagaacag agcctactct caagaggaaa tcactcaagg 1050

ttttgaagag acaggggaca cgctctatgc cccctattcc acacattttc 1100

agctgcagaa ccagcctccc caaaaggaat tctccatccc acgggcccac 1150

gcttggccga gcccttacaa agactatgaa gtaaagaaag agggcagcta 1200

actctgtcct gaagagtggg acaaatgcag ccgggcggca gatctagcgg 1250

gagctcaaag ggatgtgggc gaaatcttga gtcttctgag aaaactgtac 1300

aagacactac gggaacagtt tgcctccctc ccagcctcaa ccacaattct 1350

tccatgctgg ggctgatgtg ggctagtaag actccagttc ttagaggcgc 1400

tgtagtattt ttttttttt gtctcatcct ttagatactt cttttaagtg 1450

ggagtctcag 1460

<210> 65
<211> 1304
<212> DNA

<213> Homo sapiens

<220>
<221> unsure
<222> 1144-1145
<223> unknown base

<400> 65
tcattccatt aaacatattt ctaaataata gtaagtggta ctaacaaaat 50

aaataataat ttaatagcct tagaaataaa tgactgtata cttatacagg 100

ttgaaaaaaa ctcggtagga ataagttacc tttttgttta ctaatgttgg 150

tttcaaaaat actcagattc attttagttg gctgacatct ggaagtagtt 200

aacaactaac cagtggactt caacaatcat ttgctcccag gcttccccca 250

tcatcaccct caccacatat cctgctaata tccaacaaca aacaaatatt 300

taatattgaa atagcccatt gcctgagaat gaacacaagc taaaatacat 350

gcaagggtac ttaatggaag ccaaaccatg ttctatacct aaggagaaaa 400

catggacatg tagaatgctt ttattcatgt attcaaaatc agaacaaatc 450

agtgtatatc actagaacat cagatggagg ataacacaag aagtgataca 500

gatcagggtt cacttctctt accctctctc tgttaggacc acattcctat 550

tttagccaaa tgtttctggt acgggccatc ttttcaccat aaatggcatt 600

atgtttcaaa tggctaaaag cttattatgg gtatgctcaa aggagtaaaa 650

cccaattcac agagatgtgg cttttctaaa gaaccaaatt aagggaatac 700

ttagtatcta caactaaagt attagtcatt ctgaggatta tttgttgctt 750

taagtattag ctccctacta catccataac accctactta ctaaatttaa 800

ttacacacaa cttttcaaga attcgtattt tatttgaagg gaggtacctg 850

tctactttat ctacaataaa aacaaagggt attggctttc tctaatccat 900

gcaaactaca aattccatcg ggagtcctac atcactaaca gtggtatgaa 950

caaaactaag aaagtacttc ctatccacat gtgaatgttt taaaaaagtt 1000

tttgccataa aacctaagtg taatttagca taccacagtg ctctgaagat 1050

gggtcattga cgatgtacca tttgtatata ggtaatacac atgtaaatca 1100

ctaattgtta attaaatata agaagtaggt ttttatcttt ttannaaaaa 1150

aacaaaaaag gtgactccct ttctcattct gttctgttgt attgcgtcca 1200

aaagttgtgt gtaatttttt taaggtccaa gaaatgtaaa gaaatttgtt 1250

ggaatacctg aatttctgta aaaaaataaa aataaaaata agattttgtt 1300

actt 1304

<210> 66
<211> 3268
<212> DNA
<213> Homo sapiens

<400> 66
gggaggctgt gccaggcgag ccggaggggt gctccgcgct cccccgcccc 50

ccttccggga gcgaggatgc agactctgaa actggtgctg ctgggctgag 100

gcggaggcag gggagttgca gcgcgcgagg ctccgtgagt gtgtctcctg 150

cgcgctgaga ggcgggggga ggcggaggac caggaggagg aggaggagga 200

ggaggagggg gagaatgccc ggagccgccg ccgctgccgc cgccgccgcc 250

gccgcgatgc tcccggctca ggaggctgcc aagctgtacc acaccaacta 300

tgtgcggaac tcgcgggcca tcggcgtgct gtgggccatc ttcaccatct 350

gctttgccat cgtcaacgtg gtgtgcttca tccagcccta ctggataggc 400

gacggcgtgg acaccccgca agccggctat ttcgggctct tccactactg 450

catcggcaac ggcttctccc gggagctgac ctgcaggggc agcttcacgg 500

acttctccac gctgccctcg ggcgccttca aagccgcctc cttctttatc 550

ggcctctcca tgatgctcat cattgcctgc atcatttgct ttaccctctt 600

cttcttctgc aacacggcca ctgtgtacaa gatatgtgcc tggatgcagc 650

tcacctccgc tgcctgcctt gtgcttggct gtatgatttt ccctgatggc 700

tgggactcag atgaagtaaa acggatgtgt ggagaaaaga cagacaagta 750

cactcttggg gcttgctcag tccgctgggc atacatcctg gctattattg 800

gaattttgga tgccctgatc ctctcatttc tagcatttgt gcttggtaat 850

cgacaagaca gcttgatggc agaggaactg aaggcagaaa acaaagttct 900

gctaagccaa tattctctag aatgagcaca aaacaaatcg aataacagct 950

aaacaaatcg aataacagct aaacgaatcg aataacagct tttgtacatc 1000

aacatcaaga aggaatacgc ctgagagaga tcagagtata tagatgaata 1050

tgaacaagaa tggaacattc acttgtcaac gcactttcta aatctagatc 1100

agcagagatg ggagtgattt tctggaaaga gatgtgatca tggattaaac 1150

accagctcat tggaaactca ttggatgaga tcagaaaacg ttcatgaaaa 1200

atcatattca ggaaataagg aagaggaata taaatgctct agagttaaca 1250

tgtaaaatat atacgtactg aggtttgtaa actgtccttt ttaaatcaaa 1300

ctgaaaacaa aaagctttaa cctttcaaca gaatttttaa aaaggcagtt 1350

agttctaaat tattcctatc tcaatagcca agaggctgat caagcgtcat 1400

ttattgagga agcatcttag aaaatgcctc tgaatgtttt cataggagcc 1450

gtgacctttg gttcttcatc tctaccattc atttacttca ctgtgtaatt 1500

agttacaacc actcagttat taagagacgt aacgcttcaa acttttacc 1550

aagtctgtgt tctgtttaat ctgtccatac aagttattac tgagaaagtg 1600

tttatgccat atactattac tccatcaagc tgtatattac aggaagtaca 1650

tctttacatc ataggttccc aagcaacata gatttcccta tctttcagga 1700

aacagcatca aggaactctg aaaaatatag aaaaagttca ttttcacctt 1750

ggaagctcac gtgtaatatt ataggctact atcaaataaa cacttttttt 1800

ctaattctcc ctagtatatg cataggaatt taatatactt tataaataag 1850

tatctaaaat gtctcctact tttttcctat ttctttgcca tacatgttat 1900

cagaaatcca tgtcttctat ttcccttact gatgggcact catttttatt 1950

tttttaaaaa tcattccatt aaacatattt ctaaataata gtaagtggta 2000

ctaacaaaat aaataataat ttaatagcct tagaaataaa tgactgtata 2050

cttatacagg ttgaaaaaaa ctcggtagga ataagttacc tttttgttta 2100

ctaatgttgg tttcaaaaat actcagattc attttagttg gctgacatct 2150

ggaagtagtt aacaactaac cagtggactt caacaatcat ttgctcccag 2200

gcttccccca tcatcaccct caccacatat cctgctaata tccaacaaca 2250

aacaaatatt taatattgaa atagcccatt gcctgagaat gaacacaagc 2300

taaaatacat gcaagggtac ttaatggaag ccaaaccatg ttctatacct 2350

aaggagaaaa catggacatg tagaatgctt ttattcatgt attcaaaatc 2400

agaacaaatc agtgtatatc actagaacat cagatggagg ataacacaag 2450

aagtgataca gatcagggtt cacttctctt accctctctc tgttaggacc 2500

acattcctat tttagccaaa tgtttctggt acgggccatc ttttcaccat 2550

aaatggcatt atgtttcaaa tggctaaaag catattatgg gtatgctcaa 2600

aggagtaaaa cccaattcac agagatgtgg cttttctaaa gaaccaaatt 2650

aagggaatac ttagtatcta caactaaagt attagtcatt ctgaggatta 2700

tttgttgctt taagtattag ctccctacta catccataac accctactta 2750

ctaaatttaa ttacacacaa cttttcaaga attcgtattt tatttgaagg 2800

gaggtacctg tctactttat ctacaataaa aacaaagggt attggctttc 2850

tctaatccat gcaaactaca aattccatcg ggagtcctac atcactaaca 2900

gtggtatgaa caaaactaag aaagtacttc ctatccacat gtgaatgttt 2950

taaaaaagtt tttgccataa aacctaagtg taatttagca taccacagtg 3000

ctctgaagat gggtcattga cgatgtacca tttgtatata ggtaatacac 3050

atgtaaatca ctaattgtta attaaatata agaagtaggt ttttatcttt 3100

ttaaaaaaaa aacaaaaaag gtgactccct ttctcattct gttctgttgt 3150

attgcgtcca aaagttgtgt gtaatttttt taaggtccaa gaaatgtaaa 3200

gaaatttgtt ggaatacctg aatttctgta aaaaaataaa aataaaaata 3250

agattttgtt acttaaaa 3268

<210> 67
<211> 3268
<212> DNA
<213> Homo sapiens

<400> 67
gggaggctgt gccaggcgag ccggaggggt gctccgcgct cccccgcccc 50

ccttccggga gcgaggatgc agactctgaa actggtgctg ctgggctgag 100

gcggaggcag gggagttgca gcgcgcgagg ctccgtgagt gtgtctcctg 150

cgcgctgaga ggcgggggga ggcggaggac caggaggagg aggaggagga 200

ggaggagggg gagaatgccc ggagccgccg ccgctgccgc cgccgccgcc 250

gccgcgatgc tcccggctca ggaggctgcc aagctgtacc acaccaacta 300

tgtgcggaac tcgcgggcca tcggcgtgct gtgggccatc ttcaccatct 350

gctttgccat cgtcaacgtg gtgtgcttca tccagcccta ctggataggc 400

gacggcgtgg acaccccgca agccggctat ttcgggctct tccactactg 450

catcggcaac ggcttctccc gggagctgac ctgcaggggc agcttcacgg 500

acttctccac gctgccctcg ggcgccttca aagccgcctc cttctttatc 550

ggcctctcca tgatgctcat cattgcctgc atcatttgct ttaccctctt 600

cttcttctgc aacacggcca ctgtgtacaa gatatgtgcc tggatgcagc 650

tcacctccgc tgcctgcctt gtgcttggct gtatgatttt ccctgatggc 700

tgggactcag atgaagtaaa acggatgtgt ggagaaaaga cagacaagta 750

cactcttggg gcttgctcag tccgctgggc atacatcctg gctattattg 800

gaattttgga tgccctgatc ctctcatttc tagcatttgt gcttggtaat 850

cgacaagaca gcttgatggc agaggaactg aaggcagaaa acaaagttct 900

gctaagccaa tattctctag aatgagcaca aaacaaatcg aataacagct 950

aaacaaatcg aataacagct aaacgaatcg aataacagct tttgtacatc 1000

aacatcaaga aggaatacgc ctgagagaga tcagagtata tagatgaata 1050

tgaacaagaa tggaacattc acttgtcaac gcactttcta aatctagatc 1100

agcagagatg ggagtgattt tctggaaaga gatgtgatca tggattaaac 1150

accagctcat tggaaactca ttggatgaga tcagaaaacg ttcatgaaaa 1200

atcatattca ggaaataagg aagaggaata taaatgctct agagttaaca 1250

tgtaaaatat atacgtactg aggtttgtaa actgtccttt ttaaatcaaa 1300

ctgaaaacaa aaagctttaa cctttcaaca gaatttttaa aaaggcagtt 1350

agttctaaat tattcctatc tcaatagcca agaggctgat caagcgtcat 1400

ttattgagga agcatcttag aaaatgcctc tgaatgtttt cataggagcc 1450

gtgacctttg gttcttcatc tctaccattc atttacttca ctgtgtaatt 1500

agttacaacc actcagttat taagagacgt aacgcttcaa acttttacc 1550

aagtctgtgt tctgtttaat ctgtccatac aagttattac tgagaaagtg 1600

tttatgccat atactattac tccatcaagc tgtatattac aggaagtaca 1650

tctttacatc ataggttccc aagcaacata gatttcccta tctttcagga 1700

aacagcatca aggaactctg aaaaatatag aaaaagttca ttttcacctt 1750

ggaagctcac gtgtaatatt ataggctact atcaaataaa cacttttttt 1800

ctaattctcc ctagtatatg cataggaatt taatatactt tataaataag 1850

tatctaaaat gtctcctact tttttcctat ttctttgcca tacatgttat 1900

cagaaatcca tgtcttctat ttcccttact gatgggcact cattttatt 1950

tttttaaaaa tcattccatt aaacatattt ctaaataata gtaagtggta 2000

ctaacaaaat aaataataat ttaatagcct tagaaataaa tgactgtata 2050

cttatacagg ttgaaaaaaa ctcggtagga ataagttacc tttttgttta 2100

ctaatgttgg tttcaaaaat actcagattc attttagttg gctgacatct 2150

ggaagtagtt aacaactaac cagtggactt caacaatcat ttgctcccag 2200

gcttccccca tcatcaccct caccacatat cctgctaata tccaacaaca 2250

aacaaatatt taatattgaa atagcccatt gcctgagaat gaacacaagc 2300

taaaatacat gcaagggtac ttaatggaag ccaaaccatg ttctatacct 2350

aaggagaaaa catggacatg tagaatgctt ttattcatgt attcaaaatc 2400

agaacaaatc agtgtatatc actagaacat cagatggagg ataacacaag 2450

aagtgataca gatcagggtt cacttctctt accctctctc tgttaggacc 2500

acattcctat tttagccaaa tgtttctggt acgggccatc ttttcaccat 2550

aaatggcatt atgtttcaaa tggctaaaag catattatgg gtatgctcaa 2600

aggagtaaaa cccaattcac agagatgtgg cttttctaaa gaaccaaatt 2650

aagggaatac ttagtatcta caactaaagt attagtcatt ctgaggatta 2700

tttgttgctt taagtattag ctccctacta catccataac accctactta 2750

ctaaatttaa ttacacacaa cttttcaaga attcgtattt tatttgaagg 2800

gaggtacctg tctactttat ctacaataaa aacaaagggt attggctttc 2850

tctaatccat gcaaactaca aattccatcg ggagtcctac atcactaaca 2900

gtggtatgaa caaaactaag aaagtacttc ctatccacat gtgaatgttt 2950

taaaaaagtt tttgccataa aacctaagtg taatttagca taccacagtg 3000

ctctgaagat gggtcattga cgatgtacca tttgtatata ggtaatacac 3050

atgtaaatca ctaattgtta attaaatata agaagtaggt ttttatcttt 3100

ttaaaaaaaa aacaaaaaag gtgactccct ttctcattct gttctgttgt 3150

attgcgtcca aaagttgtgt gtaatttttt taaggtccaa gaaatgtaaa 3200

gaaatttgtt ggaatacctg aatttctgta aaaaaataaa aataaaaata 3250

agattttgtt acttaaaa 3268

<210> 68
<211> 3268
<212> DNA
<213> Homo sapiens

<400> 68
gggaggctgt gccaggcgag ccggaggggt gctccgcgct cccccgccct 50

ccttccggga gcgaggatgc agactctgaa actggtgctg ctgggctgag 100

gcggaggcag gggagttgca gcgcgcgagg ctccgtgagt gtgtctcctg 150

cgcgctgaga ggcgggggga ggcggaggac caggaggagg aggaggagga 200

ggaggagggg gagaatgccc ggagccgccg ccgctgccgc cgccgccgcc 250

gccgcgatgc tcccggctca ggaggctgcc aagctgtacc acaccaacta 300

tgtgcggaac tcgcgggcca tcggcgtgct gtgggccatc ttcaccatct 350

gctttgccat cgtcaacgtg gtgtgcttca tccagcccta ctggataggc 400

gacggcgtgg acaccccgca agccggctat ttcgggctct tccactactg 450

catcggcaac ggcttctccc gggagctgac ctgcaggggc agcttcacgg 500

acttctccac gctgccctcg ggcgccttca aagccgcctc cttctttatc 550

ggcctctcca tgatgctcat cattgcctgc atcatttgct ttaccctctt 600

cttcttctgc aacacggcca ctgtgtacaa gatatgtgcc tggatgcagc 650

tcacctccgc tgcctgcctt gtgcttggct gtatgatttt ccctgatggc 700

tgggactcag atgaagtaaa acggatgtgt ggagaaaaga cagacaagta 750

cactcttggg gcttgctcag tccgctgggc atacatcctg gctattattg 800

gaattttgga tgccctgatc ctctcatttc tagcatttgt gcttggtaat 850

cgacaagaca gcttgatggc agaggaactg aaggcagaaa acaaagttct 900

gctaagccaa tattctctag aatgagcaca aaacaaatcg aataacagct 950

aaacaaatcg aataacagct aaacgaatcg aataacagct tttgtacatc 1000

aacatcaaga aggaatacgc ctgagagaga tcagagtata tagatgaata 1050

tgaacaagaa tggaacattc acttgtcaac gcactttcta aatctagatc 1100

agcagagatg ggagtgattt tctggaaaga gatgtgatca tggattaaac 1150

accagctcat tggaaactca ttggatgaga tcagaaaacg ttcatgaaaa 1200

atcatattca ggaaataagg aagaggaata taaatgctct agagttaaca 1250

tgtaaaatat atacgtactg aggtttgtaa actgtccttt ttaaatcaaa 1300

ctgaaaacaa aaagctttaa cctttcaaca gaatttttaa aaaggcagtt 1350

agttctaaat tattcctatc tcaatagcca agaggctgat caagcgtcat 1400

ttattgagga agcatcttag aaaatgcctc tgaatgtttt cataggagcc 1450

gtgacctttg gttcttcatc tctaccattc atttacttca ctgtgtaatt 1500

agttacaacc actcagttat taagagacgt aacgcttcaa acttttacc 1550

aagtctgtgt tctgtttaat ctgtccatac aagttattac tgagaaagtg 1600

tttatgccat atactattac tccatcaagc tgtatattac aggaagtaca 1650

tctttacatc ataggttccc aagcaacata gatttcccta tctttcagga 1700

aacagcatca aggaactctg aaaaatatag aaaaagttca ttttcacctt 1750

ggaagctcac gtgtaatatt ataggctact atcaaataaa cactttttt 1800

ctaattctcc ctagtatatg cataggaatt taatatactt tataaataag 1850

tatctaaaat gtctcctact tttttcctat ttctttgcca tacatgttat 1900

cagaaatcca tgtcttctat ttcccttact gatgggcact catttttatt 1950

tttttaaaaa tcattccatt aaacatattt ctaaataata gtaagtggta 2000

ctaacaaaat aaataataat ttaatagcct tagaaataaa tgactgtata 2050

cttatacagg ttgaaaaaaa ctcggtagga ataagttacc tttttgttta 2100

ctaatgttgg tttcaaaaat actcagattc attttagttg gctgacatct 2150

ggaagtagtt aacaactaac cagtggactt caacaatcat ttgctcccag 2200

gcttccccca tcatcaccct caccacatat cctgctaata tccaacaaca 2250

aacaaatatt taatattgaa atagcccatt gcctgagaat gaacacaagc 2300

taaaatacat gcaagggtac ttaatggaag ccaaaccatg ttctatacct 2350

aaggagaaaa catggacatg tagaatgctt ttattcatgt attcaaaatc 2400

agaacaaatc agtgtatatc actagaacat cagatggagg ataacacaag 2450

aagtgataca gatcagggtt cacttctctt accctctctc tgttaggacc 2500

acattcctat tttagccaaa tgtttctggt acgggccatc ttttcaccat 2550

aaatggcatt atgtttcaaa tggctaaaag catattatgg gtatgctcaa 2600

aggagtaaaa cccaattcac agagatgtgg cttttctaaa gaaccaaatt 2650

aagggaatac ttagtatcta caactaaagt attagtcatt ctgaggatta 2700

tttgttgctt taagtattag ctccctacta catccataac accctactta 2750

ctaaatttaa ttacacacaa cttttcaaga attcgtattt tatttgaagg 2800

gaggtacctg tctactttat ctacaataaa aacaaagggt attggctttc 2850

tctaatccat gcaaactaca aattccatcg ggagtcctac atcactaaca 2900

gtggtatgaa caaaactaag aaagtacttc ctatccacat gtgaatgttt 2950

taaaaaagtt tttgccataa aacctaagtg taatttagca taccacagtg 3000

ctctgaagat gggtcattga cgatgtacca tttgtatata ggtaatacac 3050

atgtaaatca ctaattgtta attaaatata agaagtaggt ttttatcttt 3100

ttaaaaaaaa aacaaaaaag gtgactccct ttctcattct gttctgttgt 3150

attgcgtcca aaagttgtgt gtaatttttt taaggtccaa gaaatgtaaa 3200

gaaatttgtt ggaatacctg aatttctgta aaaaaataaa aataaaaata 3250

agattttgtt acttaaaa 3268

<210> 69
<211> 1755
<212> DNA
<213> Homo sapiens

<400> 69
caggcctctg aggctccctt gccgagggcc ccgagctgca gggacagtga 50

gcagtgagtc ccttgggcat cccgctcctg ggcaggtcac caataggtcc 100

ccgcagttcc caatggaact gttccagtcc tccccgaggc ctccacttca 150

acctgtctgt gtctgcccag gcctggagtt gtgtgaccct ccccaccgcc 200

tggccttctc catgggggct ggccttttct cggtggtggg caccctgctg 250

ctgcccggcc tggctgcgct tgtgcaggac tggcgtcttc tgcaggggct 300

gggtgccctg atgagtggac tcttgctgct cttttggggg ttcccggccc 350

tgttccccga gtctccctgc tggctgctgg ccacaggtca ggtagctcga 400

gccaggaaga tcctgtggcg ctttgcagaa gccagtggcg tgggccccgg 450

ggacagttcc ttggaggaga actccctggc tacagagctg accatgctgt 500

ctgcacggag ccccccagccc cggtaccact ccccactggg gcttctgcgt 550

acccgagtca cctggagaaa cgggcttatc ttgggcttca gctcgctggt 600

tggtggaggc atcagagcta gcttccgccg cagcctggca cctcaggtgc 650

cgaccttcta cctgccctac ttcctggagg ccggcctgga ggcggcagcc 700

ttggtcttcc tgctcctgac ggcagattgc tgtggacgcc gccccgtgct 750

gctgctgggc accatggtca caggcctggc atccctgctg ctcctcgctg 800

gggcccagta tctgccaggc tggactgtgc tgttcctctc tgtcctgggg 850

ctcctggcct cccgggctgt gtccgcactc agcagcctct tcgcggccga 900

ggtcttcccc acggtgatca gggggccgg gctgggcctg gtgctggggg 950

ccgggttcct gggccaggca gccggccccc tggacaccct gcacggccgg 1000

cagggcttct tcctgcaaca agtcgtcttc gcctcccttg ctgtccttgc 1050

cctgctgtgt gtcctgctgc tgcctgagag ccgaagccgg gggctgcccc 1100

agtcactgca ggacgccgac cgcctgcgcc gctccccact cctgcggggc 1150

cgccccgcc aggaccacct gcctctgctg ccgccctcca actcctactg 1200

ggccggccac acccccgagc agcactagtc ctgcctggtg gccctgggag 1250

ccaggatggg accaaagtca aggcctgggg catggctgag taccccagac 1300

gtctggtcca gggcagacac attcctctca gaagcccgtg tctcagtgca 1350

ggtggagccg tggggacagc gtgaaggtgt ctccagccag gccccaggca 1400

ctgggaggcc ctgggtctcc ccccagccac acccagtagg tgtggaggat 1450

aaaggcttct gtggaacttg gcgtcctccc tggttcttcc ctcaggccac 1500

acctggggcc tggaagggtc tcacctgtgc ctgctcccca cctctgggct 1550

caccagcaag gagccagggt gtcagcggct gctccatagc ccctgtccag 1600

gtagtggcag cccccggccg cctcctgccc cagccaagca gggtggtatg 1650

ggggcctggc cctcaggtcc ctgccaggcc aggcctcact ggccccatcc 1700

ccctcatggt cagtgtggca gtgggcgccc gtcctgcatg aaagagacta 1750

gcacg 1755

<210> 70
<211> 4021
<212> DNA
<213> Homo sapiens

<400> 70
agcggggcgg gtccccgaag cccagtacca gacctcctgt gcgaagagta 50

acctgctctt cggactgata actggatgga atggccttta gaaagggagc 100

atcttcagaa tctgcagctg ggatcagatt gttcttttgt tgaatcaact 150

cagctctgcc gagccagcat cttgctaagg gttacaaaat aaatgccatt 200

tgaacagtgc catctgtcat ggaaaaacgt gagacgtttg tacaatgccg 250

tgtctaagga gctggttgga gagtttttgc aatttgttca acttgataaa 300

gaggcctctg atcctttcag cctaaatgaa ttactagatg aattatcaag 350

gaaacagaaa gaagaattat ggcaaaggct gaagaattta ttgacagatg 400

tgttgttaga aagcccagtg gatgggtggc aggtagtgga agcccagggt 450

gaagacaata tggaaaccga acatggctca aaaatgagaa aaagcataga 500

aataatttat gcaattacat ctgtgattct tgcttctgtg tctgtaataa 550

atgaaagtga gaactacgaa gccctactgg aatgtgttat tatattaaat 600

ggtattttat atgcattacc tgagtctgaa cgaaaactac agagttctat 650

tcaggatttg tgtgttacct ggtgggagaa aggcctgcct gccaaggaag 700

acacaggaaa gacagccttt gtcatgttac taaggaggag tctggagact 750

aagacaggtg cagacgtatg tcggctttgg cgtatccatc aagctttata 800

ttgctttgat tatgatttgg aggaaagtgg agaaattaaa gatatgttac 850

ttgagtgctt cataaatatt aattatatca agaaagaaga gggaagaaga 900

tttcttagtt gtctcttcaa ctggaatatc aacttcatca aaatgatcca 950

cgggaccatt aaaaaccagt tacagggatt acaaaagtct ttgatggtat 1000

acattgcaga aatttatttc agagcttgga aaaaggcttc agggaaaata 1050

ctggaggcga ttgaaaatga ttgcatccag gacttcatgt tccacgggat 1100

acaccttccg aggaggtctc cagtgcattc caaagtgcgg gaggttctga 1150

gttactttca ccatcaaaag aaagttcggc agggagtgga agagatgctt 1200

1125
1174

tatagattat ataagcccat cctttggaga ggattaaagg ccagaaactc 1250

tgaagttcga tcaaatgctg cattgttgtt tgttgaagca tttcctatta 1300

gggatccaaa ccttcatgct attgaaatgg atagtraaat ccagaaacag 1350

tttgaagagc tctatagcct tttagaagat ccttacccga tggtccgttc 1400

cacagggatc cttggtgttt gtaaaataac ttctaagtac tgggaaatga 1450

tgccccccgac cattcttatt gacctcctga agaaggtgac tggggaactg 1500

gcatttgaca cgagctcagc tgatgttcgt tgttctgtct ttaagtgtct 1550

gccaatgatt ttggacaaca aactgagcca cccattgtta gagcagctcc 1600

ttccagctct cagatacagt ctccacgaca attcggagaa agtgagggta 1650

gcttttgtgg acatgctgtt gaagatcaaa gctgtgaggg ctgctaagtt 1700

ttggaaaata tgtcccatgg agcacattct ggttcgtctg gaaactgatt 1750

ctcgacctgt gtctcggcgc ctggtgagcc tcatctttaa ttctttcctg 1800

cctgtgaatc agccggagga ggtctggtgc gagcgctgtg tcaccctggt 1850

gcagatgaac cacgccgctg ccaggaggtt ctatcagtac gcccacgaac 1900

acaccgcctg caccaacata gcaaagctga ttcacgttat tcgtcattgc 1950

ttaaatgcct gtatccagag ggcagtgaga gagcctccag aggacgagga 2000

ggaagaggac ggaagggaga aggagaatgt gactgttctg gacaaaacac 2050

tgtcagtaaa cgatgttgca tgcatggcag gtttactaga aatcattgtg 2100

attctctgga aaagtattga cagaagtatg gaaaataata aagaggccaa 2150

actttacacg attaacaagt ttgcctctgt gcttccagag tatctgaaag 2200

tatttaagga tgatcgctgc aagatccctt tattcatgct aatgtccttt 2250

atgccggcct ctgctgtccc cccattcagc tgtggtgtga tttccacgct 2300

gagaagccgg gaggagggcg ctgtggacaa gagctactgc actttgttgg 2350

attgcctctg ctcctggggg caggtggggc acattctgga gcttgttgac 2400

aactggctgc ccacagagca tgcccaggcc aagagcaaca cagcttctaa 2450

aggtagggtg cagatccatg acacacgccc agtcaaacct gaattggcat 2500

tggtctacat tgagtatctg ctgactcatc caaagaaccg cgagtgcttg 2550

ctctctgctc ctcggaagaa acttaaccat cttttgaaag cccttgaaac 2600

gtcaaaggca gatctggagt cacttctgca gacaccgggt gggaagcctc 2650

gtggcttcag tgaagcagct gccccgcgag cctttggtct ccactgtcgc 2700

ctgagcatcc atcttcagca caagttctgc tcagaaggaa aggtgtattt 2750

gtccatgttg gaagacactg gcttttggtt agaaagcaaa attttatctt 2800

ttattcaaga tcaagaagaa gactacctga agcttcatag ggtcatttat 2850

cagcaaatta tccagaccta cctgactgtg tgtaaagatg ttgttatggt 2900

aggccttggt gaccatcagt ttcagatgca actcttacag cggagtcttg 2950

gaatcatgca aacagtgaag ggatttttt atgtttcatt acttcttgac 3000

attctgaaag agataactgg aagttccttg attcagaaaa cagattcaga 3050

tgaagaagtt gcaatgctgt tggacacagt ccagaaagta tttcagaaaa 3100

tgttggaatg tattgcacgg agcttcagga agcagccgga agaaggcctg 3150

cggctgcttt attctgttca gaggcctctt catgagttca ttactgctgt 3200

tcagtctcgg cacacagaca cccctgtgca ccggggtgta ctttctactc 3250

tgatcgctgg gcctgtggtt gagataagtc accagctacg gaaggtttct 3300

gacgtagaag agcttacccc tccagagcat ctttctgatc ttccaccatt 3350

ttcaaggtgt ttaataggaa taataataaa gtcttcgaat gtggtcaggt 3400

catttttgga tgaattaaag gcatgtgtgg cttctaatga tattgaaggc 3450

attgtgtgcc tcacggctgc tgtgcatatt atcctggtta ttaatgcagg 3500

taaacataaa agctcaaaag tgagggaggt tgcagccact gttcacagaa 3550

aactaaagac attcatggaa attactttgg aagaggatag cattgaaaga 3600

tttctctatg aatcatcatc aagaactctg ggagaacttt tgaattcata 3650

accaagccaa catctccaga catgtaaaaa tagggaaaag tgattcaaat 3700

tgaaatgcct gtgtattttc ctattgtttt taatgttaat aacccatata 3750

atagggaaag ggtgggattt ttttgtggga atgtgggaag gtgggggtta 3800

tggaggagat aactcaaaac ttcttcaatt ttgcctagtg cctgcgtaaa 3850

taatatattt aatataaagg actccaggta tgaatggtgt agaaatccat 3900

gattccaaga aaaaacactt ttctagcaaa cctggttgtt tttaaaatga 3950

cttttatata tgtaatattg cttggaaact atgagtaata aagcaatgac 4000

aacaamaaaa aaaaaaaaaa a 4021

<210> 71
<211> 3580
<212> DNA
<213> Homo sapiens

<400> 71
gcccttcaga atctgcagct gggatcagat tgttcttttg ttgaatcaac 50

tcagctctgc cgagccagca tcttgctaag ggttacaaaa taaatgccat 100

ttgaacagtg ccatctgtca tggaaaaacg tgagacgttt gtacaagccg 150

tgtctaagga gctggttgga gagtttttgc aatttgttca acttgataaa 200

gaggcctctg atcctttcag cctaaatgaa ttactagatg aattatcaag 250

gaaacagaaa gaagaattat ggcaaaggct gaagaattta ttgacagatg 300

tgttgttaga aagcccagtg gatgggtggc aggtagtgga agcccagggt 350

gaagacaata tggaaaccga acatggctca aaaatgagaa aaagcataga 400

aataatttat gcaattacat ctgtgattct tgcttctgtg tctgtaataa 450

atgaaagtga gaactacgaa gccctactgg aatgtgttat tatattaaat 500

ggtattttat atgcattacc tgagtctgaa cgaaaactac agagttctat 550

tcaggatttg tgtgttacct ggtgggagaa aggcctgcct gccaaggaag 600

acacaggaaa gacagccttt gtcatgttac taaggaggag tctggagact 650

aagacaggtg cagacgtatg tcggctttgg cgtatccatc aagctttata 700

ttgctttgat tatgatttgg aggaaagtgg agaaattaaa gatatgttac 750

ttgagtgctt cataaatatt aattatatca agaaagaaga gggaagaaga 800

tttcttagtt gtctcttcaa ctggaatatc aacttcatca aaatgatcca 850

cgggaccatt aaaaaccagt tacagggatt acaaaagtct ttgatggtat 900

acattgcaga aatttatttc agagcttgga aaaaggcttc agggaaaata 950

ctggaggcga ttgaaaatga ttgcatccag gacttcatgt tccacgggat 1000

acaccttccg aggaggtctc cagtgcattc caaagtgcgg gaggttctga 1050

gttactttca ccatcaaaag aaagttcggc agggagtgga agagatgctt 1100

tatagattat ataagcccat cctttggaga ggattaaagg ccagaaactc 1150

tgaagttcga tcaaatgctg cattgttgtt tgttgaagca tttcctatta 1200

gggatccaaa ccttcatgct attgaaatgg atagtgaaat ccagaaacag 1250

tttgaagagc tctatagcct tttagaagat ccttacccga tggtccgttc 1300

cacagggatc cttggtgttt gtaaaataac ttctaagtac tgggaaatga 1350

tgcccccgac cattcttatt gacctcctga agaaggtgac tggggaactg 1400

gcatttgaca cgagctcagc tgatgttcgt tgttctgtct ttaagtgtct 1450

gccaatgatt ttggacaaca aactgagcca cccattgtta gagcagctcc 1500

ttccagctct cagatacagt ctccacgaca attcggagaa agtgagggta 1550

gcttttgtgg acatgctgtt gaagatcaaa gctgtgaggg ctgctaagtt 1600

ttggaaaata tgtcccatgg agcacattct ggttcgtctg gaaactgatt 1650

ctcgacctgt gtctcggcgc ctggtgagcc tcatctttaa ttctttcctg 1700

cctgtgaatc agccggagga ggtctggtgc gagcgctgtg tcaccctggt 1750

gcagatgaac cacgccgctg ccaggaggtt ctatcagtac gcccacgaac 1800

acaccgcctg caccaacata gcaaagctga ttcacgttat tcgtcattgc 1850

ttaaatgcct gtatccagag ggcagtgaga gagcctccag aggacgagga 1900

ggaagaggac ggaagggaga aggagaatgt gactgttctg gacaaaacac 1950

tgtcagtaaa cgatgttgca tgcatggcag gtttactaga aatcattgtg 2000

attctctgga aaagtattga cagaagtatg gaaaataata aagaggccaa 2050

actttacacg attaacaagt ttgcctctgt gcttccagag tatctgaaag 2100

tatttaagga tgatcgctgc aagatccctt tattcatgct aatgtccttt 2150

atgccggcct ctgctgtccc cccattcagc tgtggtgtga tttccacgct 2200

gagaagccgg gaggagggcg ctgtggacaa gagctactgc actttgttgg 2250

attgcctctg ctcctggggg caggtggggc acattctgga gcttgttgac 2300

aactggctgc ccacagagca tgcccaggcc aagagcaaca cagcttctaa 2350

aggtagggtg cagatccatg acacacgccc agtcaaacct gaattggcat 2400

tggtctacat tgagtatctg ctgactcatc caaagaaccg cgagtgcttg 2450

ctctctgctc ctcggaagaa acttaaccat cttttgaaag cccttgaaac 2500

gtcaaaggca gatctggagt cacttctgca gacaccgggt gggaagcctc 2550

gtggcttcag tgaagcagct gccccgcgag cctttggtct ccactgtcgc 2600

ctgagcatcc atcttcagca caagttctgc tcagaaggaa aggtgtattt 2650

gtccatgttg gaagacactg gcttttggtt agaaagcaaa attttatctt 2700

ttattcaaga tcaagaagaa gactacctga agcttcatag ggtcatttat 2750

cagcaaatta tccagaccta cctgactgtg tgtaaagatg ttgttatggt 2800

aggccttggt gaccatcagt ttcagatgca actcttacag cggagtcttg 2850

gaatcatgca aacagtgaag ggatttttt atgtttcatt acttcttgac 2900

attctgaaag agataactgg aagttccttg attcagaaaa cagattcaga 2950

tgaagaagtt gcaatgctgt tggacacagt ccagaaagta tttcagaaaa 3000

tgttggaatg tattgcacgg agcttcagga agcagccgga agaaggcctg 3050

cggctgcttt attctgttca gaggcctctt catgagttca ttactgctgt 3100

tcagtctcgg cacacagaca cccctgtgca ccggggtgta ctttctactc 3150

tgatcgctgg gcctgtggtt gagataagtc accagctacg gaaggtttct 3200

gacgtagaag agcttacccc tccagagcat ctttctgatc ttccaccatt 3250

ttcaaggtgt ttaataggaa taataataaa gtcttcgaat gtggtcaggt 3300

catttttgga tgaattaaag gcatgtgtgg cttctaatga tattgaaggc 3350

attgtgtgcc tcacggctgc tgtgcatatt atcctggtta ttaatgcagg 3400

taaacataaa agctcaaaag tgagggaggt tgcagccact gttcacagaa 3450

aactaaagac attcatggaa attactttgg aagaggatag cattgaaaga 3500

tttctctatg aatcatcatc aagaactctg ggagaacttt tgaattcata 3550

accaagccaa catctccaga catgaagggc 3580

<210> 72
<211> 1191
<212> DNA
<213> Homo sapiens

<400> 72
tccaaccaga ttttcaccgc tatgcctcct ttcactctgg gagtcttcca 50

gaggtcttgc actcgggaga gcatgctcag gtttccccag ctctacaaaa 100

tcacccagaa tgccaaagac ttcaacgtga ccagtggtca tgccactgcc 150

tgttgatttg ttgaaaatat tgtttacacg tatgttcttg ttactgattg 200

tcagaaagct ggttttgaga ctgcagcttg gactaaattc agtcatctgg 250

ctgtctgggg aagcatgctg accagtctgg tgttctttgg catctactca 300

gccatctggt ccaccattct cattgcccca aatatgagag gacagaagaa 350

tggtaccggt actgccaatg gagatggagg aaggagacag aaagaaacag 400

agcccagacc ctagggacca ccagcatttg cagaatggat aaacagcctt 450

cttcctaaca aaggaagcac agcaactgtg atcctgagct gtgcacactt 500

ctggttggga ttatttctgg tttctacttc ctgtttgaag atgtggcatg 550

gagagtgaac aagctgctgc ccaccacctg gcatcacagc cccagaactc 600

agctatttcc atgggaccac agcatctcat ctctgggctg agccagaaag 650

acccctactg aagtccagag gcacttttct gaaaggctct gctttgacct 700

gaagtatttt atctatcctc agtctcagga cactgttgat ggaattaagg 750

ccaagcacat ctgcaaaaaa gacattgctg gaggaggtgc aaagagctgg 800

aaaccaagtc tccagtcctg ggaaaagcag tggtatggaa aagcaatgga 850

aagagcattt tgaaaatgcc attccactgt tttctggcct ttatgatttc 900

tgctgagaaa tccactgtta gtctgatggg gtctccttca tagcaccaat 950

gacctgaaga gccttgttga aggaagactc catctgatga ctcagagcaa 1000

gtatttttta gtgtgttatt gttattagca gaaagagggc cataaaatac 1050

atggggcaag ctgaatatat cttaggcaaa agaagaaaat attcaaattc 1100

ttatgttatt ttatctaatt attttatctc tttttgtgtg tgacttataa 1150

tgtgtgtatt gtattaataa aagtatataa acatgtagtt t 1191

<210> 73
<211> 2752
<212> DNA
<213> Homo sapiens

<400> 73
cattcccacc cttccttctt taataagcag gagcgaaaaa gacaaattcc 50

aaagaggatt gttcagttca agggaatgaa gaattcagaa taattttggt 100

aaatggattc caatatcggg aataagaata agctgaacag ttgacctgct 150

ttgaagaaac atactgtcca tttgtctaaa ataatctata acaaccaaac 200

caatcaaaat gaattcaaca ttattttccc aggttgaaaa tcattcagtc 250

cactctaatt tctcagagaa gaatgcccag cttctggctt ttgaaaatga 300

tgattgtcat ctgcccttgg ccatgatatt taccttagct cttgcttatg 350

gagctgtgat cattcttggt gtctctggaa acctggcctt gatcataatc 400

atcttgaaac aaaaggagat gagaaatgtt accaacatcc tgattgtgaa 450

cctttccttc tcagacttgc ttgttgccat catgtgtctc ccctttacat 500

ttgtctacac attaatggac cactgggtct ttggtgaggc gatgtgtaag 550

ttgaatcctt ttgtgcaatg tgtttcaatc actgtgtcca ttttctctct 600

ggttctcatt gctgtggaac gacatcagct gataatcaac cctcgagggt 650

ggagaccaaa taatagacat gcttatgtag gtattgctgt gatttgggtc 700

cttgctgtgg cttcttcttt gcctttcctg atctaccaag taatgactga 750

tgagccgttc caaaatgtaa cacttgatgc gtacaaagac aaatacgtgt 800

gctttgatca atttccatcg gactctcata ggttgtctta taccactctc 850

ctcttggtgc tgcagtattt tggtccactt tgttttatat ttatttgcta 900

cttcaagata tatatacgcc taaaaaggag aaacaacatg atggacaaga 950

tgagagacaa taagtacagg tccagtgaaa ccaaaagaat caatatcatg 1000

ctgctctcca ttgtggtagc atttgcagtc tgctggctcc ctcttaccat 1050

ctttaacact gtgtttgatt ggaatcatca gatcattgct acctgcaacc 1100

acaatctgtt attcctgctc tgccacctca cagcaatgat atccacttgt 1150

gtcaacccca tattttatgg gttcctgaac aaaaacttcc agagagactt 1200

gcagttcttc ttcaactttt gtgatttccg gtctcgggat gatgattatg 1250

aaacaatagc catgtccacg atgcacacag atgtttccaa aacttctttg 1300

aagcaagcaa gcccagtcgc atttaaaaaa atcaacaaca atgatgataa 1350

tgaaaaaatc tgaaactact tatagcctat ggtcccggat gacatctgtt 1400

taaaaacaag cacaacctgc aacatacttt gattacctgt tctcccaagg 1450

aatggggttg aaatcatttg aaaatgacta agattttctt gtcttgcttt 1500

ttactgcttt tgttgtagtt gtcataatta catttggaac aaaaggtgtg 1550

ggctttgggg tcttctggaa atagttttga ccagacatct ttgaagtgct 1600

ttttgtgaat ttatgcatat aatataaaga cttttatact gtacttattg 1650

gaatgaaatt tctttaaagt attacgatgc gctgacttca gaagtacctg 1700

ccatccaata cggtcattag attgggtcat cttgattaga ttagattaga 1750

ttagattgtc aacagattgg gccatcctta ctttatgata ggcatcattt 1800

tagtgtgtta caatagtaac agtatgcaaa agcagcattc aggagccgaa 1850

agatagtctt gaagtcattc agaagtggtt tgaggtttct gtttttggt 1900

ggtttttgtt tgttttttt ttttttcacc ttaagggagg ctttcatttc 1950

ctcccgactg attgtcactt aaatcaaaat ttaaaaatga ataaaaagac 2000

atacttctca gctgcaaata ttatggagaa ttgggcaccc acaggaatga 2050

agagagaaag cagctcccca acttcaaaac cattttggta cctgacaaca 2100

agagcatttt agagtaatta atttaataaa gtaaattagt attgctgcaa 2150

atagctaaat tatatttatt tgaattgatg gtcaagagat tttccatttt 2200

ttttacagac tgttcagtgt ttgtcaagct tctggtctaa tatgtactcg 2250

aaagactttc cgcttacaat ttgtagaaac acaaatatcg ttttccatac 2300

agcagtgcct atatagtgac tgattttaac tttcaatgtc catctttcaa 2350

aggaagtaac accaaggtac aatgttaaag gaatattcac tttacctagc 2400

agggaaaaat acacaaaaac tgcagatact tcatatagcc cattttaact 2450

tgtataaact gtgtgacttg tggcgtctta taaataatgc actgtaaaga 2500

ttactgaata gttgtgtcat gttaatgtgc ctaatttcat gtatcttgta 2550

atcatgattg agcctcagaa tcatttggag aaactatatt ttaaagaaca 2600

agacatactt caatgtatta tacagataaa gtattacatg tgtttgattt 2650

taaaagggcg gacattttat taaaatcaat attgtttttg ctttttctga 2700

ggagtctctt tcagtttcat tttttctcat cccatgactt ccctccgatg 2750

gt 2752

<210> 74
<211> 1001
<212> DNA
<213> Homo sapiens

<400> 74
agtatgaggc gcccaccatg gtgcgatgcg accgcgggct gcagatgctg 50

ctgaccacgg ccggagcctt cgccgccttc tcgctcatgg ccatcgccat 100

cggcaccgac tactggctgt actccagcgc gcacatctgc aacggcacca 150

acctgaccat ggacgacggg cccccgcccc gccgcgcccg cggcgacctc 200

acccactctg gtctgtggcg ggtgtgctgc atcgaaggga tctataaagg 250

gcactgcttc cggatcaatc acttcccaga ggacaatgac tacgaccacg 300

acagctcgga gtacctcctc cgcatcgtgc gagcctccag cgtcttcccc 350

atcctcagca ccatcctgct cctgctgggt ggcctgtgca tcggtgctgg 400

caggatctac agccgcaaga acaacatcgt cctcagtgcc ggcatcctct 450

tcgtggctgc aggcctcagt aacatcatcg gtatcatcgt ctacatttcc 500

agcaacacag gtgacccgag tgacaagcgg gacgaagaca aaaagaacca 550

ttacaactac ggctggtctt tttactttgg agctctgtct ttcattgtgg 600

ctgagaccgt gggcgtcctg gctgtaaaca tttacattga gaaaaataaa 650

gagttgaggt ttaagaccaa acgggaattc cttaaggcgt cttcctcttc 700

tccttatgcc aggatgccga gctacaggta ccggcgacgg cgctcgaggt 750

ccagctcaag gtccaccgag gcctcgccct ccagggacgt gtcgcccatg 800

ggcctgaaga tcacaggggc catccccatg ggggagctgt ccatgtacac 850

gctgtccagg gagcccctca aggtgaccac cgcagccagc tacagccccg 900

accaggaggc cagcttcctg caggtgcatg acttttttcca gcaggacctg 950

aaggaaggtt tccacgtcag catgctgaac cgacggacga cccctgtgtg 1000

a 1001

<210> 75
<211> 2805
<212> DNA
<213> Homo sapiens

<400> 75
cgggtctgat agtccctacc tgtcaggact ggtgttagga tgagataatg 50

tttgtgaact gtaaacatat ataaacgtgt gctactgtga gaactggaac 100

aaagaagaga gggagtgaga gaaatcaagg gagggctggg gctgggaaag 150

aacgaaaagg gagtcgcgta tagaggagag gcgacagtcg cgagccacac 200

tttgcaatga aactctttag actttctgcc gggagagcgg cccagacgcg 250

ccaggtctgt agcaggaggc cgcgagggcg ggtccccaga agcctacagg 300

tgagtatcgg ttctcccctt cccggctttc ggtccggagg aggcgggagc 350

agcttccctg ttctgatcct atcgcgggcg gcgcagggcc ggcttggcct 400

tccgtgggac ggggaggggg gcgggatgtg tcacccaaat accagtgggg 450

acggtcggtg gtggaaccag ccgggcaggt cgggtagagt ataagagccg 500

gagggagcgg ccgggggcgca gacgcctgca gaccatccca gacgccggag 550

cccgagcccc gacgagtccc cgcgcctcat ccgcccgcgt ccggtccgcg 600

ttcctccgcc ccaccatggc tcggggcccc ggcctcgcgc cgccaccgct 650

gcggctgccg ctgctgctgc tggtgctggc ggcggtgacc ggccacacgg 700

ccgcgcagga caactgcacg tgtcccacca acaagatgac cgtgtgcagc 750

cccgacggcc ccggcggccg ctgccagtgc cgcgcgctgg gctcgggcat 800

ggcggtcgac tgctccacgc tgacctccaa gtgtctgctg ctcaaggcgc 850

gcatgagcgc ccccaagaac gcccgcacgc tggtgcggcc gagtgagcac 900

gcgctcgtgg acaacgatgg cctctacgac cccgactgcg accccgaggg 950

ccgcttcaag gcgcgccagt gcaaccagac gtcggtgtgc tggtgcgtga 1000

actcggtggg cgtgcgccgc acggacaagg gcgacctgag cctacgctgc 1050

gatgagctgg tgcgcaccca ccacatcctc attgacctgc gccaccgccc 1100

caccgccggc gccttcaacc actcagacct ggacgccgag ctgaggcggc 1150

tcttccgcga gcgctatcgg ctgcacccca agttcgtggc ggccgtgcac 1200

tacgagcagc ccaccatcca gatcgagctg cggcagaaca cgtctcagaa 1250

ggccgccggt gaagtggata tcggcgatgc cgcctactac ttcgagaggg 1300

acatcaaggg cgagtctcta ttccagggcc gcggcggcct ggacttgcgc 1350

gtgcgcggag aacccctgca ggtggagcgc acgctcatct attacctgga 1400

cgagattccc ccgaagttct ccatgaagcg cctcaccgcc ggcctcatcg 1450

ccgtcatcgt ggtggtcgtg gtggccctcg tcgccggcat ggccgtcctg 1500

gtgatcacca accggagaaa gtcggggaag tacaagaagg tggagatcaa 1550

ggaactgggg gagttgagaa aggaaccgag cttgtaggta cccggcgggg 1600

caggggatgg ggtggggtac cggatttcgg tatcgtccca gacccaagtg 1650

agtcacgctt cctgattcct cggcgcaaag gagacgttta tcctttcaaa 1700

ttcctgcctt ccccctccct tttgcgcaca caccaggttt aatagatcct 1750

ggcctcaggg tctcctttct ttctcacttc tgtcttgagg gaagcatttc 1800

taaaatgtat ccccttttcgg tccaacaaca ggaaacctga ctggggcagt 1850

gaaggaaggg atggcacagc gttatgtgta aaaaacaagt atctgtatga 1900

caacccggga tcgtttgcaa gtaactgaat ccattgcgac attgtgaagg 1950

cttaaatgag tttagatggg aaatagcgtt gttatcgcct tgggtttaaa 2000

ttatttgatg agttccactt gtatcatggc ctacccgagg agaagaggag 2050

tttgttaact gggcctatgt agtagcctca tttaccatcg tttgtattac 2100

tgaccacata tgcttgtcac tgggaaagaa gcctgtttca gctgcctgaa 2150

cgcagtttgg atgtctttga ggacagacat tgcccggaaa ctcagtctat 2200

ttattcttca gcttgccctt actgccactg atattggtaa tgttcttttt 2250

tgtaaaatgt ttgtacatat gttgtctttg ataatgttgc tgtaattttt 2300

taaaataaaa cacgaattta ataaaatatg ggaaaggcac aaaccagaag 2350

tcggcatttg tgaaaagtcc ctccagattt ctatcacttt ggtctctaat 2400

ttcccaagac ttgtattttt tttttatttc aaattataac actttttttt 2450

cccccagaag tgggtgtttc atgttgctac tctggtgtgt cccaagatat 2500

cctaactggc cagtgtaaat gctattcttt ctaaataaga ttatttggaa 2550

acttccttca aactgcagga gggcgagctc tgagggcacg agaagctaaa 2600

actagctgct tttgatgaaa aagagtgcca gtctttggtc atctctaaac 2650

aaggcttatc accaatggag acagaaaact ctagttcaag agctgtacct 2700

cctttgaatc ccagccctac tcgaaataag tggtactatt tccatttagc 2750

ctttgagcaa atcacttaac tcaaaggcgt tgtggctcta agattaaacg 2800

acttt 2805

<210> 76
<211> 2637
<212> DNA
<213> Homo sapiens

<400> 76
gcggccgcgg agcaagaagg gcgccgcgtc gtgcggcccg cgcagccccc 50

ggagccatgg gcaagtgcag cgggcgctgc acgctggtcg ccttctgctg 100

cctgcagctg gtggctgcgc tggagcggca gatctttgac ttcctgggct 150

accagtgggc tcccatccta gccaacttcc tgcacatcat ggcagtcatc 200

ctgggcatct ttggcaccgt gcagtaccgc tcccggtacc tcatcctgta 250

tgcagcctgg ctggtgctct gggttggctg gaatgcattt atcatctgct 300

tctacttgga ggttggacag ctgtcccagg accgggactt catcatgacc 350

ttcaacacat ccctgcaccg ctcctggtgg atggagaatg ggccaggctg 400

cctggtgaca cctgttctga actcccgcct ggctctggag gaccaccatg 450

tcatctctgt cactggctgc ctgcttgact acccctacat tgaagccctc 500

agcagcgccc tgcagatctt cctggcactg ttcggcttcg tgttcgcctg 550

ctacgtgagc aaagtgttcc tggaggagga ggacagcttt gacttcatcg 600

gcggctttga ctcctacgga taccaggcgc cccagaagac gtcgcattta 650

cagctgcagc ctctgtacac gtcggggtag cctctgcccc gcgcccaccc 700

cggcgcctcg ccctgggctg accgcagctg ccgcgagctc gggccaaggc 750

gcaggcgtgt ccccctggtg gcccgcgctc tcactgcagc ctgtgcccaa 800

ccccgcgtct gcatctggag atgcggactt ggacgtggac ttggacttgg 850

acttggattt gagcttggct cttcgcagcc cggacttcgg aggagtgggg 900

cggggcgggg gaggggcacc acgggttttt tgtttttttgt ttgtttgttt 950

ttaatctcag ccttggcgtg agctggggcc ttcctctctt ctccagcctc 1000

tccctttcac tcttcgccca gcatcctgcc cccctgtcca aaaacagcag 1050

gacatcaggc ccatcccatc ccaccacact cactcaccag ctctggggaa 1100

agctactgtg aactaggagc aggattcctg ggttctaatc gcaggtccat 1150

cactgactgt gacgtctagc aaagcccttg ccctctctga gcctcggttt 1200

ccgcacctca agtaattaat cccttagcaa atggactctt ttagacttct 1250

catttaactc aattccctga gctagactgg gattaaaatt ctcattttgc 1300

agtacattaa aactgatgcc cagagatgtg atttgcttga ggccacacag 1350

ctagattttt ggtggaagtg ggccttgaac acagtgtact ttctgcagtt 1400

tctgactgta aaacccagtg tctgctctct gagttccatt tccaagcccc 1450

cctccatctt ggacctatgt ggtctccacc atattcacac accaccacca 1500

ccacttgcca atgcctctct taaagcaata tacccattcg ttctcttatt 1550

gggaactgga tggatgaagc cccaaattca gccccaccca cagagaagcc 1600

ttcctacact cagcctctgt ccacccttgg caaatctttc aagctctctc 1650

ctccaggaaa gtggggcccc aactcagtca ctccaccccc ttccaggtcc 1700

ctgaggctgg ttctactgta tccccatcac ctccacaact ccactcaccc 1750

ctgacggctc catccacctc accagttgga aggcttgtgg tttcagagag 1800

gagcaatgct ggtcagcgct gcccagactc cagtgtttac agatcaccag 1850

catttacaac caatccaatg gccagaagcc tcctctaaca agcccagaag 1900

gagttctgaa ggggcagatg ggggtgtgag tagtcgggga gtcgggattg 1950

ccagcaccct caccettcct tgggggcaag tagaggtgag aacactttcc 2000

ccacctccct ccacagacac tcctgaggac gctgcatccc acgcactgcc 2050

tggtgcgtcc atagagagag gatcaggtct cagcatttca tctgtgaaag 2100

aggcatggcc ctgggttaga aaggagggca ggagacatgg aggaactggg 2150

gggcacccag atggtgcaga tggtttgcac acctgagcct gtctgtggtg 2200

accattccgc tcctctccca ctaccctcca atctatcatt ccctactctc 2250

taaggccaaa atatcctgag caaggctggc aaccccaccc caccatccca 2300

aatgcaagca gccaggccca ggagttcctc tggcccccac aggcatggag 2350

ctcccagctg gtgggtacag cttgagaggg gggcagctcc ctcaggctaa 2400

gctactgccc ttcactgggc cagccctgcc tccagccctc acctctctca 2450

ccccaactct cccccaagcc cctttctact caacgggtgt agccactggt 2500

gctttgaagc cttttgtttt tataagatgg tttttgcaag gggaccaggt 2550

tctcttttca ctgggacctt gcaaggaggg gagtgctctc ctggtttctg 2600

tgcaggcggg ttgattaaag atggtgtttt cttctct 2637

<210> 77
<211> 1829
<212> DNA
<213> Homo sapiens

<400> 77
ggcggggaga ggagaggaga gaagagccgc ggggggccca gcccggagcc 50

aggatgcccg cgccgcgcgc ccgggagcag ccccgcgtgc ccggggagcg 100

ccagccgctg ctgcctcgcg gtgcgcgggg ccctcgacgg tggcggcggg 150

cggcgggcgc ggccgtgctg ctggtggaga tgctggagcg cgccgccttc 200

ttcggcgtca ccgccaacct cgtgctgtac ctcaacagca ccaacttcaa 250

ctggaccggc gagcaggcga cgcgcgccgc gctggtattc ctgggcgcct 300

cctacctgct ggcgcccgtg ggcggctggc tggccgacgt gtacctgggc 350

cgctaccgcg cggtcgcgct cagcctgctg ctctacctgg ccgcctcggg 400

cctgctgccc gccaccgcct tccccgacgg ccgcagctcc ttctgcggag 450

agatgcccgc gtcgccgctg ggacctgcct gcccctcggc cggctgcccg 500

cgctcctcgc ccagccccta ctgcgcgccc gtcctctacg cgggcctgct 550

gctactcggc ctggccgcca gctccgtccg gagcaacctc acctccttcg 600

gtgccgacca ggtgatggat ctcggccgcg acgccacccg ccgcttcttc 650

aactggtttt actggagcat caacctgggt gctgtgctgt cgctgctggt 700

ggtggcgttt attcagcaga acatcagctt cctgctgggc tacagcatcc 750

ctgtgggctg tgtgggcctg gcatttttca tcttcctctt tgccaccccc 800

gtcttcatca ccaagccccc gatgggcagc caagtgtcct ctatgcttaa 850

gctcgctctc caaaactgct gcccccagct gtggcaacga cactcggcca 900

gagaccgtca atgtgcccgc gtgctggccg acgagaggtc tccccagcca 950

ggggcttccc cgcaagagga catcgccaac ttccaggtgc tggtgaagat 1000

cttgcccgtc atggtgaccc tggtgcccta ctggatggtc tacttccaga 1050

tgcagtccac ctatgtcctg cagggtcttc acctccacat cccaaacatt 1100

ttcccagcca acccggccaa catctctgtg gccctgagag cccagggcag 1150

cagctacacg atcccggaag cctggctcct cctggccaat gttgtggtgg 1200

tgctgattct ggtccctctg aaggaccgct tgatcgaccc tttactgctg 1250

cggtgcaagc tgcttccctc tgctctgcag aagatggcgc tggggatgtt 1300

ctttggtttt acctccgtca ttgtggcagg agtcctggag atggagcgct 1350

tacactacat ccaccacaac gagaccgtgt cccagcagat tggggaggtc 1400

ctgtacaacg cggcaccact gtccatctgg tggcagatcc ctcagtacct 1450

gctcattggg atcagtgaga tctttgccag catcccaggc ctggagtttg 1500

cctactcaga ggccccgcgc tccatgcagg gcgccatcat gggcatcttc 1550

ttctgcctgt cgggggtggg ctcactgttg ggctccagcc tagtggcact 1600

gctgtccttg cccgggggct ggctgcactg ccccaaggac tttgggaaca 1650

tcaacaattg ccggatggac ctctacttct tcctgctggc tggcattcag 1700

gccgtcacgg ctctcctatt tgtctggatc gctggacgct atgagagggc 1750

gtcccagggc ccagcctccc acagccgttt cagcagggac aggggctgaa 1800

caggccctat tccagccccc ttgcttcac 1829

<210> 78
<211> 5746
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 5621,, 5664
<223> unknown base

<400> 78
ggccgcgggc tctcgcgggg cggcgacgcc gcggggagga tgctgcttgc 50

cgcgcccgcg tcctcaccgt cctcccgggc cgcctgctgg ggctttgttg 100

tggcccggac gccgcgggcc acccccctgaa gtcgcctgcc gccgccgccg 150

ccgcacctag cggacgggcg ggcgggcgcg cgtgtgccca ggagtgcgcg 200

cctgtcgcgg tggtgggtgc aggactggac ccacgggccc attgtgcgcc 250

cgcccgcggc agccaggacc atgtgggtga acccggagga ggtgttgctg 300

gccaacgcgc tgtggatcac cgagagggcc aacccatact tcatcctgca 350

gcggaggaag ggccacgccg gcgatggagg cggcggcggc ggactggcgg 400

gcctgctggt gggtaccctt gatgttgtgt tggactccag cgcccgggtc 450

gctccttacc gaatcttgta ccagactcca gactccctgg tctactggac 500

catcgcctgt ggtggttcca ggaaagaaat cactgaacac tgggaatggc 550

ttgagcaaaa tctcttgcag acactctcca tctttgaaaa tgagaatgat 600

atcaccacat ttgtgagagg aaaaatacag ggcatcattg cagaatacaa 650

caaaatcaat gatgtaaagg aagatgatga cacggagaag tttaaagaag 700

ccattgtgaa atttcatagg ctgtttggga tgccagagga agagaaactc 750

gtcaactatt actcttgcag ctattggaag gggaaggtcc cccgtcaggg 800

ttggatgtac ctcagcatta accacctttg cttttattct tttcttatgg 850

gaagggaagc gaaactggtc atccggtggg tagacatcac tcagcttgag 900

aagaatgcca ccctgcttct gcctgatgtg atcaaagtga gcacacggtc 950

cagtgagcat ttcttctctg tattcctcaa catcaacgag accttcaagt 1000

taatggagca gcttgccaac atagccatga ggcaactctt agacaatgag 1050

ggatttgaac aagatcgatc cctgcccaaa ctcaaaagga aatctcctaa 1100

aaaagtgtct gctctaaaac gtgatcttga tgccagggca aagagtgaga 1150

gataccgtgc acttttccgg ctgcccaaag atgaaaaatt agatggccac 1200

acagactgca ctctctggac tccatttaac aaaatgcaca ttttggggca 1250

gatgtttgtg tccacaaatt acatctgttt taccagcaag gaggagaact 1300

tatgtagcct cattatcccg ctccgtgagg tgacaattgt ggaaaaggca 1350

gacagctcca gtgtgctccc cagtccctta tccatcagca cccgaaacag 1400

gatgaccttc ctatttgcca acttgaaaga tagagacttt ctagtgcaga 1450

ggatctcaga tttcctgcaa cagactactt ccaaaatata ttctgacaag 1500

gagtttgcag gaagttacaa cagttcagat gatgaggtgt actctcgacc 1550

cagcagcctc gtctcctcca gcccccagag aagcacgagc tctgatgctg 1600

atggagagcg ccagtttaac ctaaatggca acagcgtccc cacagccaca 1650

cagaccctga tgaccatgta tcggcggcgg tctcccgagg agttcaaccc 1700

gaaattggcc aaagagtttc tgaaagagca agcctggaag attcactttg 1750

ctgagtatgg gcaagggatc tgcatgtacc gcacagagaa aacgcgggag 1800

ctggtgttga agggcatccc ggagagcatg cgtggggagc tctggctgct 1850

gctgtcaggt gccatcaatg agaaggccac acatcctggg tactatgaag 1900

acctagtgga gaagtccatg gggaagtata atctcgccac ggaggagatt 1950

gagagggatt tacaccgctc ccttccagaa cacccagctt ttcagaatga 2000

aatgggcatt gctgcactaa ggagagtctt aacagcttat gcttttcgaa 2050

atcccaacat agggtattgc caggccatga atattgtcac ttcagtgctg 2100

ctgctttatg ccaaagagga ggaagctttc tggctgcttg tggctttgtg 2150

tgagcgcatg ctcccagatt actacaacac cagagttgtg ggtgcactgg 2200

tggaccaagg tgtctttgag gagctagcac gagactacgt cccacagctg 2250

tacgactgca tgcaagacct gggcgtgatt tccaccatct ccctgtcttg 2300

gttcctcaca ctatttctca gtgtgatgcc ttttgagagt gcagttgtgg 2350

ttgttgactg tttcttctat gaaggaatta aagtgatatt ccagttggcc 2400

ctagctgtgc tggatgcaaa tgtggacaaa ctgttgaact gcaaggatga 2450

tggggaggcc atgaccgttt tgggaaggta tttagacagt gtgaccaata 2500

aagacagcac actgcctccc attcctcacc tccactcctt gctcagcgat 2550

gatgtggaac cttaccctga ggtagacatc tttagactca tcagaacttc 2600

ctacgagaaa ttcggaacta tccgggcaga tttgattgaa cagatgagat 2650

tcaaacagag actgaaagtg atccagacgc tggaggatac tacgaaacgc 2700

aacgtggtac gaaccattgt gacagaaact tcctttacca ttgatgagct 2750

ggaagaactt tatgctcttt tcaaggcaga acatctcacc agctgctact 2800

ggggcgggag cagcaacgcg ctggaccggc atgaccccag cctgccctac 2850

ctggaacagt atcgcattga cttcgagcag ttcaagggaa tgtttgctct 2900

tctctttcct tgggcatgtg gaactcactc tgacgttctg gcctcccgct 2950

tgttccagtt attagatgaa aatggagact ctttgattaa cttccgggag 3000

tttgtctctg ggctaagtgc tgcatgccat ggggacctca cagagaagct 3050

caaactcctg tacaaaatgc acgtcttgcc tgagccatcc tctgatcaag 3100

atgaaccaga ttctgctttt gaagcaactc agtacttctt tgaagatatt 3150

accccagaat gtacacatgt tgttggattg gatagcagaa gcaaacaggg 3200

tgcagatgat ggctttgtta cggtgagcct aaagccagac aaagggaaga 3250

gagcaaattc ccaagaaaat cgtaattatt tgagactgtg gactccagaa 3300

aataaatcta agtcaaagaa tgcaaaggat ttacccaaat taaatcaggg 3350

gcagttcatt gaactgtgta agacaatgta taacatgttc agcgaagacc 3400

ccaatgagca ggagctgtac catgccacgg cagcagtgac cagcctcctg 3450

ctggagattg gggaggtcgg caagttgttc gtggcccagc ctgcaaagga 3500

gggcgggagc ggaggcagtg ggccgtcctg ccaccagggc atcccaggcg 3550

tgctcttccc caagaaaggg ccaggccagc cttacgtggt ggagtctgtt 3600

gagcccctgc cggccagcct ggcccccgac agcgaggaac actcccttgg 3650

aggacaaatg gaggacatca agctggagga ctcctcgccc cgggacaacg 3700

gggcctgctc ctccatgctg atctctgacg acgacaccaa ggacgacagc 3750

tccatgtcct catactcggt gctgagtgcc ggctcccacg aggaggacaa 3800

gctgcactgc gaggacatcg gagaggacac ggtcctggtg cggagcggcc 3850

agggcacggc ggcactgccc cggagcacca gcctggaccg ggactgggcc 3900

atcaccttcg agcagttcct ggcctccctc ttaactgagc ctgccctggt 3950

caagtacttt gacaagcccg tgtgcatgat ggccaggatt accagtgcaa 4000

aaaacatccg gatgatgggc aagcccctca cctcggccag tgactatgaa 4050

atctcggcca tgtccggctg acacgggcgc cttcccgggg gagtgggagg 4100

agagggaggg gagggatttt ttatgttctt ctgtgttgag ttttttcttt 4150

ctttctttta aattaaatat ttattagtac ctggcttgaa gcctagtgtt 4200

ttcataatgt aattcaatga aaactgttgg agaaatattt aaacacctca 4250

atgtaggtac attacactct tgttgcgggg aggggattta ccagaataca 4300

gtttatttcg tgaattctaa aaaacaaaaa gatgaatctg tcagtgatat 4350

gtgtgtatta taacttatta atcttgctgt tgagctgtat acatggttta 4400

aaaaatagta ctgtttaatg ctaagtaagg cagcagtcat ttgtgtattc 4450

aggcttttta aataaaatta gagctgtaag gaaaatgaaa agccacaaat 4500

gcaagactgt tcttaaatgg aaggcatagt cagcgagggt aaatcctata 4550

ccactttagg aagtattaaa aatattttta agatttgaaa tatatttcat 4600

agaagtcctc tattcaaaat catattccac agatgttccc cttcaaaggg 4650

aaaacatttg gggttctaaa cagttatgaa agtaagtgat ttttacatga 4700

ttccagaata acacttgtat tgaccaattt agacagatac cagaccaatt 4750

ttgcatttaa gaaattgttc tgattattta cgtcaactca ttagaattca 4800

gtgaaaagta acagtctttt gtcacagaga atctgaaagt agcagcaaag 4850

acagagggct catgacaggt ttttgctttt gctttgcttt tgtttttgaa 4900

agagtaaaag tactgatgct tctgatactg gatgtttagc ttcttactgc 4950

aaaaacataa gtaaaacagt caactttacc atttccgtat tctccataga 5000

ttgaagaaat ttataccaca tatcgcatat gaccatcttt ccatcaaatc 5050

aatgtagaga taatgtaaac tgaaaaaaaa tctgcaagat aatgtaactg 5100

aatgttttaa aaacagaact tgtcacttta tataaaagaa tagtatgctc 5150

tatttcctga atggatgtgg aaatgaaagc tagcgcacct gcactttgaa 5200

ttcttgcttc ttttttatta ctgttatgat tttgcttttt acagatgttg 5250

gacgattttt tcttctgatt gttgaattca taatcatggt ctcatttcct 5300

ttgcttcttt ggaatatttc tttcaacaca ttcctttatt ttattataca 5350

ttgtgtcctt tttttagcta ttgctgctgt tgtttttttat tctatttaca 5400

ggatgatttt taaactgtca aatgaagtag tgttaacctc aaataggcta 5450

aatgtgaaca aataaaatac agcaaatact cagatacagc tttttatctt 5500

tgtgcttgag ttcctgccta aggaataaca ttattctttt gacaactttt 5550

gcaggggaaa ttatatcagg caaccatttt gattaagtaa ataaatttta 5600

taggcaaaca tatagagaga natacaattt gtagtatatc aatgactata 5650

tttaaaataa ggantataat tgttatcagt tatctaactt aaaatgctta 5700

tccataatga tcagtgatat tcagcttttt aaaatatgct tgttgg 5746

<210> 79
<211> 183
<212> PRT
<213> Homo sapiens

<400> 79

Met	Arg	Ala	Ala	Tyr	Leu	Phe	Leu	Leu	Phe	Leu	Pro	Ala	Gly	Leu
1				5					10					15
Leu	Ala	Gln	Gly	Gln	Tyr	Asp	Leu	Asp	Pro	Leu	Pro	Pro	Phe	Pro
				20					25					30
Asp	His	Val	Gln	Tyr	Thr	His	Tyr	Ser	Asp	Gln	Ile	Asp	Asn	Pro
				35					40					45
Asp	Tyr	Tyr	Asp	Tyr	Gln	Glu	Val	Thr	Pro	Arg	Pro	Ser	Glu	Glu
				50					55					60
Gln	Phe	Gln	Phe	Gln	Ser	Gln	Gln	Gln	Val	Gln	Gln	Glu	Val	Ile
				65					70					75
Pro	Ala	Pro	Thr	Pro	Glu	Pro	Gly	Asn	Ala	Glu	Leu	Glu	Pro	Thr
				80					85					90
Glu	Pro	Gly	Pro	Leu	Asp	Cys	Arg	Glu	Glu	Gln	Tyr	Pro	Cys	Thr
				95					100					105

Arg Leu Tyr Ser Ile His Arg Pro Cys Lys Gln Cys Leu Asn Glu
110 115 120

Val Cys Phe Tyr Ser Leu Arg Arg Val Tyr Val Ile Asn Lys Glu
125 130 135

Ile Cys Val Arg Thr Val Cys Ala His Glu Glu Leu Leu Arg Ala
140 145 150

Asp Leu Cys Arg Asp Lys Phe Ser Lys Cys Gly Val Met Ala Ser
155 160 165

Ser Gly Leu Cys Gln Ser Val Ala Ala Ser Cys Ala Arg Ser Cys
170 175 180

Gly Ser Cys

<210> 80
<211> 909
<212> PRT
<213> Homo sapiens

<400> 80
Met Ala Ala Arg Pro Leu Pro Val Ser Pro Ala Arg Ala Leu Leu
1 5 10 15

Leu Ala Leu Ala Gly Ala Leu Leu Ala Pro Cys Glu Ala Arg Gly
20 25 30

Val Ser Leu Trp Asn Glu Gly Arg Ala Asp Glu Val Val Ser Ala
35 40 45

Ser Val Arg Ser Gly Asp Leu Trp Ile Pro Val Lys Ser Phe Asp
50 55 60

Ser Lys Asn His Pro Glu Val Leu Asn Ile Arg Leu Gln Arg Glu
65 70 75

Ser Lys Glu Leu Ile Ile Asn Leu Glu Arg Asn Glu Gly Leu Ile
80 85 90

Ala Ser Ser Phe Thr Glu Thr His Tyr Leu Gln Asp Gly Thr Asp
95 100 105

Val Ser Leu Ala Arg Asn Tyr Thr Val Ile Leu Gly His Cys Tyr
110 115 120

Tyr His Gly His Val Arg Gly Tyr Ser Asp Ser Ala Val Ser Leu
125 130 135

Ser Thr Cys Ser Gly Leu Arg Gly Leu Ile Val Phe Glu Asn Glu
140 145 150

Ser Tyr Val Leu Glu Pro Met Lys Ser Ala Thr Asn Arg Tyr Lys
155 160 165

Leu Phe Pro Ala Lys Lys Leu Lys Ser Val Arg Gly Ser Cys Gly
170 175 180

Ser His His Asn Thr Pro Asn Leu Ala Ala Lys Asn Val Phe Pro
185 190 195

Pro Pro Ser Gln Thr Trp Ala Arg Arg His Lys Arg Glu Thr Leu
200 205 210

Lys Ala Thr Lys Tyr Val Glu Leu Val Ile Val Ala Asp Asn Arg
215 220 225

Glu Phe Gln Arg Gln Gly Lys Asp Leu Glu Lys Val Lys Gln Arg
230 235 240

Leu Ile Glu Ile Ala Asn His Val Asp Lys Phe Tyr Arg Pro Leu
245 250 255

Asn Ile Arg Ile Val Leu Val Gly Val Glu Val Trp Asn Asp Met

260 265 270

Asp Lys Cys Ser Val Ser Gln Asp Pro Phe Thr Ser Leu His Glu
275 280 285

Phe Leu Asp Trp Arg Lys Met Lys Leu Leu Pro Arg Lys Ser His
290 295 300

Asp Asn Ala Gln Leu Val Ser Gly Val Tyr Phe Gln Gly Thr Thr
305 310 315

Ile Gly Met Ala Pro Ile Met Ser Met Cys Thr Ala Asp Gln Ser
320 325 330

Gly Gly Ile Val Met Asp His Ser Asp Asn Pro Leu Gly Ala Ala
335 340 345

Val Thr Leu Ala His Glu Leu Gly His Asn Phe Gly Met Asn His
350 355 360

Asp Thr Leu Asp Arg Gly Cys Ser Cys Gln Met Ala Val Glu Lys
365 370 375

Gly Gly Cys Ile Met Asn Ala Ser Thr Gly Tyr Pro Phe Pro Met
380 385 390

Val Phe Ser Ser Cys Ser Arg Lys Asp Leu Glu Thr Ser Leu Glu
395 400 405

Lys Gly Met Gly Val Cys Leu Phe Asn Leu Pro Glu Val Arg Glu
410 415 420

Ser Phe Gly Gly Gln Lys Cys Gly Asn Arg Phe Val Glu Glu Gly
425 430 435

Glu Glu Cys Asp Cys Gly Glu Pro Glu Glu Cys Met Asn Arg Cys
440 445 450

Cys Asn Ala Thr Thr Cys Thr Leu Lys Pro Asp Ala Val Cys Ala
455 460 465

His Gly Leu Cys Cys Glu Asp Cys Gln Leu Lys Pro Ala Gly Thr
470 475 480

Ala Cys Arg Asp Ser Ser Asn Ser Cys Asp Leu Pro Glu Phe Cys
485 490 495

Thr Gly Ala Ser Pro His Cys Pro Ala Asn Val Tyr Leu His Asp
500 505 510

Gly His Ser Cys Gln Asp Val Asp Gly Tyr Cys Tyr Asn Gly Ile
515 520 525

Cys Gln Thr His Glu Gln Gln Cys Val Thr Leu Trp Gly Pro Gly
530 535 540

Ala Lys Pro Ala Pro Gly Ile Cys Phe Glu Arg Val Asn Ser Ala
545 550 555

Gly Asp Pro Tyr Gly Asn Cys Gly Lys Val Ser Lys Ser Ser Phe
560 565 570

Ala Lys Cys Glu Met Arg Asp Ala Lys Cys Gly Lys Ile Gln Cys
575 580 585

Gln Gly Gly Ala Ser Arg Pro Val Ile Gly Thr Asn Ala Val Ser
590 595 600

Ile Glu Thr Asn Ile Pro Leu Gln Gln Gly Gly Arg Ile Leu Cys
605 610 615

Arg Gly Thr His Val Tyr Leu Gly Asp Asp Met Pro Asp Pro Gly
620 625 630

Leu Val Leu Ala Gly Thr Lys Cys Ala Asp Gly Lys Ile Cys Leu
635 640 645

Asn Arg Gln Cys Gln Asn Ile Ser Val Phe Gly Val His Glu Cys
650 655 660

Ala Met Gln Cys His Gly Arg Gly Val Cys Asn Asn Arg Lys Asn
665 670 675

Cys His Cys Glu Ala His Trp Ala Pro Pro Phe Cys Asp Lys Phe
680 685 690

Gly Phe Gly Gly Ser Thr Asp Ser Gly Pro Ile Arg Gln Ala Asp
695 700 705

Asn Gln Gly Leu Thr Ile Gly Ile Leu Val Thr Ile Leu Cys Leu
710 715 720

Leu Ala Ala Gly Phe Val Val Tyr Leu Lys Arg Lys Thr Leu Ile
725 730 735

Arg Leu Leu Phe Thr Asn Lys Lys Thr Thr Ile Glu Lys Leu Arg
740 745 750

Cys Val Arg Pro Ser Arg Pro Pro Arg Gly Phe Gln Pro Cys Gln
755 760 765

Ala His Leu Gly His Leu Gly Lys Gly Leu Met Arg Lys Pro Pro
770 775 780

Asp Ser Tyr Pro Pro Lys Asp Asn Pro Arg Arg Leu Leu Gln Cys
785 790 795

Gln Asn Val Asp Ile Ser Arg Pro Leu Asn Gly Leu Asn Val Pro
800 805 810

Gln Pro Gln Ser Thr Gln Arg Val Leu Pro Pro Leu His Arg Ala
815 820 825

Pro Arg Ala Pro Ser Val Pro Ala Arg Pro Leu Pro Ala Lys Pro
830 835 840

Ala Leu Arg Gln Ala Gln Gly Thr Cys Lys Pro Asn Pro Pro Gln
845 850 855

Lys Pro Leu Pro Ala Asp Pro Leu Ala Arg Thr Thr Arg Leu Thr
860 865 870

His Ala Leu Ala Arg Thr Pro Gly Gln Trp Glu Thr Gly Leu Arg
875 880 885

Leu Ala Pro Leu Arg Pro Ala Pro Gln Tyr Pro His Gln Val Pro
890 895 900

Arg Ser Thr His Thr Ala Tyr Ile Lys
905

<210> 81
<211> 158
<212> PRT
<213> Homo sapiens

<400> 81
Met Ala Ser Arg Ser Met Arg Leu Leu Leu Leu Leu Ser Cys Leu
1 5 10 15

Ala Lys Thr Gly Val Leu Gly Asp Ile Ile Met Arg Pro Ser Cys
20 25 30

Ala Pro Gly Trp Phe Tyr His Lys Ser Asn Cys Tyr Gly Tyr Phe
35 40 45

Arg Lys Leu Arg Asn Trp Ser Asp Ala Glu Leu Glu Cys Gln Ser
50 55 60

Tyr Gly Asn Gly Ala His Leu Ala Ser Ile Leu Ser Leu Lys Glu
65 70 75

1191

Ala Ser Thr Ile Ala Glu Tyr Ile Ser Gly Tyr Gln Arg Ser Gln
80 85 90

Pro Ile Trp Ile Gly Leu His Asp Pro Gln Lys Arg Gln Gln Trp
95 100 105

Gln Trp Ile Asp Gly Ala Met Tyr Leu Tyr Arg Ser Trp Ser Gly
110 115 120

Lys Ser Met Gly Gly Asn Lys His Cys Ala Glu Ile Ser Ser Asn
125 130 135

Asn Asn Phe Leu Thr Trp Ser Ser Asn Glu Cys Asn Lys Arg Gln
140 145 150

His Phe Leu Cys Lys Tyr Arg Pro
155

<210> 82
<211> 84
<212> PRT
<213> Homo sapiens

<400> 82
Met Ala Thr Met Glu Asn Lys Val Ile Cys Ala Leu Val Leu Val
1 5 10 15

Ser Met Leu Ala Leu Gly Thr Leu Ala Glu Ala Gln Thr Glu Thr
20 25 30

Cys Thr Val Ala Pro Arg Glu Arg Gln Asn Cys Gly Phe Pro Gly
35 40 45

Val Thr Pro Ser Gln Cys Ala Asn Lys Gly Cys Cys Phe Asp Asp
50 55 60

Thr Val Arg Gly Val Pro Trp Cys Phe Tyr Pro Asn Thr Ile Asp
65 70 75

Val Pro Pro Glu Glu Glu Cys Glu Phe
80

<210> 83
<211> 680
<212> PRT
<213> Homo sapiens

<400> 83
Met Leu Pro Gln Ile Pro Phe Leu Leu Leu Val Ser Leu Asn Leu
1 5 10 15

Val His Gly Val Phe Tyr Ala Glu Arg Tyr Gln Thr Pro Thr Gly
20 25 30

Ile Lys Gly Pro Leu Pro Asn Thr Lys Thr Gln Phe Phe Ile Pro
35 40 45

Tyr Thr Ile Lys Ser Lys Gly Ile Ala Val Arg Gly Glu Gln Gly
50 55 60

Thr Pro Gly Pro Pro Gly Pro Ala Gly Pro Arg Gly His Pro Gly
65 70 75

Pro Ser Gly Pro Pro Gly Lys Pro Gly Tyr Gly Ser Pro Gly Leu
80 85 90

Gln Gly Glu Pro Gly Leu Pro Gly Pro Pro Gly Pro Ser Ala Val
95 100 105

Gly Lys Pro Gly Val Pro Gly Leu Pro Gly Lys Pro Gly Glu Arg
110 115 120

Gly Pro Tyr Gly Pro Lys Gly Asp Val Gly Pro Ala Gly Leu Pro
125 130 135

Gly Pro Arg Gly Pro Pro Gly Pro Pro Gly Ile Pro Gly Pro Ala

```
                  140                 145                 150

Gly Ile Ser Val Pro Gly Lys Pro Gly Gln Gln Gly Pro Thr Gly
                  155                 160                 165

Ala Pro Gly Pro Arg Gly Phe Pro Gly Glu Lys Gly Ala Pro Gly
                  170                 175                 180

Val Pro Gly Met Asn Gly Gln Lys Gly Glu Met Gly Tyr Gly Ala
                  185                 190                 195

Pro Gly Arg Pro Gly Glu Arg Gly Leu Pro Gly Pro Gln Gly Pro
                  200                 205                 210

Thr Gly Pro Ser Gly Pro Pro Gly Val Gly Lys Arg Gly Glu Asn
                  215                 220                 225

Gly Val Pro Gly Gln Pro Gly Ile Lys Gly Asp Arg Gly Phe Pro
                  230                 235                 240

Gly Glu Met Gly Pro Ile Gly Pro Pro Gly Pro Gln Gly Pro Pro
                  245                 250                 255

Gly Glu Arg Gly Pro Glu Gly Ile Gly Lys Pro Gly Ala Ala Gly
                  260                 265                 270

Ala Pro Gly Gln Pro Gly Ile Pro Gly Thr Lys Gly Leu Pro Gly
                  275                 280                 285

Ala Pro Gly Ile Ala Gly Pro Pro Gly Pro Pro Gly Phe Gly Lys
                  290                 295                 300

Pro Gly Leu Pro Gly Leu Lys Gly Glu Arg Gly Pro Ala Gly Leu
                  305                 310                 315

Pro Gly Gly Pro Gly Ala Lys Gly Glu Gln Gly Pro Ala Gly Leu
                  320                 325                 330

Pro Gly Lys Pro Gly Leu Thr Gly Pro Pro Gly Asn Met Gly Pro
                  335                 340                 345

Gln Gly Pro Lys Gly Ile Pro Gly Ser His Gly Leu Pro Gly Pro
                  350                 355                 360

Lys Gly Glu Thr Gly Pro Ala Gly Pro Ala Gly Tyr Pro Gly Ala
                  365                 370                 375

Lys Gly Glu Arg Gly Ser Pro Gly Ser Asp Gly Lys Pro Gly Tyr
                  380                 385                 390

Pro Gly Lys Pro Gly Leu Asp Gly Pro Lys Gly Asn Pro Gly Leu
                  395                 400                 405

Pro Gly Pro Lys Gly Asp Pro Gly Val Gly Gly Pro Pro Gly Leu
                  410                 415                 420

Pro Gly Pro Val Gly Pro Ala Gly Ala Lys Gly Met Pro Gly His
                  425                 430                 435

Asn Gly Glu Ala Gly Pro Arg Gly Ala Pro Gly Ile Pro Gly Thr
                  440                 445                 450

Arg Gly Pro Ile Gly Pro Pro Gly Ile Pro Gly Phe Pro Gly Ser
                  455                 460                 465

Lys Gly Asp Pro Gly Ser Pro Gly Pro Pro Gly Pro Ala Gly Ile
                  470                 475                 480

Ala Thr Lys Gly Leu Asn Gly Pro Thr Gly Pro Pro Gly Pro Pro
                  485                 490                 495

Gly Pro Arg Gly Pro Ser Gly Glu Pro Gly Leu Pro Gly Pro Pro
                  500                 505                 510

Gly Pro Pro Gly Pro Pro Gly Gln Ala Val Met Pro Glu Gly Phe
                  515                 520                 525
```

```
Ile Lys Ala Gly Gln Arg Pro Ser Leu Ser Gly Thr Pro Leu Val
                530                 535                 540

Ser Ala Asn Gln Gly Val Thr Gly Met Pro Val Ser Ala Phe Thr
                545                 550                 555

Val Ile Leu Ser Lys Ala Tyr Pro Ala Ile Gly Thr Pro Ile Pro
                560                 565                 570

Phe Asp Lys Ile Leu Tyr Asn Arg Gln Gln His Tyr Asp Pro Arg
                575                 580                 585

Thr Gly Ile Phe Thr Cys Gln Ile Pro Gly Ile Tyr Tyr Phe Ser
                590                 595                 600

Tyr His Val His Val Lys Gly Thr His Val Trp Val Gly Leu Tyr
                605                 610                 615

Lys Asn Gly Thr Pro Val Met Tyr Thr Tyr Asp Glu Tyr Thr Lys
                620                 625                 630

Gly Tyr Leu Asp Gln Ala Ser Gly Ser Ala Ile Ile Asp Leu Thr
                635                 640                 645

Glu Asn Asp Gln Val Trp Leu Gln Leu Pro Asn Ala Glu Ser Asn
                650                 655                 660

Gly Leu Tyr Ser Ser Glu Tyr Val His Ser Ser Phe Ser Gly Phe
                665                 670                 675

Leu Val Ala Pro Met
                680
```

<210> 84
<211> 542
<212> PRT
<213> Homo sapiens

<400> 84

```
Met Pro Leu Pro Pro Leu Leu Leu Leu Leu Leu Ala Ala Pro Trp
  1                 5                  10                  15

Gly Arg Ala Val Pro Cys Val Ser Gly Gly Leu Pro Lys Pro Ala
                 20                  25                  30

Asn Ile Thr Phe Leu Ser Ile Asn Met Lys Asn Val Leu Gln Trp
                 35                  40                  45

Thr Pro Pro Glu Gly Leu Gln Gly Val Lys Val Thr Tyr Thr Val
                 50                  55                  60

Gln Tyr Phe Ile Tyr Gly Gln Lys Lys Trp Leu Asn Lys Ser Glu
                 65                  70                  75

Cys Arg Asn Ile Asn Arg Thr Tyr Cys Asp Leu Ser Ala Glu Thr
                 80                  85                  90

Ser Asp Tyr Glu His Gln Tyr Tyr Ala Lys Val Lys Ala Ile Trp
                 95                 100                 105

Gly Thr Lys Cys Ser Lys Trp Ala Glu Ser Gly Arg Phe Tyr Pro
                110                 115                 120

Phe Leu Glu Thr Gln Ile Gly Pro Pro Glu Val Ala Leu Thr Thr
                125                 130                 135

Asp Glu Lys Ser Ile Ser Val Val Leu Thr Ala Pro Glu Lys Trp
                140                 145                 150

Lys Arg Asn Pro Glu Asp Leu Pro Val Ser Met Gln Gln Ile Tyr
                155                 160                 165

Ser Asn Leu Lys Tyr Asn Val Ser Val Leu Asn Thr Lys Ser Asn
                170                 175                 180
```

```
Arg Thr Trp Ser Gln Cys Val Thr Asn His Thr Leu Val Leu Thr
                185                 190                 195

Trp Leu Glu Pro Asn Thr Leu Tyr Cys Val His Val Glu Ser Phe
                200                 205                 210

Val Pro Gly Pro Pro Arg Arg Ala Gln Pro Ser Glu Lys Gln Cys
                215                 220                 225

Ala Arg Thr Leu Lys Asp Gln Ser Ser Glu Phe Lys Ala Lys Ile
                230                 235                 240

Ile Phe Trp Tyr Val Leu Pro Ile Ser Ile Thr Val Phe Leu Phe
                245                 250                 255

Ser Val Met Gly Tyr Ser Ile Tyr Arg Tyr Ile His Val Gly Lys
                260                 265                 270

Glu Lys His Pro Ala Asn Leu Ile Leu Ile Tyr Gly Asn Glu Phe
                275                 280                 285

Asp Lys Arg Phe Phe Val Pro Ala Glu Lys Ile Val Ile Asn Phe
                290                 295                 300

Ile Thr Leu Asn Ile Ser Asp Asp Ser Lys Ile Ser His Gln Asp
                305                 310                 315

Met Ser Leu Leu Gly Lys Ser Ser Asp Val Ser Ser Leu Asn Asp
                320                 325                 330

Pro Gln Pro Ser Gly Asn Leu Arg Pro Pro Gln Glu Glu Glu Glu
                335                 340                 345

Val Lys His Leu Gly Tyr Ala Ser His Leu Met Glu Ile Phe Cys
                350                 355                 360

Asp Ser Glu Glu Asn Thr Glu Gly Thr Ser Leu Thr Gln Gln Glu
                365                 370                 375

Ser Leu Ser Arg Thr Ile Pro Pro Asp Lys Thr Val Ile Glu Tyr
                380                 385                 390

Glu Tyr Asp Val Arg Thr Thr Asp Ile Cys Ala Gly Pro Glu Glu
                395                 400                 405

Gln Glu Leu Ser Leu Gln Glu Glu Val Ser Thr Gln Gly Thr Leu
                410                 415                 420

Leu Glu Ser Gln Ala Ala Leu Ala Val Leu Gly Pro Gln Thr Leu
                425                 430                 435

Gln Tyr Ser Tyr Thr Pro Gln Leu Gln Asp Leu Asp Pro Leu Ala
                440                 445                 450

Gln Glu His Thr Asp Ser Glu Glu Gly Pro Glu Glu Glu Pro Ser
                455                 460                 465

Thr Thr Leu Val Asp Trp Asp Pro Gln Thr Gly Arg Leu Cys Ile
                470                 475                 480

Pro Ser Leu Ser Ser Phe Asp Gln Asp Ser Glu Gly Cys Glu Pro
                485                 490                 495

Ser Glu Gly Asp Gly Leu Gly Glu Glu Gly Leu Leu Ser Arg Leu
                500                 505                 510

Tyr Glu Glu Pro Ala Pro Asp Arg Pro Pro Gly Glu Asn Glu Thr
                515                 520                 525

Tyr Leu Met Gln Phe Met Glu Glu Trp Gly Leu Tyr Val Gln Met
                530                 535                 540

Glu Asn
```

<210> 85

<211> 553
<212> PRT
<213> Homo sapiens

<400> 85

```
Met Arg Ala Pro Gly Arg Pro Ala Leu Arg Pro Leu Pro Leu Pro
  1               5                  10                  15

Pro Leu Leu Leu Leu Leu Leu Ala Ala Pro Trp Gly Arg Ala Val
             20                  25                  30

Pro Cys Val Ser Gly Gly Leu Pro Lys Pro Ala Asn Ile Thr Phe
             35                  40                  45

Leu Ser Ile Asn Met Lys Asn Val Leu Gln Trp Thr Pro Pro Glu
             50                  55                  60

Gly Leu Gln Gly Val Lys Val Thr Tyr Thr Val Gln Tyr Phe Ile
             65                  70                  75

Tyr Gly Gln Lys Lys Trp Leu Asn Lys Ser Glu Cys Arg Asn Ile
             80                  85                  90

Asn Arg Thr Tyr Cys Asp Leu Ser Ala Glu Thr Ser Asp Tyr Glu
             95                 100                 105

His Gln Tyr Tyr Ala Lys Val Lys Ala Ile Trp Gly Thr Lys Cys
            110                 115                 120

Ser Lys Trp Ala Glu Ser Gly Arg Phe Tyr Pro Phe Leu Glu Thr
            125                 130                 135

Gln Ile Gly Pro Pro Glu Val Ala Leu Thr Thr Asp Glu Lys Ser
            140                 145                 150

Ile Ser Val Val Leu Thr Ala Pro Glu Lys Trp Lys Arg Asn Pro
            155                 160                 165

Glu Asp Leu Pro Val Ser Met Gln Gln Ile Tyr Ser Asn Leu Lys
            170                 175                 180

Tyr Asn Val Ser Val Leu Asn Thr Lys Ser Asn Arg Thr Trp Ser
            185                 190                 195

Gln Cys Val Thr Asn His Thr Leu Val Leu Thr Trp Leu Glu Pro
            200                 205                 210

Asn Thr Leu Tyr Cys Val His Val Glu Ser Phe Val Pro Gly Pro
            215                 220                 225

Pro Arg Arg Ala Gln Pro Ser Glu Lys Gln Cys Ala Arg Thr Leu
            230                 235                 240

Lys Asp Gln Ser Ser Glu Phe Lys Ala Lys Ile Ile Phe Trp Tyr
            245                 250                 255

Val Leu Pro Ile Ser Ile Thr Val Phe Leu Phe Ser Val Met Gly
            260                 265                 270

Tyr Ser Ile Tyr Arg Tyr Ile His Val Gly Lys Glu Lys His Pro
            275                 280                 285

Ala Asn Leu Ile Leu Ile Tyr Gly Asn Glu Phe Asp Lys Arg Phe
            290                 295                 300

Phe Val Pro Ala Glu Lys Ile Val Ile Asn Phe Ile Thr Leu Asn
            305                 310                 315

Ile Ser Asp Asp Ser Lys Ile Ser His Gln Asp Met Ser Leu Leu
            320                 325                 330

Gly Lys Ser Ser Asp Val Ser Ser Leu Asn Asp Pro Gln Pro Ser
            335                 340                 345

Gly Asn Leu Arg Pro Pro Gln Glu Glu Glu Glu Val Lys His Leu
            350                 355                 360
```

```
 Gly Tyr Ala Ser His Leu Met Glu Ile Phe Cys Asp Ser Glu Glu
                  365                 370                 375

 Asn Thr Glu Gly Thr Ser Leu Thr Gln Gln Glu Ser Leu Ser Arg
                  380                 385                 390

 Thr Ile Pro Pro Asp Lys Thr Val Ile Glu Tyr Glu Tyr Asp Val
                  395                 400                 405

 Arg Thr Thr Asp Ile Cys Ala Gly Pro Glu Glu Gln Glu Leu Ser
                  410                 415                 420

 Leu Gln Glu Glu Val Ser Thr Gln Gly Thr Leu Leu Glu Ser Gln
                  425                 430                 435

 Ala Ala Leu Ala Val Leu Gly Pro Gln Thr Leu Gln Tyr Ser Tyr
                  440                 445                 450

 Thr Pro Gln Leu Gln Asp Leu Asp Pro Leu Ala Gln Glu His Thr
                  455                 460                 465

 Asp Ser Glu Glu Gly Pro Glu Glu Glu Pro Ser Thr Thr Leu Val
                  470                 475                 480

 Asp Trp Asp Pro Gln Thr Gly Arg Leu Cys Ile Pro Ser Leu Ser
                  485                 490                 495

 Ser Phe Asp Gln Asp Ser Glu Gly Cys Glu Pro Ser Glu Gly Asp
                  500                 505                 510

 Gly Leu Gly Glu Glu Gly Leu Leu Ser Arg Leu Tyr Glu Glu Pro
                  515                 520                 525

 Ala Pro Asp Arg Pro Pro Gly Glu Asn Glu Thr Tyr Leu Met Gln
                  530                 535                 540

 Phe Met Glu Glu Trp Gly Leu Tyr Val Gln Met Glu Asn
                  545                 550

<210> 86
<211> 95
<212> PRT
<213> Homo sapiens

<400> 86
 Met Ala Ala Ala Arg Leu Cys Leu Ser Leu Leu Leu Leu Ser Thr
   1               5                  10                  15

 Cys Val Ala Leu Leu Leu Gln Pro Leu Leu Gly Ala Gln Gly Ala
                  20                  25                  30

 Pro Leu Glu Pro Val Tyr Pro Gly Asp Asn Ala Thr Pro Glu Gln
                  35                  40                  45

 Met Ala Gln Tyr Ala Ala Asp Leu Arg Arg Tyr Ile Asn Met Leu
                  50                  55                  60

 Thr Arg Pro Arg Tyr Gly Lys Arg His Lys Glu Asp Thr Leu Ala
                  65                  70                  75

 Phe Ser Glu Trp Gly Ser Pro His Ala Ala Val Pro Arg Glu Leu
                  80                  85                  90

 Ser Pro Leu Asp Leu
                  95

<210> 87
<211> 344
<212> PRT
<213> Homo sapiens

<400> 87
 Met Gly Pro Pro Ser Ala Pro Pro Cys Arg Leu His Val Pro Trp
   1               5                  10                  15
```

Lys Glu Val Leu Leu Thr Ala Ser Leu Leu Thr Phe Trp Asn Pro
20 25 30

Pro Thr Thr Ala Lys Leu Thr Ile Glu Ser Thr Pro Phe Asn Val
35 40 45

Ala Glu Gly Lys Glu Val Leu Leu Leu Ala His Asn Leu Pro Gln
50 55 60

Asn Arg Ile Gly Tyr Ser Trp Tyr Lys Gly Glu Arg Val Asp Gly
65 70 75

Asn Ser Leu Ile Val Gly Tyr Val Ile Gly Thr Gln Gln Ala Thr
80 85 90

Pro Gly Pro Ala Tyr Ser Gly Arg Glu Thr Ile Tyr Pro Asn Ala
95 100 105

Ser Leu Leu Ile Gln Asn Val Thr Gln Asn Asp Thr Gly Phe Tyr
110 115 120

Thr Leu Gln Val Ile Lys Ser Asp Leu Val Asn Glu Glu Ala Thr
125 130 135

Gly Gln Phe His Val Tyr Pro Glu Leu Pro Lys Pro Ser Ile Ser
140 145 150

Ser Asn Asn Ser Asn Pro Val Glu Asp Lys Asp Ala Val Ala Phe
155 160 165

Thr Cys Glu Pro Glu Val Gln Asn Thr Thr Tyr Leu Trp Trp Val
170 175 180

Asn Gly Gln Ser Leu Pro Val Ser Pro Arg Leu Gln Leu Ser Asn
185 190 195

Gly Asn Met Thr Leu Thr Leu Leu Ser Val Lys Arg Asn Asp Ala
200 205 210

Gly Ser Tyr Glu Cys Glu Ile Gln Asn Pro Ala Ser Ala Asn Arg
215 220 225

Ser Asp Pro Val Thr Leu Asn Val Leu Tyr Gly Pro Asp Gly Pro
230 235 240

Thr Ile Ser Pro Ser Lys Ala Asn Tyr Arg Pro Gly Glu Asn Leu
245 250 255

Asn Leu Ser Cys His Ala Ala Ser Asn Pro Pro Ala Gln Tyr Ser
260 265 270

Trp Phe Ile Asn Gly Thr Phe Gln Gln Ser Thr Gln Glu Leu Phe
275 280 285

Ile Pro Asn Ile Thr Val Asn Asn Ser Gly Ser Tyr Met Cys Gln
290 295 300

Ala His Asn Ser Ala Thr Gly Leu Asn Arg Thr Thr Val Thr Met
305 310 315

Ile Thr Val Ser Gly Ser Ala Pro Val Leu Ser Ala Val Ala Thr
320 325 330

Val Gly Ile Thr Ile Gly Val Leu Ala Arg Val Ala Leu Ile
335 340

<210> 88
<211> 987
<212> PRT
<213> Homo sapiens

<400> 88
Met Glu Leu Arg Val Leu Leu Cys Trp Ala Ser Leu Ala Ala Ala
1 5 10 15

Leu Glu Glu Thr Leu Leu Asn Thr Lys Leu Glu Thr Ala Asp Leu

```
                  20                  25                  30

Lys Trp Val Thr Phe Pro Gln Val Asp Gly Gln Trp Glu Glu Leu
                  35                  40                  45

Ser Gly Leu Asp Glu Glu Gln His Ser Val Arg Thr Tyr Glu Val
                  50                  55                  60

Cys Asp Val Gln Arg Ala Pro Gly Gln Ala His Trp Leu Arg Thr
                  65                  70                  75

Gly Trp Val Pro Arg Arg Gly Ala Val His Val Tyr Ala Thr Leu
                  80                  85                  90

Arg Phe Thr Met Leu Glu Cys Leu Ser Leu Pro Arg Ala Gly Arg
                  95                 100                 105

Ser Cys Lys Glu Thr Phe Thr Val Phe Tyr Tyr Glu Ser Asp Ala
                 110                 115                 120

Asp Thr Ala Thr Ala Leu Thr Pro Ala Trp Met Glu Asn Pro Tyr
                 125                 130                 135

Ile Lys Val Asp Thr Val Ala Ala Glu His Leu Thr Arg Lys Arg
                 140                 145                 150

Pro Gly Ala Glu Ala Thr Gly Lys Val Asn Val Lys Thr Leu Arg
                 155                 160                 165

Leu Gly Pro Leu Ser Lys Ala Gly Phe Tyr Leu Ala Phe Gln Asp
                 170                 175                 180

Gln Gly Ala Cys Met Ala Leu Leu Ser Leu His Leu Phe Tyr Lys
                 185                 190                 195

Lys Cys Ala Gln Leu Thr Val Asn Leu Thr Arg Phe Pro Glu Thr
                 200                 205                 210

Val Pro Arg Glu Leu Val Val Pro Val Ala Gly Ser Cys Val Val
                 215                 220                 225

Asp Ala Val Pro Ala Pro Gly Pro Ser Pro Ser Leu Tyr Cys Arg
                 230                 235                 240

Glu Asp Gly Gln Trp Ala Glu Gln Pro Val Thr Gly Cys Ser Cys
                 245                 250                 255

Ala Pro Gly Phe Glu Ala Ala Glu Gly Asn Thr Lys Cys Arg Ala
                 260                 265                 270

Cys Ala Gln Gly Thr Phe Lys Pro Leu Ser Gly Glu Gly Ser Cys
                 275                 280                 285

Gln Pro Cys Pro Ala Asn Ser His Ser Asn Thr Ile Gly Ser Ala
                 290                 295                 300

Val Cys Gln Cys Arg Val Gly Tyr Phe Arg Ala Arg Thr Asp Pro
                 305                 310                 315

Arg Gly Ala Pro Cys Thr Thr Pro Pro Ser Ala Pro Arg Ser Val
                 320                 325                 330

Val Ser Arg Leu Asn Gly Ser Ser Leu His Leu Glu Trp Ser Ala
                 335                 340                 345

Pro Leu Glu Ser Gly Gly Arg Glu Asp Leu Thr Tyr Ala Leu Arg
                 350                 355                 360

Cys Arg Glu Cys Arg Pro Gly Gly Ser Cys Ala Pro Cys Gly Gly
                 365                 370                 375

Asp Leu Thr Phe Asp Pro Gly Pro Arg Asp Leu Val Glu Pro Trp
                 380                 385                 390

Val Val Val Arg Gly Leu Arg Pro Asp Phe Thr Tyr Thr Phe Glu
                 395                 400                 405
```

Val Thr Ala Leu Asn Gly Val Ser Ser Leu Ala Thr Gly Pro Val
410 415 420

Pro Phe Glu Pro Val Asn Val Thr Thr Asp Arg Glu Val Pro Pro
425 430 435

Ala Val Ser Asp Ile Arg Val Thr Arg Ser Ser Pro Ser Ser Leu
440 445 450

Ser Leu Ala Trp Ala Val Pro Arg Ala Pro Ser Gly Ala Val Leu
455 460 465

Asp Tyr Glu Val Lys Tyr His Glu Lys Gly Ala Glu Gly Pro Ser
470 475 480

Ser Val Arg Phe Leu Lys Thr Ser Glu Asn Arg Ala Glu Leu Arg
485 490 495

Gly Leu Lys Arg Gly Ala Ser Tyr Leu Val Gln Val Arg Ala Arg
500 505 510

Ser Glu Ala Gly Tyr Gly Pro Phe Gly Gln Glu His His Ser Gln
515 520 525

Thr Gln Leu Asp Glu Ser Glu Gly Trp Arg Glu Gln Leu Ala Leu
530 535 540

Ile Ala Gly Thr Ala Val Val Gly Val Val Leu Val Leu Val Val
545 550 555

Ile Val Val Ala Val Leu Cys Leu Arg Lys Gln Ser Asn Gly Arg
560 565 570

Glu Ala Glu Tyr Ser Asp Lys His Gly Gln Tyr Leu Ile Gly His
575 580 585

Gly Thr Lys Val Tyr Ile Asp Pro Phe Thr Tyr Glu Asp Pro Asn
590 595 600

Glu Ala Val Arg Glu Phe Ala Lys Glu Ile Asp Val Ser Tyr Val
605 610 615

Lys Ile Glu Glu Val Ile Gly Ala Gly Glu Phe Gly Glu Val Cys
620 625 630

Arg Gly Arg Leu Lys Ala Pro Gly Lys Lys Glu Ser Cys Val Ala
635 640 645

Ile Lys Thr Leu Lys Gly Gly Tyr Thr Glu Arg Gln Arg Arg Glu
650 655 660

Phe Leu Ser Glu Ala Ser Ile Met Gly Gln Phe Glu His Pro Asn
665 670 675

Ile Ile Arg Leu Glu Gly Val Val Thr Asn Ser Met Pro Val Met
680 685 690

Ile Leu Thr Glu Phe Met Glu Asn Gly Ala Leu Asp Ser Phe Leu
695 700 705

Arg Leu Asn Asp Gly Gln Phe Thr Val Ile Gln Leu Val Gly Met
710 715 720

Leu Arg Gly Ile Ala Ser Gly Met Arg Tyr Leu Ala Glu Met Ser
725 730 735

Tyr Val His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Asn Ser
740 745 750

Asn Leu Val Cys Lys Val Ser Asp Phe Gly Leu Ser Arg Phe Leu
755 760 765

Glu Glu Asn Ser Ser Asp Pro Thr Tyr Thr Ser Ser Leu Gly Gly
770 775 780

```
Lys Ile Pro Ile Arg Trp Thr Ala Pro Glu Ala Ile Ala Phe Arg
                785                 790                 795

Lys Phe Thr Ser Ala Ser Asp Ala Trp Ser Tyr Gly Ile Val Met
                800                 805                 810

Trp Glu Val Met Ser Phe Gly Glu Arg Pro Tyr Trp Asp Met Ser
                815                 820                 825

Asn Gln Asp Val Ile Asn Ala Ile Glu Gln Asp Tyr Arg Leu Pro
                830                 835                 840

Pro Pro Pro Asp Cys Pro Thr Ser Leu His Gln Leu Met Leu Asp
                845                 850                 855

Cys Trp Gln Lys Asp Arg Asn Ala Arg Pro Arg Phe Pro Gln Val
                860                 865                 870

Val Ser Ala Leu Asp Lys Met Ile Arg Asn Pro Ala Ser Leu Lys
                875                 880                 885

Ile Val Ala Arg Glu Asn Gly Gly Ala Ser His Pro Leu Leu Asp
                890                 895                 900

Gln Arg Gln Pro His Tyr Ser Ala Phe Gly Ser Val Gly Glu Trp
                905                 910                 915

Leu Arg Ala Ile Lys Met Gly Arg Tyr Glu Glu Ser Phe Ala Ala
                920                 925                 930

Ala Gly Phe Gly Ser Phe Glu Leu Val Ser Gln Ile Ser Ala Glu
                935                 940                 945

Asp Leu Leu Arg Ile Gly Val Thr Leu Ala Gly His Gln Lys Lys
                950                 955                 960

Ile Leu Ala Ser Val Gln His Met Lys Ser Gln Ala Lys Pro Gly
                965                 970                 975

Thr Pro Gly Gly Thr Gly Gly Pro Ala Pro Gln Tyr
                980                 985

<210> 89
<211> 911
<212> PRT
<213> Homo sapiens

<400> 89
 Met Ala Gln Leu Phe Leu Pro Leu Leu Ala Ala Leu Val Leu Ala
   1               5                  10                  15

 Gln Ala Pro Ala Ala Leu Ala Asp Val Leu Glu Gly Asp Ser Ser
                20                  25                  30

 Glu Asp Arg Ala Phe Arg Val Arg Ile Ala Gly Asp Ala Pro Leu
                35                  40                  45

 Gln Gly Val Leu Gly Gly Ala Leu Thr Ile Pro Cys His Val His
                50                  55                  60

 Tyr Leu Arg Pro Pro Pro Ser Arg Arg Ala Val Leu Gly Ser Pro
                65                  70                  75

 Arg Val Lys Trp Thr Phe Leu Ser Arg Gly Arg Glu Ala Glu Val
                80                  85                  90

 Leu Val Ala Arg Gly Val Arg Val Lys Val Asn Glu Ala Tyr Arg
                95                 100                 105

 Phe Arg Val Ala Leu Pro Ala Tyr Pro Ala Ser Leu Thr Asp Val
               110                 115                 120

 Ser Leu Ala Leu Ser Glu Leu Arg Pro Asn Asp Ser Gly Ile Tyr
               125                 130                 135

 Arg Cys Glu Val Gln His Gly Ile Asp Asp Ser Ser Asp Ala Val
```

```
                  140                 145                 150

Glu Val Lys Val Lys Gly Val Val Phe Leu Tyr Arg Glu Gly Ser
                  155                 160                 165

Ala Arg Tyr Ala Phe Ser Phe Ser Gly Ala Gln Glu Ala Cys Ala
                  170                 175                 180

Arg Ile Gly Ala His Ile Ala Thr Pro Glu Gln Leu Tyr Ala Ala
                  185                 190                 195

Tyr Leu Gly Gly Tyr Glu Gln Cys Asp Ala Gly Trp Leu Ser Asp
                  200                 205                 210

Gln Thr Val Arg Tyr Pro Ile Gln Thr Pro Arg Glu Ala Cys Tyr
                  215                 220                 225

Gly Asp Met Asp Gly Phe Pro Gly Val Arg Asn Tyr Gly Val Val
                  230                 235                 240

Asp Pro Asp Asp Leu Tyr Asp Val Tyr Cys Tyr Ala Glu Asp Leu
                  245                 250                 255

Asn Gly Glu Leu Phe Leu Gly Asp Pro Pro Glu Lys Leu Thr Leu
                  260                 265                 270

Glu Glu Ala Arg Ala Tyr Cys Gln Glu Arg Gly Ala Glu Ile Ala
                  275                 280                 285

Thr Thr Gly Gln Leu Tyr Ala Ala Trp Asp Gly Gly Leu Asp His
                  290                 295                 300

Cys Ser Pro Gly Trp Leu Ala Asp Gly Ser Val Arg Tyr Pro Ile
                  305                 310                 315

Val Thr Pro Ser Gln Arg Cys Gly Gly Gly Leu Pro Gly Val Lys
                  320                 325                 330

Thr Leu Phe Leu Phe Pro Asn Gln Thr Gly Phe Pro Asn Lys His
                  335                 340                 345

Ser Arg Phe Asn Val Tyr Cys Phe Arg Asp Ser Ala Gln Pro Ser
                  350                 355                 360

Ala Ile Pro Glu Ala Ser Asn Pro Ala Ser Asn Pro Ala Ser Asp
                  365                 370                 375

Gly Leu Glu Ala Ile Val Thr Val Thr Glu Thr Leu Glu Glu Leu
                  380                 385                 390

Gln Leu Pro Gln Glu Ala Thr Glu Ser Glu Ser Arg Gly Ala Ile
                  395                 400                 405

Tyr Ser Ile Pro Ile Met Glu Asp Gly Gly Gly Gly Ser Ser Thr
                  410                 415                 420

Pro Glu Asp Pro Ala Glu Ala Pro Arg Thr Leu Leu Glu Phe Glu
                  425                 430                 435

Thr Gln Ser Met Val Pro Pro Thr Gly Phe Ser Glu Glu Glu Gly
                  440                 445                 450

Lys Ala Leu Glu Glu Glu Glu Lys Tyr Glu Asp Glu Glu Glu Lys
                  455                 460                 465

Glu Glu Glu Glu Glu Glu Glu Glu Val Glu Asp Glu Ala Leu Trp
                  470                 475                 480

Ala Trp Pro Ser Glu Leu Ser Ser Pro Gly Pro Glu Ala Ser Leu
                  485                 490                 495

Pro Thr Glu Pro Ala Ala Gln Glu Lys Ser Leu Ser Gln Ala Pro
                  500                 505                 510

Ala Arg Ala Val Leu Gln Pro Gly Ala Ser Pro Leu Pro Asp Gly
                  515                 520                 525
```

1203
1202

Glu Ser Glu Ala Ser Arg Pro Pro Arg Val His Gly Pro Pro Thr
530 535 540

Glu Thr Leu Pro Thr Pro Arg Glu Arg Asn Leu Ala Ser Pro Ser
545 550 555

Pro Ser Thr Leu Val Glu Ala Arg Glu Val Gly Glu Ala Thr Gly
560 565 570

Gly Pro Glu Leu Ser Gly Val Pro Arg Gly Glu Ser Glu Glu Thr
575 580 585

Gly Ser Ser Glu Gly Ala Pro Ser Leu Leu Pro Ala Thr Arg Ala
590 595 600

Pro Glu Gly Thr Arg Glu Leu Glu Ala Pro Ser Glu Asp Asn Ser
605 610 615

Gly Arg Thr Ala Pro Ala Gly Thr Ser Val Gln Ala Gln Pro Val
620 625 630

Leu Pro Thr Asp Ser Ala Ser Arg Gly Gly Val Ala Val Val Pro
635 640 645

Ala Ser Gly Asp Cys Val Pro Ser Pro Cys His Asn Gly Gly Thr
650 655 660

Cys Leu Glu Glu Glu Glu Gly Val Arg Cys Leu Cys Leu Pro Gly
665 670 675

Tyr Gly Gly Asp Leu Cys Asp Val Gly Leu Arg Phe Cys Asn Pro
680 685 690

Gly Trp Asp Ala Phe Gln Gly Ala Cys Tyr Lys His Phe Ser Thr
695 700 705

Arg Arg Ser Trp Glu Glu Ala Glu Thr Gln Cys Arg Met Tyr Gly
710 715 720

Ala His Leu Ala Ser Ile Ser Thr Pro Glu Glu Gln Asp Phe Ile
725 730 735

Asn Asn Arg Tyr Arg Glu Tyr Gln Trp Ile Gly Leu Asn Asp Arg
740 745 750

Thr Ile Glu Gly Asp Phe Leu Trp Ser Asp Gly Val Pro Leu Leu
755 760 765

Tyr Glu Asn Trp Asn Pro Gly Gln Pro Asp Ser Tyr Phe Leu Ser
770 775 780

Gly Glu Asn Cys Val Val Met Val Trp His Asp Gln Gly Gln Trp
785 790 795

Ser Asp Val Pro Cys Asn Tyr His Leu Ser Tyr Thr Cys Lys Met
800 805 810

Gly Leu Val Ser Cys Gly Pro Pro Pro Glu Leu Pro Leu Ala Gln
815 820 825

Val Phe Gly Arg Pro Arg Leu Arg Tyr Glu Val Asp Thr Val Leu
830 835 840

Arg Tyr Arg Cys Arg Glu Gly Leu Ala Gln Arg Asn Leu Pro Leu
845 850 855

Ile Arg Cys Gln Glu Asn Gly Arg Trp Glu Ala Pro Gln Ile Ser
860 865 870

Cys Val Pro Arg Arg Pro Ala Arg Ala Leu His Pro Glu Glu Asp
875 880 885

Pro Glu Gly Arg Gln Gly Arg Leu Leu Gly Arg Trp Lys Ala Leu
890 895 900

1207
1803

Leu Ile Pro Pro Ser Ser Pro Met Pro Gly Pro
905 910

<210> 90
<211> 671
<212> PRT
<213> Homo sapiens

<400> 90
Met Ala Gln Leu Phe Leu Pro Leu Leu Ala Ala Leu Val Leu Ala
1 5 10 15

Gln Ala Pro Ala Ala Leu Ala Asp Val Leu Glu Gly Asp Ser Ser
20 25 30

Glu Asp Arg Ala Phe Arg Val Arg Ile Ala Gly Asp Ala Pro Leu
35 40 45

Gln Gly Val Leu Gly Gly Ala Leu Thr Ile Pro Cys His Val His
50 55 60

Tyr Leu Arg Pro Pro Pro Ser Arg Arg Ala Val Leu Gly Ser Pro
65 70 75

Arg Val Lys Trp Thr Phe Leu Ser Arg Gly Arg Glu Ala Glu Val
80 85 90

Leu Val Ala Arg Gly Val Arg Val Lys Val Asn Glu Ala Tyr Arg
95 100 105

Phe Arg Val Ala Leu Pro Ala Tyr Pro Ala Ser Leu Thr Asp Val
110 115 120

Ser Leu Ala Leu Ser Glu Leu Arg Pro Asn Asp Ser Gly Ile Tyr
125 130 135

Arg Cys Glu Val Gln His Gly Ile Asp Asp Ser Ser Asp Ala Val
140 145 150

Glu Val Lys Val Lys Gly Val Val Phe Leu Tyr Arg Glu Gly Ser
155 160 165

Ala Arg Tyr Ala Phe Ser Phe Ser Gly Ala Gln Glu Ala Cys Ala
170 175 180

Arg Ile Gly Ala His Ile Ala Thr Pro Glu Gln Leu Tyr Ala Ala
185 190 195

Tyr Leu Gly Gly Tyr Glu Gln Cys Asp Ala Gly Trp Leu Ser Asp
200 205 210

Gln Thr Val Arg Tyr Pro Ile Gln Thr Pro Arg Glu Ala Cys Tyr
215 220 225

Gly Asp Met Asp Gly Phe Pro Gly Val Arg Asn Tyr Gly Val Val
230 235 240

Asp Pro Asp Asp Leu Tyr Asp Val Tyr Cys Tyr Ala Glu Asp Leu
245 250 255

Asn Gly Glu Leu Phe Leu Gly Asp Pro Pro Glu Lys Leu Thr Leu
260 265 270

Glu Glu Ala Arg Ala Tyr Cys Gln Glu Arg Gly Ala Glu Ile Ala
275 280 285

Thr Thr Gly Gln Leu Tyr Ala Ala Trp Asp Gly Gly Leu Asp His
290 295 300

Cys Ser Pro Gly Trp Leu Ala Asp Gly Ser Val Arg Tyr Pro Ile
305 310 315

Val Thr Pro Ser Gln Arg Cys Gly Gly Gly Leu Pro Gly Val Lys
320 325 330

Thr Leu Phe Leu Phe Pro Asn Gln Thr Gly Phe Pro Asn Lys His

1205
1204

335 340 345

Ser Arg Phe Asn Val Tyr Cys Phe Arg Asp Ser Ala Gln Pro Ser
350 355 360

Ala Ile Pro Glu Ala Ser Asn Pro Ala Ser Asn Pro Ala Ser Asp
365 370 375

Gly Leu Glu Ala Ile Val Thr Val Thr Glu Thr Leu Glu Glu Leu
380 385 390

Gln Leu Pro Gln Glu Ala Thr Glu Ser Glu Ser Arg Gly Ala Ile
395 400 405

Tyr Ser Ile Pro Ile Met Glu Asp Gly Gly Gly Gly Ser Ser Thr
410 415 420

Pro Glu Asp Pro Ala Glu Ala Pro Arg Thr Leu Leu Glu Phe Glu
425 430 435

Thr Gln Ser Met Val Pro Pro Thr Gly Phe Ser Glu Glu Glu Gly
440 445 450

Lys Ala Leu Glu Glu Glu Glu Lys Tyr Glu Asp Glu Glu Glu Lys
455 460 465

Glu Glu Glu Glu Glu Glu Glu Glu Val Glu Asp Glu Ala Leu Trp
470 475 480

Ala Trp Pro Ser Glu Leu Ser Ser Pro Gly Pro Glu Ala Ser Leu
485 490 495

Pro Thr Glu Pro Ala Ala Gln Glu Glu Ser Leu Ser Gln Ala Pro
500 505 510

Ala Arg Ala Val Leu Gln Pro Gly Ala Ser Pro Leu Pro Asp Gly
515 520 525

Glu Ser Glu Ala Ser Arg Pro Pro Arg Val His Gly Pro Pro Thr
530 535 540

Glu Thr Leu Pro Thr Pro Arg Glu Arg Asn Leu Ala Ser Pro Ser
545 550 555

Pro Ser Thr Leu Val Glu Ala Arg Glu Val Gly Glu Ala Thr Gly
560 565 570

Gly Pro Glu Leu Ser Gly Val Pro Arg Gly Glu Ser Glu Glu Thr
575 580 585

Gly Ser Ser Glu Gly Ala Pro Ser Leu Leu Pro Ala Thr Arg Ala
590 595 600

Pro Glu Gly Thr Arg Glu Leu Glu Ala Pro Ser Glu Asp Asn Ser
605 610 615

Gly Arg Thr Ala Pro Ala Gly Thr Ser Val Gln Ala Gln Pro Val
620 625 630

Leu Pro Thr Asp Ser Ala Ser Arg Gly Gly Val Ala Val Val Pro
635 640 645

Ala Ser Gly Asn Ser Ala Gln Gly Ser Thr Ala Leu Ser Ile Leu
650 655 660

Leu Leu Phe Phe Pro Leu Gln Leu Trp Val Thr
665 670

<210> 91
<211> 86
<212> PRT
<213> Homo sapiens

<400> 91
Met Ala Val Arg Gln Trp Val Ile Ala Leu Ala Leu Ala Ala Leu
1 5 10 15

L806
1205

Leu Val Val Asp Arg Glu Val Pro Val Ala Ala Gly Lys Leu Pro
20 25 30

Phe Ser Arg Met Pro Ile Cys Glu His Met Val Glu Ser Pro Thr
35 40 45

Cys Ser Gln Met Ser Asn Leu Val Cys Gly Thr Asp Gly Leu Thr
50 55 60

Tyr Thr Asn Glu Cys Gln Leu Cys Leu Ala Arg Ile Lys Thr Lys
65 70 75

Gln Asp Ile Gln Ile Met Lys Asp Gly Lys Cys
80 85

<210> 92
<211> 459
<212> PRT
<213> Homo sapiens

<400> 92
Met Ala Pro Leu Cys Pro Ser Pro Trp Leu Pro Leu Leu Ile Pro
1 5 10 15

Ala Pro Ala Pro Gly Leu Thr Val Gln Leu Leu Leu Ser Leu Leu
20 25 30

Leu Leu Val Pro Val His Pro Gln Arg Leu Pro Arg Met Gln Glu
35 40 45

Asp Ser Pro Leu Gly Gly Gly Ser Ser Gly Glu Asp Asp Pro Leu
50 55 60

Gly Glu Glu Asp Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu
65 70 75

Asp Pro Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro
80 85 90

Gly Glu Glu Asp Leu Pro Glu Val Lys Pro Lys Ser Glu Glu Glu
95 100 105

Gly Ser Leu Lys Leu Glu Asp Leu Pro Thr Val Glu Ala Pro Gly
110 115 120

Asp Pro Gln Glu Pro Gln Asn Asn Ala His Arg Asp Lys Glu Gly
125 130 135

Asp Asp Gln Ser His Trp Arg Tyr Gly Gly Asp Pro Pro Trp Pro
140 145 150

Arg Val Ser Pro Ala Cys Ala Gly Arg Phe Gln Ser Pro Val Asp
155 160 165

Ile Arg Pro Gln Leu Ala Ala Phe Cys Pro Ala Leu Arg Pro Leu
170 175 180

Glu Leu Leu Gly Phe Gln Leu Pro Pro Leu Pro Glu Leu Arg Leu
185 190 195

Arg Asn Asn Gly His Ser Val Gln Leu Thr Leu Pro Pro Gly Leu
200 205 210

Glu Met Ala Leu Gly Pro Gly Arg Glu Tyr Arg Ala Leu Gln Leu
215 220 225

His Leu His Trp Gly Ala Ala Gly Arg Pro Gly Ser Glu His Thr
230 235 240

Val Glu Gly His Arg Phe Pro Ala Glu Ile His Val Val His Leu
245 250 255

Ser Thr Ala Phe Ala Arg Val Asp Glu Ala Leu Gly Arg Pro Gly
260 265 270

Gly Leu Ala Val Leu Ala Ala Phe Leu Glu Glu Gly Pro Glu Glu
275 280 285

Asn Ser Ala Tyr Glu Gln Leu Leu Ser Arg Leu Glu Glu Ile Ala
290 295 300

Glu Glu Gly Ser Glu Thr Gln Val Pro Gly Leu Asp Ile Ser Ala
305 310 315

Leu Leu Pro Ser Asp Phe Ser Arg Tyr Phe Arg Tyr Glu Gly Ser
320 325 330

Leu Thr Thr Pro Pro Cys Ala Gln Gly Val Ile Trp Thr Val Phe
335 340 345

Asn Gln Thr Val Met Leu Ser Ala Lys Gln Leu His Thr Leu Ser
350 355 360

Asp Thr Leu Trp Gly Pro Gly Asp Ser Arg Leu Gln Leu Asn Phe
365 370 375

Arg Ala Thr Gln Pro Leu Asn Gly Arg Val Ile Glu Ala Ser Phe
380 385 390

Pro Ala Gly Val Asp Ser Ser Pro Arg Ala Ala Glu Pro Val Gln
395 400 405

Leu Asn Ser Cys Leu Ala Ala Gly Asp Ile Leu Ala Leu Val Phe
410 415 420

Gly Leu Leu Phe Ala Val Thr Ser Val Ala Phe Leu Val Gln Met
425 430 435

Arg Arg Gln His Arg Arg Gly Thr Lys Gly Gly Val Ser Tyr Arg
440 445 450

Pro Ala Glu Val Ala Glu Thr Gly Ala
455

<210> 93
<211> 375
<212> PRT
<213> Homo sapiens

<400> 93
Met Asp Ala Leu Gln Leu Ala Asn Ser Ala Phe Ala Val Asp Leu
1 5 10 15

Phe Lys Gln Leu Cys Glu Lys Glu Pro Leu Gly Asn Val Leu Phe
20 25 30

Ser Pro Ile Cys Leu Ser Thr Ser Leu Ser Leu Ala Gln Val Gly
35 40 45

Ala Lys Gly Asp Thr Ala Asn Glu Ile Gly Gln Val Leu His Phe
50 55 60

Glu Asn Val Lys Asp Val Pro Phe Gly Phe Gln Thr Val Thr Ser
65 70 75

Asp Val Asn Lys Leu Ser Ser Phe Tyr Ser Leu Lys Leu Ile Lys
80 85 90

Arg Leu Tyr Val Asp Lys Ser Leu Asn Leu Ser Thr Glu Phe Ile
95 100 105

Ser Ser Thr Lys Arg Pro Tyr Ala Lys Glu Leu Glu Thr Val Asp
110 115 120

Phe Lys Asp Lys Leu Glu Glu Thr Lys Gly Gln Ile Asn Asn Ser
125 130 135

Ile Lys Asp Leu Thr Asp Gly His Phe Glu Asn Ile Leu Ala Asp
140 145 150

Asn Ser Val Asn Asp Gln Thr Lys Ile Leu Val Val Asn Ala Ala

155 160 165

Tyr Phe Val Gly Lys Trp Met Lys Lys Phe Ser Glu Ser Glu Thr
170 175 180

Lys Glu Cys Pro Phe Arg Val Asn Lys Thr Asp Thr Lys Pro Val
185 190 195

Gln Met Met Asn Met Glu Ala Thr Phe Cys Met Gly Asn Ile Asp
200 205 210

Ser Ile Asn Cys Lys Ile Ile Glu Leu Pro Phe Gln Asn Lys His
215 220 225

Leu Ser Met Phe Ile Leu Leu Pro Lys Asp Val Glu Asp Glu Ser
230 235 240

Thr Gly Leu Glu Lys Ile Glu Lys Gln Leu Asn Ser Glu Ser Leu
245 250 255

Ser Gln Trp Thr Asn Pro Ser Thr Met Ala Asn Ala Lys Val Lys
260 265 270

Leu Ser Ile Pro Lys Phe Lys Val Glu Lys Met Ile Asp Pro Lys
275 280 285

Ala Cys Leu Glu Asn Leu Gly Leu Lys His Ile Phe Ser Glu Asp
290 295 300

Thr Ser Asp Phe Ser Gly Met Ser Glu Thr Lys Gly Val Ala Leu
305 310 315

Ser Asn Val Ile His Lys Val Cys Leu Glu Ile Thr Glu Asp Gly
320 325 330

Gly Asp Ser Ile Glu Val Pro Gly Ala Arg Ile Leu Gln His Lys
335 340 345

Asp Glu Leu Asn Ala Asp His Pro Phe Ile Tyr Ile Ile Arg His
350 355 360

Asn Lys Thr Arg Asn Ile Ile Phe Phe Gly Lys Phe Cys Ser Pro
365 370 375

<210> 94
<211> 260
<212> PRT
<213> Homo sapiens

<400> 94
Met Asp Arg His Ser Ser Tyr Ile Phe Ile Trp Leu Gln Leu Glu
1 5 10 15

Leu Cys Ala Met Ala Val Leu Leu Thr Lys Gly Glu Ile Arg Cys
20 25 30

Tyr Cys Asp Ala Ala His Cys Val Ala Thr Gly Tyr Met Cys Lys
35 40 45

Ser Glu Leu Ser Ala Cys Phe Ser Arg Leu Leu Asp Pro Gln Asn
50 55 60

Ser Asn Ser Pro Leu Thr His Gly Cys Leu Asp Ser Leu Ala Ser
65 70 75

Thr Thr Asp Ile Cys Gln Ala Lys Gln Ala Arg Asn His Ser Gly
80 85 90

Thr Thr Ile Pro Thr Leu Glu Cys Cys His Glu Asp Met Cys Asn
95 100 105

Tyr Arg Gly Leu His Asp Val Leu Ser Pro Pro Arg Gly Glu Ala
110 115 120

Ser Gly Gln Gly Asn Arg Tyr Gln His Asp Gly Ser Arg Asn Leu
125 130 135

1204
1208

Ile Thr Lys Val Gln Glu Leu Thr Ser Ser Lys Glu Leu Trp Phe
140 145 150

Arg Ala Ala Val Ile Ala Val Pro Ile Ala Gly Gly Leu Ile Leu
155 160 165

Val Leu Leu Ile Met Leu Ala Leu Arg Met Leu Arg Ser Glu Asn
170 175 180

Lys Arg Leu Gln Asp Gln Arg Gln Gln Met Leu Ser Arg Leu His
185 190 195

Tyr Ser Phe His Gly His His Ser Lys Lys Gly Gln Val Ala Lys
200 205 210

Leu Asp Leu Glu Cys Met Val Pro Val Ser Gly His Glu Asn Cys
215 220 225

Cys Leu Thr Cys Asp Lys Met Arg Gln Ala Asp Leu Ser Asn Asp
230 235 240

Lys Ile Leu Ser Leu Val His Trp Gly Met Tyr Ser Gly His Gly
245 250 255

Lys Leu Glu Phe Val
260

<210> 95
<211> 1095
<212> PRT
<213> Homo sapiens

<400> 95
Met Arg Asn Arg Arg Asn Asp Thr Leu Asp Ser Thr Arg Thr Leu
1 5 10 15

Tyr Ser Ser Ala Ser Arg Ser Thr Asp Leu Ser Tyr Ser Glu Ser
20 25 30

Asp Leu Val Asn Phe Ile Gln Ala Asn Phe Lys Lys Arg Glu Cys
35 40 45

Val Phe Phe Thr Lys Asp Ser Lys Ala Thr Glu Asn Val Cys Lys
50 55 60

Cys Gly Tyr Ala Gln Ser Gln His Met Glu Gly Thr Gln Ile Asn
65 70 75

Gln Ser Glu Lys Trp Asn Tyr Lys Lys His Thr Lys Glu Phe Pro
80 85 90

Thr Asp Ala Phe Gly Asp Ile Gln Phe Glu Thr Leu Gly Lys Lys
95 100 105

Gly Lys Tyr Ile Arg Leu Ser Cys Asp Thr Asp Ala Glu Ile Leu
110 115 120

Tyr Glu Leu Leu Thr Gln His Trp His Leu Lys Thr Pro Asn Leu
125 130 135

Val Ile Ser Val Thr Gly Gly Ala Lys Asn Phe Ala Leu Lys Pro
140 145 150

Arg Met Arg Lys Ile Phe Ser Arg Leu Ile Tyr Ile Ala Gln Ser
155 160 165

Lys Gly Ala Trp Ile Leu Thr Gly Gly Thr His Tyr Gly Leu Thr
170 175 180

Lys Tyr Ile Gly Glu Val Val Arg Asp Asn Thr Ile Ser Arg Ser
185 190 195

Ser Glu Glu Asn Ile Val Ala Ile Gly Ile Ala Ala Trp Gly Met
200 205 210

B40
10209

```
Val Ser Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu Gly
                215                 220                 225

Tyr Phe Leu Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro
                230                 235                 240

Leu Tyr Ile Leu Asp Asn Asn His Thr His Leu Leu Leu Val Asp
                245                 250                 255

Asn Gly Cys His Gly His Pro Thr Val Glu Ala Lys Leu Arg Asn
                260                 265                 270

Gln Leu Glu Lys His Ile Ser Glu Arg Thr Ile Gln Asp Ser Asn
                275                 280                 285

Tyr Gly Gly Lys Ile Pro Ile Val Cys Phe Ala Gln Gly Gly Gly
                290                 295                 300

Lys Glu Thr Leu Lys Ala Ile Asn Thr Ser Ile Lys Asn Lys Ile
                305                 310                 315

Pro Cys Val Val Val Glu Gly Ser Gly Arg Ile Ala Asp Val Ile
                320                 325                 330

Ala Ser Leu Val Glu Val Glu Asp Ala Pro Thr Ser Ser Ala Val
                335                 340                 345

Lys Glu Lys Leu Val Arg Phe Leu Pro Arg Thr Val Ser Arg Leu
                350                 355                 360

Ser Glu Glu Glu Thr Glu Ser Trp Ile Lys Trp Leu Lys Glu Ile
                365                 370                 375

Leu Glu Cys Ser His Leu Leu Thr Val Ile Lys Met Glu Glu Ala
                380                 385                 390

Gly Asp Glu Ile Val Ser Asn Ala Ile Ser Tyr Ala Leu Tyr Lys
                395                 400                 405

Ala Phe Ser Thr Ser Glu Gln Asp Lys Asp Asn Trp Asn Gly Gln
                410                 415                 420

Leu Lys Leu Leu Leu Glu Trp Asn Gln Leu Asp Leu Ala Asn Asp
                425                 430                 435

Glu Ile Phe Thr Asn Asp Arg Arg Trp Glu Ser Ala Asp Leu Gln
                440                 445                 450

Glu Val Met Phe Thr Ala Leu Ile Lys Asp Arg Pro Lys Phe Val
                455                 460                 465

Arg Leu Phe Leu Glu Asn Gly Leu Asn Leu Arg Lys Phe Leu Thr
                470                 475                 480

His Asp Val Leu Thr Glu Leu Phe Ser Asn His Phe Ser Thr Leu
                485                 490                 495

Val Tyr Arg Asn Leu Gln Ile Ala Lys Asn Ser Tyr Asn Asp Ala
                500                 505                 510

Leu Leu Thr Phe Val Trp Lys Leu Val Ala Asn Phe Arg Arg Gly
                515                 520                 525

Phe Arg Lys Glu Asp Arg Asn Gly Arg Asp Glu Met Asp Ile Glu
                530                 535                 540

Leu His Asp Val Ser Pro Ile Thr Arg His Pro Leu Gln Ala Leu
                545                 550                 555

Phe Ile Trp Ala Ile Leu Gln Asn Lys Lys Glu Leu Ser Lys Val
                560                 565                 570

Ile Trp Glu Gln Thr Arg Gly Cys Thr Leu Ala Ala Leu Gly Ala
                575                 580                 585

Ser Lys Leu Leu Lys Thr Leu Ala Lys Val Lys Asn Asp Ile Asn
```

590 595 600

Ala Ala Gly Glu Ser Glu Glu Leu Ala Asn Glu Tyr Glu Thr Arg
605 610 615

Ala Val Glu Leu Phe Thr Glu Cys Tyr Ser Ser Asp Glu Asp Leu
620 625 630

Ala Glu Gln Leu Leu Val Tyr Ser Cys Glu Ala Trp Gly Gly Ser
635 640 645

Asn Cys Leu Glu Leu Ala Val Glu Ala Thr Asp Gln His Phe Thr
650 655 660

Ala Gln Pro Gly Val Gln Asn Phe Leu Ser Lys Gln Trp Tyr Gly
665 670 675

Glu Ile Ser Arg Asp Thr Lys Asn Trp Lys Ile Ile Leu Cys Leu
680 685 690

Phe Ile Ile Pro Leu Val Gly Cys Gly Phe Val Ser Phe Arg Lys
695 700 705

Lys Pro Val Asp Lys His Lys Lys Leu Leu Trp Tyr Tyr Val Ala
710 715 720

Phe Phe Thr Ser Pro Phe Val Val Phe Ser Trp Asn Val Val Phe
725 730 735

Tyr Ile Ala Phe Leu Leu Leu Phe Ala Tyr Val Leu Leu Met Asp
740 745 750

Phe His Ser Val Pro His Pro Pro Glu Leu Val Leu Tyr Ser Leu
755 760 765

Val Phe Val Leu Phe Cys Asp Glu Val Arg Gln Trp Tyr Val Asn
770 775 780

Gly Val Asn Tyr Phe Thr Asp Leu Trp Asn Val Met Asp Thr Leu
785 790 795

Gly Leu Phe Tyr Phe Ile Ala Gly Ile Val Phe Arg Leu His Ser
800 805 810

Ser Asn Lys Ser Ser Leu Tyr Ser Gly Arg Val Ile Phe Cys Leu
815 820 825

Asp Tyr Ile Ile Phe Thr Leu Arg Leu Ile His Ile Phe Thr Val
830 835 840

Ser Arg Asn Leu Gly Pro Lys Ile Ile Met Leu Gln Arg Met Leu
845 850 855

Ile Asp Val Phe Phe Phe Leu Phe Leu Phe Ala Val Trp Met Val
860 865 870

Ala Phe Gly Val Ala Arg Gln Gly Ile Leu Arg Gln Asn Glu Gln
875 880 885

Arg Trp Arg Trp Ile Phe Arg Ser Val Ile Tyr Glu Pro Tyr Leu
890 895 900

Ala Met Phe Gly Gln Val Pro Ser Asp Val Asp Gly Thr Thr Tyr
905 910 915

Asp Phe Ala His Cys Thr Phe Thr Gly Asn Glu Ser Lys Pro Leu
920 925 930

Cys Val Glu Leu Asp Glu His Asn Leu Pro Arg Phe Pro Glu Trp
935 940 945

Ile Thr Ile Pro Leu Val Cys Ile Tyr Met Leu Ser Thr Asn Ile
950 955 960

Leu Leu Val Asn Leu Leu Val Ala Met Phe Gly Tyr Thr Val Gly
965 970 975

Thr Val Gln Glu Asn Asn Asp Gln Val Trp Lys Phe Gln Arg Tyr
980 985 990

Phe Leu Val Gln Glu Tyr Cys Ser Arg Leu Asn Ile Pro Phe Pro
995 1000 1005

Phe Ile Val Phe Ala Tyr Phe Tyr Met Val Val Lys Lys Cys Phe
1010 1015 1020

Lys Cys Cys Cys Lys Glu Lys Asn Met Glu Ser Ser Val Cys Cys
1025 1030 1035

Phe Lys Asn Glu Asp Asn Glu Thr Leu Ala Trp Glu Gly Val Met
1040 1045 1050

Lys Glu Asn Tyr Leu Val Lys Ile Asn Thr Lys Ala Asn Asp Thr
1055 1060 1065

Ser Glu Glu Met Arg His Arg Phe Arg Gln Leu Asp Thr Lys Leu
1070 1075 1080

Asn Asp Leu Lys Gly Leu Leu Lys Glu Ile Ala Asn Lys Ile Lys
1085 1090 1095

<210> 96
<211> 1095
<212> PRT
<213> Homo sapiens

<400> 96
Met Arg Asn Arg Arg Asn Asp Thr Leu Asp Ser Thr Arg Thr Leu
1 5 10 15

Tyr Ser Ser Ala Ser Arg Ser Thr Asp Leu Ser Tyr Ser Glu Ser
20 25 30

Asp Leu Val Asn Phe Ile Gln Ala Asn Phe Lys Lys Arg Glu Cys
35 40 45

Val Phe Phe Thr Lys Asp Ser Lys Ala Thr Glu Asn Val Cys Lys
50 55 60

Cys Gly Tyr Ala Gln Ser Gln His Met Glu Gly Thr Gln Ile Asn
65 70 75

Gln Ser Glu Lys Trp Asn Tyr Lys Lys His Thr Lys Glu Phe Pro
80 85 90

Thr Asp Ala Phe Gly Asp Ile Gln Phe Glu Thr Leu Gly Lys Lys
95 100 105

Gly Lys Tyr Ile Arg Leu Ser Cys Asp Thr Asp Ala Glu Ile Leu
110 115 120

Tyr Glu Leu Leu Thr Gln His Trp His Leu Lys Thr Pro Asn Leu
125 130 135

Val Ile Ser Val Thr Gly Gly Ala Lys Asn Phe Ala Leu Lys Pro
140 145 150

Arg Met Arg Lys Ile Phe Ser Arg Leu Ile Tyr Ile Ala Gln Ser
155 160 165

Lys Gly Ala Trp Ile Leu Thr Gly Gly Thr His Tyr Gly Leu Met
170 175 180

Lys Tyr Ile Gly Glu Val Val Arg Asp Asn Thr Ile Ser Arg Ser
185 190 195

Ser Glu Glu Asn Ile Val Ala Ile Gly Ile Ala Ala Trp Gly Met
200 205 210

Val Ser Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu Gly
215 220 225

Tyr Phe Leu Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro
230 235 240

Leu Tyr Ile Leu Asp Asn Asn His Thr His Leu Leu Leu Val Asp
245 250 255

Asn Gly Cys His Gly His Pro Thr Val Glu Ala Lys Leu Arg Asn
260 265 270

Gln Leu Glu Lys Tyr Ile Ser Glu Arg Thr Ile Gln Asp Ser Asn
275 280 285

Tyr Gly Gly Lys Ile Pro Ile Val Cys Phe Ala Gln Gly Gly Gly
290 295 300

Lys Glu Thr Leu Lys Ala Ile Asn Thr Ser Ile Lys Asn Lys Ile
305 310 315

Pro Cys Val Val Val Glu Gly Ser Gly Gln Ile Ala Asp Val Ile
320 325 330

Ala Ser Leu Val Glu Val Glu Asp Ala Leu Thr Ser Ser Ala Val
335 340 345

Lys Glu Lys Leu Val Arg Phe Leu Pro Arg Thr Val Ser Arg Leu
350 355 360

Pro Glu Glu Glu Thr Glu Ser Trp Ile Lys Trp Leu Lys Glu Ile
365 370 375

Leu Glu Cys Ser His Leu Leu Thr Val Ile Lys Met Glu Glu Ala
380 385 390

Gly Asp Glu Ile Val Ser Asn Ala Ile Ser Tyr Ala Leu Tyr Lys
395 400 405

Ala Phe Ser Thr Ser Glu Gln Asp Lys Asp Asn Trp Asn Gly Gln
410 415 420

Leu Lys Leu Leu Leu Glu Trp Asn Gln Leu Asp Leu Ala Asn Asp
425 430 435

Glu Ile Phe Thr Asn Asp Arg Arg Trp Glu Ser Ala Asp Leu Gln
440 445 450

Glu Val Met Phe Thr Ala Leu Ile Lys Asp Arg Pro Lys Phe Val
455 460 465

Arg Leu Phe Leu Glu Asn Gly Leu Asn Leu Arg Lys Phe Leu Thr
470 475 480

His Asp Val Leu Thr Glu Leu Phe Ser Asn His Phe Ser Thr Leu
485 490 495

Val Tyr Arg Asn Leu Gln Ile Ala Lys Asn Ser Tyr Asn Asp Ala
500 505 510

Leu Leu Thr Phe Val Trp Lys Leu Val Ala Asn Phe Arg Arg Gly
515 520 525

Phe Arg Lys Glu Asp Arg Asn Gly Arg Asp Glu Met Asp Ile Glu
530 535 540

Leu His Asp Val Ser Pro Ile Thr Arg His Pro Leu Gln Ala Leu
545 550 555

Phe Ile Trp Ala Ile Leu Gln Asn Lys Lys Glu Leu Ser Lys Val
560 565 570

Ile Trp Glu Gln Thr Arg Gly Cys Thr Leu Ala Ala Leu Gly Ala
575 580 585

Ser Lys Leu Leu Lys Thr Leu Ala Lys Val Lys Asn Asp Ile Asn
590 595 600

Ala Ala Gly Glu Ser Glu Glu Leu Ala Asn Glu Tyr Glu Thr Arg

605 610 615

Ala Val Glu Leu Phe Thr Glu Cys Tyr Ser Ser Asp Glu Asp Leu
620 625 630

Ala Glu Gln Leu Leu Val Tyr Ser Cys Glu Ala Trp Gly Gly Ser
635 640 645

Asn Cys Leu Glu Leu Ala Val Glu Ala Thr Asp Gln His Phe Ile
650 655 660

Ala Gln Pro Gly Val Gln Asn Phe Leu Ser Lys Gln Trp Tyr Gly
665 670 675

Glu Ile Ser Arg Asp Thr Lys Asn Trp Lys Ile Ile Leu Cys Leu
680 685 690

Phe Ile Ile Pro Leu Val Gly Cys Gly Phe Val Ser Phe Arg Lys
695 700 705

Lys Pro Val Asp Lys His Lys Lys Leu Leu Trp Tyr Tyr Val Ala
710 715 720

Phe Phe Thr Ser Pro Phe Val Val Phe Ser Trp Asn Val Val Phe
725 730 735

Tyr Ile Ala Phe Leu Leu Leu Phe Ala Tyr Val Leu Leu Met Asp
740 745 750

Phe His Ser Val Pro His Pro Pro Glu Leu Val Leu Tyr Ser Leu
755 760 765

Val Phe Val Leu Phe Cys Asp Glu Val Arg Gln Trp Tyr Val Asn
770 775 780

Gly Val Asn Tyr Phe Thr Asp Leu Trp Asn Val Met Asp Thr Leu
785 790 795

Gly Leu Phe Tyr Phe Ile Ala Gly Ile Val Phe Arg Leu His Ser
800 805 810

Ser Asn Lys Ser Ser Leu Tyr Ser Gly Arg Val Ile Phe Cys Leu
815 820 825

Asp Tyr Ile Ile Phe Thr Leu Arg Leu Ile His Ile Phe Thr Val
830 835 840

Ser Arg Asn Leu Gly Pro Lys Ile Ile Met Leu Gln Arg Met Leu
845 850 855

Ile Asp Val Phe Phe Phe Leu Phe Leu Phe Ala Val Trp Met Val
860 865 870

Ala Phe Gly Val Ala Arg Gln Gly Ile Leu Arg Gln Asn Glu Gln
875 880 885

Arg Trp Arg Trp Ile Phe Arg Ser Val Ile Tyr Glu Pro Tyr Leu
890 895 900

Ala Met Phe Gly Gln Val Pro Ser Asp Val Asp Gly Thr Thr Tyr
905 910 915

Asp Phe Ala His Cys Thr Phe Thr Gly Asn Glu Ser Lys Pro Leu
920 925 930

Cys Val Glu Leu Asp Glu His Asn Leu Pro Arg Phe Pro Glu Trp
935 940 945

Ile Thr Ile Pro Leu Val Cys Ile Tyr Met Leu Ser Thr Asn Ile
950 955 960

Leu Leu Val Asn Leu Leu Val Ala Met Phe Gly Tyr Thr Val Gly
965 970 975

Thr Val Gln Glu Asn Asn Asp Gln Val Trp Lys Phe Gln Arg Tyr
980 985 990

Phe Leu Val Gln Glu Tyr Cys Ser Arg Leu Asn Ile Pro Phe Pro
995 1000 1005

Phe Ile Val Phe Ala Tyr Phe Tyr Met Val Val Lys Lys Cys Phe
1010 1015 1020

Lys Cys Cys Cys Lys Glu Lys Asn Met Glu Ser Ser Val Cys Cys
1025 1030 1035

Phe Lys Asn Glu Asp Asn Glu Thr Leu Ala Trp Glu Gly Val Met
1040 1045 1050

Lys Glu Asn Tyr Leu Val Lys Ile Asn Thr Lys Ala Asn Asp Thr
1055 1060 1065

Ser Glu Glu Met Arg His Arg Phe Arg Gln Leu Asp Thr Lys Leu
1070 1075 1080

Asn Asp Leu Lys Gly Leu Leu Lys Glu Ile Ala Asn Lys Ile Lys
1085 1090 1095

<210> 97
<211> 347
<212> PRT
<213> Homo sapiens

<400> 97
Arg Arg Gly Arg Gly Glu Val Leu Ala Val Glu Arg Gly Ser Gly
1 5 10 15

Ser Gly Gly Gly Gly Thr Arg Ser Gly Trp Pro Ala Pro Ala Ala
20 25 30

Gly Ala Asp Lys Lys Pro Ser Arg Cys Gly Ser Gly Arg Glu Gly
35 40 45

Glu Gly Val Ser Glu Gly His Lys Ser Met Thr Gly Leu Tyr Glu
50 55 60

Leu Val Trp Arg Val Leu His Ala Leu Leu Cys Leu His Arg Thr
65 70 75

Leu Thr Ser Trp Leu Arg Val Arg Phe Gly Thr Trp Asn Trp Ile
80 85 90

Trp Arg Arg Cys Cys Arg Ala Ala Ser Ala Ala Val Leu Ala Pro
95 100 105

Leu Gly Phe Thr Leu Arg Lys Pro Pro Ala Val Gly Arg Asn Arg
110 115 120

Arg His His Arg His Pro Arg Gly Gly Ser Cys Leu Ala Ala Ala
125 130 135

His His Arg Met Arg Trp Arg Ala Asp Gly Arg Ser Leu Glu Lys
140 145 150

Leu Pro Val His Met Gly Leu Val Ile Thr Glu Val Glu Gln Glu
155 160 165

Pro Ser Phe Ser Asp Ile Ala Ser Leu Val Val Trp Cys Met Ala
170 175 180

Val Gly Ile Ser Tyr Ile Ser Val Tyr Asp His Gln Gly Ile Phe
185 190 195

Lys Arg Asn Asn Ser Arg Leu Met Asp Glu Ile Leu Lys Gln Gln
200 205 210

Gln Glu Leu Leu Gly Leu Asp Cys Ser Lys Tyr Ser Pro Glu Phe
215 220 225

Ala Asn Ser Asn Asp Lys Asp Asp Gln Val Leu Asn Cys His Leu
230 235 240

Ala Val Lys Val Leu Ser Pro Glu Asp Gly Lys Ala Asp Ile Val
245 250 255

Arg Ala Ala Gln Asp Phe Cys Gln Leu Val Ala Gln Lys Gln Arg
260 265 270

Arg Pro Thr Asp Leu Asp Val Asp Thr Leu Ala Ser Leu Leu Ser
275 280 285

Ser Asn Gly Cys Pro Asp Pro Asp Leu Val Leu Lys Phe Gly Pro
290 295 300

Val Asp Ser Thr Leu Gly Phe Leu Pro Trp His Ile Arg Leu Thr
305 310 315

Glu Ile Val Ser Leu Pro Ser His Leu Asn Ile Ser Tyr Glu Asp
320 325 330

Phe Phe Ser Ala Leu Arg Gln Tyr Ala Ala Cys Glu Gln Arg Leu
335 340 345

Gly Lys

<210> 98
<211> 293
<212> PRT
<213> Homo sapiens

<400> 98
Met Thr Gly Leu Tyr Glu Leu Val Trp Arg Val Leu His Ala Leu
1 5 10 15

Leu Cys Leu His Arg Thr Leu Thr Ser Trp Leu Arg Val Arg Phe
20 25 30

Gly Thr Trp Asn Trp Ile Trp Arg Arg Cys Cys Arg Ala Ala Ser
35 40 45

Ala Ala Val Leu Ala Pro Leu Gly Phe Thr Leu Arg Lys Pro Pro
50 55 60

Ala Val Gly Arg Asn Arg Arg His His Arg His Pro Arg Gly Gly
65 70 75

Ser Cys Leu Ala Ala Ala His His Arg Met Arg Trp Arg Ala Asp
80 85 90

Gly Arg Ser Leu Glu Lys Leu Pro Val His Met Gly Leu Val Ile
95 100 105

Thr Glu Val Glu Gln Glu Pro Ser Phe Ser Asp Ile Ala Ser Leu
110 115 120

Val Val Trp Cys Met Ala Val Gly Ile Ser Tyr Ile Ser Val Tyr
125 130 135

Asp His Gln Gly Ile Phe Lys Arg Asn Asn Ser Arg Leu Met Asp
140 145 150

Glu Ile Leu Lys Gln Gln Gln Glu Leu Leu Gly Leu Asp Cys Ser
155 160 165

Lys Tyr Ser Pro Glu Phe Ala Asn Ser Asn Asp Lys Asp Asp Gln
170 175 180

Val Leu Asn Cys His Leu Ala Val Lys Val Leu Ser Pro Glu Asp
185 190 195

Gly Lys Ala Asp Ile Val Arg Ala Ala Gln Asp Phe Cys Gln Leu
200 205 210

Val Ala Gln Lys Gln Lys Arg Pro Thr Asp Leu Asp Val Asp Thr
215 220 225

Leu Ala Ser Leu Leu Ser Ser Asn Gly Cys Pro Asp Pro Asp Leu

230 235 240

Val Leu Lys Phe Gly Pro Val Asp Ser Thr Leu Gly Phe Leu Pro
245 250 255

Trp His Ile Arg Leu Thr Glu Ile Val Ser Leu Pro Ser His Leu
260 265 270

Asn Ile Ser Tyr Glu Asp Phe Phe Ser Ala Leu Arg Gln Tyr Ala
275 280 285

Ala Cys Glu Gln Arg Leu Gly Lys
290

<210> 99
<211> 553
<212> PRT
<213> Homo sapiens

<400> 99
Met Val Gln Arg Leu Trp Val Ser Arg Leu Leu Arg His Arg Lys
1 5 10 15

Ala Gln Leu Leu Leu Val Asn Leu Leu Thr Phe Gly Leu Glu Val
20 25 30

Cys Leu Ala Ala Gly Ile Thr Tyr Val Pro Pro Leu Leu Leu Glu
35 40 45

Val Gly Val Glu Glu Lys Phe Met Thr Met Val Leu Gly Ile Gly
50 55 60

Pro Val Leu Gly Leu Val Cys Val Pro Leu Leu Gly Ser Ala Ser
65 70 75

Asp His Trp Arg Gly Arg Tyr Gly Arg Arg Arg Pro Phe Ile Trp
80 85 90

Ala Leu Ser Leu Gly Ile Leu Leu Ser Leu Phe Leu Ile Pro Arg
95 100 105

Ala Gly Trp Leu Ala Gly Leu Leu Cys Pro Asp Pro Arg Pro Leu
110 115 120

Glu Leu Ala Leu Leu Ile Leu Gly Val Gly Leu Leu Asp Phe Cys
125 130 135

Gly Gln Val Cys Phe Thr Pro Leu Glu Ala Leu Leu Ser Asp Leu
140 145 150

Phe Arg Asp Pro Asp His Cys Arg Gln Ala Tyr Ser Val Tyr Ala
155 160 165

Phe Met Ile Ser Leu Gly Gly Cys Leu Gly Tyr Leu Leu Pro Ala
170 175 180

Ile Asp Trp Asp Thr Ser Ala Leu Ala Pro Tyr Leu Gly Thr Gln
185 190 195

Glu Glu Cys Leu Phe Gly Leu Leu Thr Leu Ile Phe Leu Thr Cys
200 205 210

Val Ala Ala Thr Leu Leu Val Ala Glu Glu Ala Ala Leu Gly Pro
215 220 225

Thr Glu Pro Ala Glu Gly Leu Ser Ala Pro Ser Leu Ser Pro His
230 235 240

Cys Cys Pro Cys Arg Ala Arg Leu Ala Phe Arg Asn Leu Gly Ala
245 250 255

Leu Leu Pro Arg Leu His Gln Leu Cys Cys Arg Met Pro Arg Thr
260 265 270

Leu Arg Arg Leu Phe Val Ala Glu Leu Cys Ser Trp Met Ala Leu
275 280 285

Met Thr Phe Thr Leu Phe Tyr Thr Asp Phe Val Gly Glu Gly Leu
290 295 300

Tyr Gln Gly Val Pro Arg Ala Glu Pro Gly Thr Glu Ala Arg Arg
305 310 315

His Tyr Asp Glu Gly Val Arg Met Gly Ser Leu Gly Leu Phe Leu
320 325 330

Gln Cys Ala Ile Ser Leu Val Phe Ser Leu Val Met Asp Arg Leu
335 340 345

Val Gln Arg Phe Gly Thr Arg Ala Val Tyr Leu Ala Ser Val Ala
350 355 360

Ala Phe Pro Val Ala Ala Gly Ala Thr Cys Leu Ser His Ser Val
365 370 375

Ala Val Val Thr Ala Ser Ala Ala Leu Thr Gly Phe Thr Phe Ser
380 385 390

Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala Ser Leu Tyr His Arg
395 400 405

Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly Asp Thr Gly Gly
410 415 420

Ala Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu Pro Gly Pro
425 430 435

Lys Pro Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala Gly Gly
440 445 450

Ser Gly Leu Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser Ala
455 460 465

Cys Asp Val Ser Val Arg Val Val Val Gly Glu Pro Thr Glu Ala
470 475 480

Arg Val Val Pro Gly Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu
485 490 495

Asp Ser Ala Phe Leu Leu Ser Gln Val Ala Pro Ser Leu Phe Met
500 505 510

Gly Ser Ile Val Gln Leu Ser Gln Ser Val Thr Ala Tyr Met Val
515 520 525

Ser Ala Ala Gly Leu Gly Leu Val Ala Ile Tyr Phe Ala Thr Gln
530 535 540

Val Val Phe Asp Lys Ser Asp Leu Ala Lys Tyr Ser Ala
545 550

<210> 100
<211> 987
<212> PRT
<213> Homo sapiens

<400> 100
Met Ala Leu Arg Arg Leu Gly Ala Ala Leu Leu Leu Leu Pro Leu
1 5 10 15

Leu Ala Ala Val Glu Glu Thr Leu Met Asp Ser Thr Thr Ala Thr
20 25 30

Ala Glu Leu Gly Trp Met Val His Pro Pro Ser Gly Trp Glu Glu
35 40 45

Val Ser Gly Tyr Asp Glu Asn Met Asn Thr Ile Arg Thr Tyr Gln
50 55 60

Val Cys Asn Val Phe Glu Ser Ser Gln Asn Asn Trp Leu Arg Thr
65 70 75

Lys Phe Ile Arg Arg Arg Gly Ala His Arg Ile His Val Glu Met
80 85 90

Lys Phe Ser Val Arg Asp Cys Ser Ser Ile Pro Ser Val Pro Gly
95 100 105

Ser Cys Lys Glu Thr Phe Asn Leu Tyr Tyr Tyr Glu Ala Asp Phe
110 115 120

Asp Ser Ala Thr Lys Thr Phe Pro Asn Trp Met Glu Asn Pro Trp
125 130 135

Val Lys Val Asp Thr Ile Ala Ala Asp Glu Ser Phe Ser Gln Val
140 145 150

Asp Leu Gly Gly Arg Val Met Lys Ile Asn Thr Glu Val Arg Ser
155 160 165

Phe Gly Pro Val Ser Arg Ser Gly Phe Tyr Leu Ala Phe Gln Asp
170 175 180

Tyr Gly Gly Cys Met Ser Leu Ile Ala Val Arg Val Phe Tyr Arg
185 190 195

Lys Cys Pro Arg Ile Ile Gln Asn Gly Ala Ile Phe Gln Glu Thr
200 205 210

Leu Ser Gly Ala Glu Ser Thr Ser Leu Val Ala Ala Arg Gly Ser
215 220 225

Cys Ile Ala Asn Ala Glu Glu Val Asp Val Pro Ile Lys Leu Tyr
230 235 240

Cys Asn Gly Asp Gly Glu Trp Leu Val Pro Ile Gly Arg Cys Met
245 250 255

Cys Lys Ala Gly Phe Glu Ala Val Glu Asn Gly Thr Val Cys Arg
260 265 270

Gly Cys Pro Ser Gly Thr Phe Lys Ala Asn Gln Gly Asp Glu Ala
275 280 285

Cys Thr His Cys Pro Ile Asn Ser Arg Thr Thr Ser Glu Gly Ala
290 295 300

Thr Asn Cys Val Cys Arg Asn Gly Tyr Tyr Arg Ala Asp Leu Asp
305 310 315

Pro Leu Asp Met Pro Cys Thr Thr Ile Pro Ser Ala Pro Gln Ala
320 325 330

Val Ile Ser Ser Val Asn Glu Thr Ser Leu Met Leu Glu Trp Thr
335 340 345

Pro Pro Arg Asp Ser Gly Gly Arg Glu Asp Leu Val Tyr Asn Ile
350 355 360

Ile Cys Lys Ser Cys Gly Ser Gly Arg Gly Ala Cys Thr Arg Cys
365 370 375

Gly Asp Asn Val Gln Tyr Ala Pro Arg Gln Leu Gly Leu Thr Glu
380 385 390

Pro Arg Ile Tyr Ile Ser Asp Leu Leu Ala His Thr Gln Tyr Thr
395 400 405

Phe Glu Ile Gln Ala Val Asn Gly Val Thr Asp Gln Ser Pro Phe
410 415 420

Ser Pro Gln Phe Ala Ser Val Asn Ile Thr Thr Asn Gln Ala Ala
425 430 435

Pro Ser Ala Val Ser Ile Met His Gln Val Ser Arg Thr Val Asp
440 445 450

Ser Ile Thr Leu Ser Trp Ser Gln Pro Asp Gln Pro Asn Gly Val

455 460 465

Ile Leu Asp Tyr Glu Leu Gln Tyr Tyr Glu Lys Glu Leu Ser Glu
470 475 480

Tyr Asn Ala Thr Ala Ile Lys Ser Pro Thr Asn Thr Val Thr Val
485 490 495

Gln Gly Leu Lys Ala Gly Ala Ile Tyr Val Phe Gln Val Arg Ala
500 505 510

Arg Thr Val Ala Gly Tyr Gly Arg Tyr Ser Gly Lys Met Tyr Phe
515 520 525

Gln Thr Met Thr Glu Ala Glu Tyr Gln Thr Ser Ile Gln Glu Lys
530 535 540

Leu Pro Leu Ile Ile Gly Ser Ser Ala Ala Gly Leu Val Phe Leu
545 550 555

Ile Ala Val Val Val Ile Ala Ile Val Cys Asn Arg Arg Arg Gly
560 565 570

Phe Glu Arg Ala Asp Ser Glu Tyr Thr Asp Lys Leu Gln His Tyr
575 580 585

Thr Ser Gly His Met Thr Pro Gly Met Lys Ile Tyr Ile Asp Pro
590 595 600

Phe Thr Tyr Glu Asp Pro Asn Glu Ala Val Arg Glu Phe Ala Lys
605 610 615

Glu Ile Asp Ile Ser Cys Val Lys Ile Glu Gln Val Ile Gly Ala
620 625 630

Gly Glu Phe Gly Glu Val Cys Ser Gly His Leu Lys Leu Pro Gly
635 640 645

Lys Arg Glu Ile Phe Val Ala Ile Lys Thr Leu Lys Ser Gly Tyr
650 655 660

Thr Glu Lys Gln Arg Arg Asp Phe Leu Ser Glu Ala Ser Ile Met
665 670 675

Gly Gln Phe Asp His Pro Asn Val Ile His Leu Glu Gly Val Val
680 685 690

Thr Lys Ser Thr Pro Val Met Ile Ile Thr Glu Phe Met Glu Asn
695 700 705

Gly Ser Leu Asp Ser Phe Leu Arg Gln Asn Asp Gly Gln Phe Thr
710 715 720

Val Ile Gln Leu Val Gly Met Leu Arg Gly Ile Ala Ala Gly Met
725 730 735

Lys Tyr Leu Ala Asp Met Asn Tyr Val His Arg Asp Leu Ala Ala
740 745 750

Arg Asn Ile Leu Val Asn Ser Asn Leu Val Cys Lys Val Ser Asp
755 760 765

Phe Gly Leu Ser Arg Phe Leu Glu Asp Asp Thr Ser Asp Pro Thr
770 775 780

Tyr Thr Ser Ala Leu Gly Gly Lys Ile Pro Ile Arg Trp Thr Ala
785 790 795

Pro Glu Ala Ile Gln Tyr Arg Lys Phe Thr Ser Ala Ser Asp Val
800 805 810

Trp Ser Tyr Gly Ile Val Met Trp Glu Val Met Ser Tyr Gly Glu
815 820 825

Arg Pro Tyr Trp Asp Met Thr Asn Gln Asp Val Ile Asn Ala Ile
830 835 840

1220

```
Glu Gln Asp Tyr Arg Leu Pro Pro Pro Met Asp Cys Pro Ser Ala
                845                 850                 855

Leu His Gln Leu Met Leu Asp Cys Trp Gln Lys Asp Arg Asn His
                860                 865                 870

Arg Pro Lys Phe Gly Gln Ile Val Asn Thr Leu Asp Lys Met Ile
                875                 880                 885

Arg Asn Pro Asn Ser Leu Lys Ala Met Ala Pro Leu Ser Ser Gly
                890                 895                 900

Ile Asn Leu Pro Leu Leu Asp Arg Thr Ile Pro Asp Tyr Thr Ser
                905                 910                 915

Phe Asn Thr Val Asp Glu Trp Leu Glu Ala Ile Lys Met Gly Gln
                920                 925                 930

Tyr Lys Glu Ser Phe Ala Asn Ala Gly Phe Thr Ser Phe Asp Val
                935                 940                 945

Val Ser Gln Met Met Met Glu Asp Ile Leu Arg Val Gly Leu Thr
                950                 955                 960

Leu Ala Gly His Gln Lys Lys Ile Leu Asn Ser Ile Gln Val Met
                965                 970                 975

Arg Ala Gln Met Asn Gln Ile Gln Ser Val Glu Val
                980                 985
```

<210> 101
<211> 986
<212> PRT
<213> Homo sapiens

<400> 101

```
Met Ala Leu Arg Arg Leu Gly Ala Ala Leu Leu Leu Leu Pro Leu
  1               5                  10                  15

Leu Ala Ala Val Glu Glu Thr Leu Met Asp Ser Thr Thr Ala Thr
                20                  25                  30

Ala Glu Leu Gly Trp Met Val His Pro Pro Ser Gly Trp Glu Glu
                35                  40                  45

Val Ser Gly Tyr Asp Glu Asn Met Asn Thr Ile Arg Thr Tyr Gln
                50                  55                  60

Val Cys Asn Val Phe Glu Ser Ser Gln Asn Asn Trp Leu Arg Thr
                65                  70                  75

Lys Phe Ile Arg Arg Arg Gly Ala His Arg Ile His Val Glu Met
                80                  85                  90

Lys Phe Ser Val Arg Asp Cys Ser Ser Ile Pro Ser Val Pro Gly
                95                 100                 105

Ser Cys Lys Glu Thr Phe Asn Leu Tyr Tyr Tyr Glu Ala Asp Phe
               110                 115                 120

Asp Ser Ala Thr Lys Thr Phe Pro Asn Trp Met Glu Asn Pro Trp
               125                 130                 135

Val Lys Val Asp Thr Ile Ala Ala Asp Glu Ser Phe Ser Gln Val
               140                 145                 150

Asp Leu Gly Gly Arg Val Met Lys Ile Asn Thr Glu Val Arg Ser
               155                 160                 165

Phe Gly Pro Val Ser Arg Ser Gly Phe Tyr Leu Ala Phe Gln Asp
               170                 175                 180

Tyr Gly Gly Cys Met Ser Leu Ile Ala Val Arg Val Phe Tyr Arg
               185                 190                 195
```

Lys Cys Pro Arg Ile Ile Gln Asn Gly Ala Ile Phe Gln Glu Thr
200 205 210

Leu Ser Gly Ala Glu Ser Thr Ser Leu Val Ala Ala Arg Gly Ser
215 220 225

Cys Ile Ala Asn Ala Glu Glu Val Asp Val Pro Ile Lys Leu Tyr
230 235 240

Cys Asn Gly Asp Gly Glu Trp Leu Val Pro Ile Gly Arg Cys Met
245 250 255

Cys Lys Ala Gly Phe Glu Ala Val Glu Asn Gly Thr Val Cys Arg
260 265 270

Gly Cys Pro Ser Gly Thr Phe Lys Ala Asn Gln Gly Asp Glu Ala
275 280 285

Cys Thr His Cys Pro Ile Asn Ser Arg Thr Thr Ser Glu Gly Ala
290 295 300

Thr Asn Cys Val Cys Arg Asn Gly Tyr Tyr Arg Ala Asp Leu Asp
305 310 315

Pro Leu Asp Met Pro Cys Thr Thr Ile Pro Ser Ala Pro Gln Ala
320 325 330

Val Ile Ser Ser Val Asn Glu Thr Ser Leu Met Leu Glu Trp Thr
335 340 345

Pro Pro Arg Asp Ser Gly Gly Arg Glu Asp Leu Val Tyr Asn Ile
350 355 360

Ile Cys Lys Ser Cys Gly Ser Gly Arg Gly Ala Cys Thr Arg Cys
365 370 375

Gly Asp Asn Val Gln Tyr Ala Pro Arg Gln Leu Gly Leu Thr Glu
380 385 390

Pro Arg Ile Tyr Ile Ser Asp Leu Leu Ala His Thr Gln Tyr Thr
395 400 405

Phe Glu Ile Gln Ala Val Asn Gly Val Thr Asp Gln Ser Pro Phe
410 415 420

Ser Pro Gln Phe Ala Ser Val Asn Ile Thr Thr Asn Gln Ala Ala
425 430 435

Pro Ser Ala Val Ser Ile Met His Gln Val Ser Arg Thr Val Asp
440 445 450

Ser Ile Thr Leu Ser Trp Ser Gln Pro Asp Gln Pro Asn Gly Val
455 460 465

Ile Leu Asp Tyr Glu Leu Gln Tyr Tyr Glu Lys Glu Leu Ser Glu
470 475 480

Tyr Asn Ala Thr Ala Ile Lys Ser Pro Thr Asn Thr Val Thr Val
485 490 495

Gln Gly Leu Lys Ala Gly Ala Ile Tyr Val Phe Gln Val Arg Ala
500 505 510

Arg Thr Val Ala Gly Tyr Gly Arg Tyr Ser Gly Lys Met Tyr Phe
515 520 525

Gln Thr Met Thr Glu Ala Glu Tyr Gln Thr Ser Ile Gln Glu Lys
530 535 540

Leu Pro Leu Ile Ile Gly Ser Ser Ala Ala Gly Leu Val Phe Leu
545 550 555

Ile Ala Val Val Val Ile Ala Ile Val Cys Asn Arg Arg Gly Phe
560 565 570

Glu Arg Ala Asp Ser Glu Tyr Thr Asp Lys Leu Gln His Tyr Thr

1224
1222

575 580 585

Ser Gly His Met Thr Pro Gly Met Lys Ile Tyr Ile Asp Pro Phe
590 595 600

Thr Tyr Glu Asp Pro Asn Glu Ala Val Arg Glu Phe Ala Lys Glu
605 610 615

Ile Asp Ile Ser Cys Val Lys Ile Glu Gln Val Ile Gly Ala Gly
620 625 630

Glu Phe Gly Glu Val Cys Ser Gly His Leu Lys Leu Pro Gly Lys
635 640 645

Arg Glu Ile Phe Val Ala Ile Lys Thr Leu Lys Ser Gly Tyr Thr
650 655 660

Glu Lys Gln Arg Arg Asp Phe Leu Ser Glu Ala Ser Ile Met Gly
665 670 675

Gln Phe Asp His Pro Asn Val Ile His Leu Glu Gly Val Val Thr
680 685 690

Lys Ser Thr Pro Val Met Ile Ile Thr Glu Phe Met Glu Asn Gly
695 700 705

Ser Leu Asp Ser Phe Leu Arg Gln Asn Asp Gly Gln Phe Thr Val
710 715 720

Ile Gln Leu Val Gly Met Leu Arg Gly Ile Ala Ala Gly Met Lys
725 730 735

Tyr Leu Ala Asp Met Asn Tyr Val His Arg Asp Leu Ala Ala Arg
740 745 750

Asn Ile Leu Val Asn Ser Asn Leu Val Cys Lys Val Ser Asp Phe
755 760 765

Gly Leu Ser Arg Phe Leu Glu Asp Asp Thr Ser Asp Pro Thr Tyr
770 775 780

Thr Ser Ala Leu Gly Gly Lys Ile Pro Ile Arg Trp Thr Ala Pro
785 790 795

Glu Ala Ile Gln Tyr Arg Lys Phe Thr Ser Ala Ser Asp Val Trp
800 805 810

Ser Tyr Gly Ile Val Met Trp Glu Val Met Ser Tyr Gly Glu Arg
815 820 825

Pro Tyr Trp Asp Met Thr Asn Gln Asp Val Ile Asn Ala Ile Glu
830 835 840

Gln Asp Tyr Arg Leu Pro Pro Pro Met Asp Cys Pro Ser Ala Leu
845 850 855

His Gln Leu Met Leu Asp Cys Trp Gln Lys Asp Arg Asn His Arg
860 865 870

Pro Lys Phe Gly Gln Ile Val Asn Thr Leu Asp Lys Met Ile Arg
875 880 885

Asn Pro Asn Ser Leu Lys Ala Met Ala Pro Leu Ser Ser Gly Ile
890 895 900

Asn Leu Pro Leu Leu Asp Arg Thr Ile Pro Asp Tyr Thr Ser Phe
905 910 915

Asn Thr Val Asp Glu Trp Leu Glu Ala Ile Lys Met Gly Gln Tyr
920 925 930

Lys Glu Ser Phe Ala Asn Ala Gly Phe Thr Ser Phe Asp Val Val
935 940 945

Ser Gln Met Met Met Glu Asp Ile Leu Arg Val Gly Val Thr Leu
950 955 960

Ala Gly His Gln Lys Lys Ile Leu Asn Ser Ile Gln Val Met Arg
965 970 975

Ala Gln Met Asn Gln Ile Gln Ser Val Glu Val
980 985

<210> 102
<211> 153
<212> PRT
<213> Homo sapiens

<400> 102
Met Asn Phe Ser Thr Ser Ser Ser Ser Phe Ala Tyr Asp Arg Glu
1 5 10 15

Phe Leu Arg Thr Leu Pro Gly Phe Leu Ile Val Ala Glu Ile Val
20 25 30

Leu Gly Leu Leu Val Trp Thr Leu Ile Ala Gly Thr Glu Tyr Phe
35 40 45

Arg Val Pro Ala Phe Gly Trp Val Met Phe Val Ala Val Phe Tyr
50 55 60

Trp Val Leu Thr Val Phe Phe Leu Ile Ile Tyr Ile Thr Met Thr
65 70 75

Tyr Thr Arg Ile Pro Gln Val Pro Trp Thr Thr Val Gly Leu Cys
80 85 90

Phe Asn Gly Ser Ala Phe Val Leu Tyr Leu Ser Ala Ala Val Val
95 100 105

Asp Ala Ser Ser Val Ser Pro Glu Arg Asp Ser His Asn Phe Asn
110 115 120

Ser Trp Ala Ala Ser Ser Phe Phe Ala Phe Leu Val Asn Ile Cys
125 130 135

Tyr Ala Gly Asn Thr Tyr Phe Ser Phe Ile Ala Trp Arg Ser Arg
140 145 150

Thr Ile Gln

<210> 103
<211> 153
<212> PRT
<213> Homo sapiens

<400> 103
Met Asn Phe Ser Thr Ser Ser Ser Ser Phe Ala Tyr Asp Arg Glu
1 5 10 15

Phe Leu Arg Thr Leu Pro Gly Phe Leu Ile Val Ala Glu Ile Val
20 25 30

Leu Gly Leu Leu Val Trp Thr Leu Ile Ala Gly Thr Glu Tyr Phe
35 40 45

Arg Val Pro Ala Phe Gly Trp Val Met Phe Val Ala Val Phe Tyr
50 55 60

Trp Val Leu Thr Val Phe Phe Leu Ile Ile Tyr Ile Thr Met Thr
65 70 75

Tyr Thr Arg Ile Pro Gln Val Pro Trp Thr Thr Val Gly Leu Cys
80 85 90

Phe Asn Gly Ser Ala Phe Val Leu Tyr Leu Ser Ala Ala Val Val
95 100 105

Asp Ala Ser Ser Val Ser Pro Glu Arg Asp Ser His Asn Phe Asn
110 115 120

+220
1224

Ser Trp Ala Ala Ser Ser Phe Phe Ala Phe Leu Val Thr Ile Cys
125 130 135

Tyr Ala Gly Asn Thr Tyr Phe Ser Phe Ile Ala Trp Arg Ser Arg
140 145 150

Thr Ile Gln

<210> 104
<211> 555
<212> PRT
<213> Homo sapiens

<400> 104
Met Pro Thr Thr Val Asp Asp Val Leu Glu His Gly Gly Glu Phe
1 5 10 15

His Phe Phe Gln Lys Gln Met Phe Phe Leu Leu Ala Leu Leu Ser
20 25 30

Ala Thr Phe Ala Pro Ile Tyr Val Gly Ile Val Phe Leu Gly Phe
35 40 45

Thr Pro Asp His Arg Cys Arg Ser Pro Gly Val Ala Glu Leu Ser
50 55 60

Leu Arg Cys Gly Trp Ser Pro Ala Glu Glu Leu Asn Tyr Thr Val
65 70 75

Pro Gly Pro Gly Pro Ala Gly Glu Ala Ser Pro Arg Gln Cys Arg
80 85 90

Arg Tyr Glu Val Asp Trp Asn Gln Ser Thr Phe Asp Cys Val Asp
95 100 105

Pro Leu Ala Ser Leu Asp Thr Asn Arg Ser Arg Leu Pro Leu Gly
110 115 120

Pro Cys Arg Asp Gly Trp Val Tyr Glu Thr Pro Gly Ser Ser Ile
125 130 135

Val Thr Glu Phe Asn Leu Val Cys Ala Asn Ser Trp Met Leu Asp
140 145 150

Leu Phe Gln Ser Ser Val Asn Val Gly Phe Phe Ile Gly Ser Met
155 160 165

Ser Ile Gly Tyr Ile Ala Asp Arg Phe Gly Arg Lys Leu Cys Leu
170 175 180

Leu Thr Thr Val Leu Ile Asn Ala Ala Ala Gly Val Leu Met Ala
185 190 195

Ile Ser Pro Thr Tyr Thr Trp Met Leu Ile Phe Arg Leu Ile Gln
200 205 210

Gly Leu Val Ser Lys Ala Gly Trp Leu Ile Gly Tyr Ile Leu Ile
215 220 225

Thr Glu Phe Val Gly Arg Arg Tyr Arg Arg Thr Val Gly Ile Phe
230 235 240

Tyr Gln Val Ala Tyr Thr Val Gly Leu Leu Val Leu Ala Gly Val
245 250 255

Ala Tyr Ala Leu Pro His Trp Arg Trp Leu Gln Phe Thr Val Ala
260 265 270

Leu Pro Asn Phe Phe Phe Leu Leu Tyr Tyr Trp Cys Ile Pro Glu
275 280 285

Ser Pro Arg Trp Leu Ile Ser Gln Asn Lys Asn Ala Glu Ala Met
290 295 300

Arg Ile Ile Lys His Ile Ala Lys Lys Asn Gly Lys Ser Leu Pro

305 310 315

Ala Ser Leu Gln Arg Leu Arg Leu Glu Glu Glu Thr Gly Lys Lys
320 325 330

Leu Asn Pro Ser Phe Leu Asp Leu Val Arg Thr Pro Gln Ile Arg
335 340 345

Lys His Thr Met Ile Leu Met Tyr Asn Trp Phe Thr Ser Ser Val
350 355 360

Leu Tyr Gln Gly Leu Ile Met His Met Gly Leu Ala Gly Asp Asn
365 370 375

Ile Tyr Leu Asp Phe Phe Tyr Ser Ala Leu Val Glu Phe Pro Ala
380 385 390

Ala Phe Met Ile Ile Leu Thr Ile Asp Arg Ile Gly Arg Arg Tyr
395 400 405

Pro Trp Ala Ala Ser Asn Met Val Ala Gly Ala Ala Cys Leu Ala
410 415 420

Ser Val Phe Ile Pro Gly Asp Leu Gln Trp Leu Lys Ile Ile Ile
425 430 435

Ser Cys Leu Gly Arg Met Gly Ile Thr Met Ala Tyr Glu Ile Val
440 445 450

Cys Leu Val Asn Ala Glu Leu Tyr Pro Thr Phe Ile Arg Asn Leu
455 460 465

Gly Val His Ile Cys Ser Ser Met Cys Asp Ile Gly Gly Ile Ile
470 475 480

Thr Pro Phe Leu Val Tyr Arg Leu Thr Asn Ile Trp Leu Glu Leu
485 490 495

Pro Leu Met Val Phe Gly Val Leu Gly Leu Val Ala Gly Gly Leu
500 505 510

Val Leu Leu Leu Pro Glu Thr Lys Gly Lys Ala Leu Pro Glu Thr
515 520 525

Ile Glu Glu Ala Glu Asn Met Gln Arg Pro Arg Lys Asn Lys Glu
530 535 540

Lys Met Ile Tyr Leu Gln Val Gln Lys Leu Asp Ile Pro Leu Asn
545 550 555

<210> 105
<211> 555
<212> PRT
<213> Homo sapiens

<400> 105
Met Pro Thr Thr Val Asp Asp Val Leu Glu His Gly Gly Glu Phe
1 5 10 15

His Phe Phe Gln Lys Gln Met Phe Phe Leu Leu Ala Leu Leu Ser
20 25 30

Ala Thr Phe Ala Pro Ile Tyr Val Gly Ile Val Phe Leu Gly Phe
35 40 45

Thr Pro Asp His Arg Cys Arg Ser Pro Gly Val Ala Glu Leu Ser
50 55 60

Leu Arg Cys Gly Trp Ser Pro Ala Glu Glu Leu Asn Tyr Thr Val
65 70 75

Pro Gly Pro Gly Pro Ala Gly Glu Ala Ser Pro Arg Gln Cys Arg
80 85 90

Arg Tyr Glu Val Asp Trp Asn Gln Ser Thr Phe Asp Cys Val Asp
95 100 105

~~1228~~
1226

```
Pro Leu Ala Ser Leu Asp Thr Asn Arg Ser Arg Leu Pro Leu Gly
                110                 115                 120

Pro Cys Arg Asp Gly Trp Val Tyr Glu Thr Pro Gly Ser Ser Ile
                125                 130                 135

Val Thr Glu Phe Asn Leu Val Cys Ala Asn Ser Trp Met Leu Asp
                140                 145                 150

Leu Phe Gln Ser Ser Val Asn Val Gly Phe Phe Ile Gly Ser Met
                155                 160                 165

Ser Ile Gly Tyr Ile Ala Asp Arg Phe Gly Arg Lys Leu Cys Leu
                170                 175                 180

Leu Thr Thr Val Leu Ile Asn Ala Ala Ala Gly Val Leu Met Ala
                185                 190                 195

Ile Ser Pro Thr Tyr Thr Trp Met Leu Ile Phe Arg Leu Ile Gln
                200                 205                 210

Gly Leu Val Ser Lys Ala Gly Trp Leu Ile Gly Tyr Ile Leu Ile
                215                 220                 225

Thr Glu Phe Val Gly Arg Arg Tyr Arg Arg Thr Val Gly Ile Phe
                230                 235                 240

Tyr Gln Val Ala Tyr Thr Val Gly Leu Leu Val Leu Ala Gly Val
                245                 250                 255

Ala Tyr Ala Leu Pro His Trp Arg Trp Leu Gln Phe Thr Val Ser
                260                 265                 270

Leu Pro Asn Phe Phe Phe Leu Leu Tyr Tyr Trp Cys Ile Pro Glu
                275                 280                 285

Ser Pro Arg Trp Leu Ile Ser Gln Asn Lys Asn Ala Glu Ala Met
                290                 295                 300

Arg Ile Ile Lys His Ile Ala Lys Lys Asn Gly Lys Ser Leu Pro
                305                 310                 315

Ala Ser Leu Gln Arg Leu Arg Leu Glu Glu Glu Thr Gly Lys Lys
                320                 325                 330

Leu Asn Pro Ser Phe Leu Asp Leu Val Arg Thr Pro Gln Ile Arg
                335                 340                 345

Lys His Thr Met Ile Leu Met Tyr Asn Trp Phe Thr Ser Ser Val
                350                 355                 360

Leu Tyr Gln Gly Leu Ile Met His Met Gly Leu Ala Gly Asp Asn
                365                 370                 375

Ile Tyr Leu Asp Phe Phe Tyr Ser Ala Leu Val Glu Phe Pro Ala
                380                 385                 390

Ala Phe Met Ile Ile Leu Thr Ile Asp Arg Ile Gly Arg Arg Tyr
                395                 400                 405

Pro Trp Ala Ala Ser Asn Met Val Ala Gly Ala Ala Cys Leu Ala
                410                 415                 420

Ser Val Phe Ile Pro Gly Asp Leu Gln Trp Leu Lys Ile Ile Ile
                425                 430                 435

Ser Cys Leu Gly Arg Met Gly Ile Thr Met Ala Tyr Glu Ile Val
                440                 445                 450

Cys Leu Val Asn Ala Glu Leu Tyr Pro Thr Phe Ile Arg Asn Leu
                455                 460                 465

Gly Val His Ile Cys Ser Ser Met Cys Asp Ile Gly Gly Ile Ile
                470                 475                 480
```

1228
1229

```
Thr Pro Phe Leu Val Tyr Arg Leu Thr Asn Ile Trp Leu Glu Leu
                485                 490                 495

Pro Leu Met Val Phe Gly Val Leu Gly Leu Val Ala Gly Gly Leu
                500                 505                 510

Val Leu Leu Leu Pro Glu Thr Lys Gly Lys Ala Leu Pro Glu Thr
                515                 520                 525

Ile Glu Glu Ala Glu Asn Met Gln Arg Pro Arg Lys Asn Lys Glu
                530                 535                 540

Lys Met Ile Tyr Leu Gln Val Gln Lys Leu Asp Ile Pro Leu Asn
                545                 550                 555
```

<210> 106
<211> 690
<212> PRT
<213> Homo sapiens

<400> 106

```
Met Ala Pro Trp Pro Glu Leu Gly Asp Ala Gln Pro Asn Pro Asp
  1               5                  10                  15

Lys Tyr Leu Glu Gly Ala Ala Gly Gln Gln Pro Thr Ala Pro Asp
                20                  25                  30

Lys Ser Lys Glu Thr Asn Lys Thr Asp Asn Thr Glu Ala Pro Val
                35                  40                  45

Thr Lys Ile Glu Leu Leu Pro Ser Tyr Ser Thr Ala Thr Leu Ile
                50                  55                  60

Asp Glu Pro Thr Glu Val Asp Asp Pro Trp Asn Leu Pro Thr Leu
                65                  70                  75

Gln Asp Ser Gly Ile Lys Trp Ser Glu Arg Asp Thr Lys Gly Lys
                80                  85                  90

Ile Leu Cys Phe Phe Gln Gly Ile Gly Arg Leu Ile Leu Leu Leu
                95                  100                 105

Gly Phe Leu Tyr Phe Phe Val Cys Ser Leu Asp Ile Leu Ser Ser
                110                 115                 120

Ala Phe Gln Leu Val Gly Gly Lys Met Ala Gly Gln Phe Phe Ser
                125                 130                 135

Asn Ser Ser Ile Met Ser Asn Pro Leu Leu Gly Leu Val Ile Gly
                140                 145                 150

Val Leu Val Thr Val Leu Val Gln Ser Ser Ser Thr Ser Thr Ser
                155                 160                 165

Ile Val Val Ser Met Val Ser Ser Ser Leu Leu Thr Val Arg Ala
                170                 175                 180

Ala Ile Pro Ile Ile Met Gly Ala Asn Ile Gly Thr Ser Ile Thr
                185                 190                 195

Asn Thr Ile Val Ala Leu Met Gln Val Gly Asp Arg Ser Glu Phe
                200                 205                 210

Arg Arg Ala Phe Ala Gly Ala Thr Val His Asp Phe Phe Asn Trp
                215                 220                 225

Leu Ser Val Leu Val Leu Leu Pro Val Glu Val Ala Thr His Tyr
                230                 235                 240

Leu Glu Ile Ile Thr Gln Leu Ile Val Glu Ser Phe His Phe Lys
                245                 250                 255

Asn Gly Glu Asp Ala Pro Asp Leu Leu Lys Val Ile Thr Lys Pro
                260                 265                 270

Phe Thr Lys Leu Ile Val Gln Leu Asp Lys Lys Val Ile Ser Gln
```

275 280 285

Ile Ala Met Asn Asp Glu Lys Ala Lys Asn Lys Ser Leu Val Lys
290 295 300

Ile Trp Cys Lys Thr Phe Thr Asn Lys Thr Gln Ile Asn Val Thr
305 310 315

Val Pro Ser Thr Ala Asn Cys Thr Ser Pro Ser Leu Cys Trp Thr
320 325 330

Asp Gly Ile Gln Asn Trp Thr Met Lys Asn Val Thr Tyr Lys Glu
335 340 345

Asn Ile Ala Lys Cys Gln His Ile Phe Val Asn Phe His Leu Pro
350 355 360

Asp Leu Ala Val Gly Thr Ile Leu Leu Ile Leu Ser Leu Leu Val
365 370 375

Leu Cys Gly Cys Leu Ile Met Ile Val Lys Ile Leu Gly Ser Val
380 385 390

Leu Lys Gly Gln Val Ala Thr Val Ile Lys Lys Thr Ile Asn Thr
395 400 405

Asp Phe Pro Phe Pro Phe Ala Trp Leu Thr Gly Tyr Leu Ala Ile
410 415 420

Leu Val Gly Ala Gly Met Thr Phe Ile Val Gln Ser Ser Ser Val
425 430 435

Phe Thr Ser Ala Leu Thr Pro Leu Ile Gly Ile Gly Val Ile Thr
440 445 450

Ile Glu Arg Ala Tyr Pro Leu Thr Leu Gly Ser Asn Ile Gly Thr
455 460 465

Thr Thr Thr Ala Ile Leu Ala Ala Leu Ala Ser Pro Gly Asn Ala
470 475 480

Leu Arg Ser Ser Leu Gln Ile Ala Leu Cys His Phe Phe Phe Asn
485 490 495

Ile Ser Gly Ile Leu Leu Trp Tyr Pro Ile Pro Phe Thr Arg Leu
500 505 510

Pro Ile Arg Met Ala Lys Gly Leu Gly Asn Ile Ser Ala Lys Tyr
515 520 525

Arg Trp Phe Ala Val Phe Tyr Leu Ile Ile Phe Phe Phe Leu Ile
530 535 540

Pro Leu Thr Val Phe Gly Leu Ser Leu Ala Gly Trp Arg Val Leu
545 550 555

Val Gly Val Gly Val Pro Val Val Phe Ile Ile Ile Leu Val Leu
560 565 570

Cys Leu Arg Leu Leu Gln Ser Arg Cys Pro Arg Val Leu Pro Lys
575 580 585

Lys Leu Gln Asn Trp Asn Phe Leu Pro Leu Trp Met Arg Ser Leu
590 595 600

Lys Pro Trp Asp Ala Val Val Ser Lys Phe Thr Gly Cys Phe Gln
605 610 615

Met Arg Cys Cys Tyr Cys Cys Arg Val Cys Cys Arg Ala Cys Cys
620 625 630

Leu Leu Cys Gly Cys Pro Lys Cys Cys Arg Cys Ser Lys Cys Cys
635 640 645

Glu Asp Leu Glu Glu Ala Gln Glu Gly Gln Asp Val Pro Val Lys
650 655 660

Ala Pro Glu Thr Phe Asp Asn Ile Thr Ile Ser Arg Glu Ala Gln
665 670 675

Gly Glu Val Pro Ala Ser Asp Ser Lys Thr Glu Cys Thr Ala Leu
680 685 690

<210> 107
<211> 690
<212> PRT
<213> Homo sapiens

<400> 107
Met Ala Pro Trp Pro Glu Leu Gly Asp Ala Gln Pro Asn Pro Asp
1 5 10 15

Lys Tyr Leu Glu Gly Ala Ala Gly Gln Gln Pro Thr Ala Pro Asp
20 25 30

Lys Ser Lys Glu Thr Asn Lys Thr Asp Asn Thr Glu Ala Pro Val
35 40 45

Thr Lys Ile Glu Leu Leu Pro Ser Tyr Ser Thr Ala Thr Leu Ile
50 55 60

Asp Glu Pro Thr Glu Val Asp Asp Pro Trp Asn Leu Pro Thr Leu
65 70 75

Gln Asp Ser Gly Ile Lys Trp Ser Glu Arg Asp Thr Lys Gly Lys
80 85 90

Ile Leu Cys Phe Phe Gln Gly Ile Gly Arg Leu Ile Leu Leu Leu
95 100 105

Gly Phe Leu Tyr Phe Phe Val Cys Ser Leu Asp Ile Leu Ser Ser
110 115 120

Ala Phe Gln Leu Val Gly Gly Lys Met Ala Gly Gln Phe Phe Ser
125 130 135

Asn Ser Ser Ile Met Ser Asn Pro Leu Leu Gly Leu Val Ile Gly
140 145 150

Val Leu Val Thr Val Leu Val Gln Ser Ser Ser Thr Ser Thr Ser
155 160 165

Ile Val Val Ser Met Val Ser Ser Ser Leu Leu Thr Val Arg Ala
170 175 180

Ala Ile Pro Ile Ile Met Gly Ala Asn Ile Gly Thr Ser Ile Thr
185 190 195

Asn Thr Ile Val Ala Leu Met Gln Val Gly Asp Arg Ser Glu Phe
200 205 210

Arg Arg Ala Phe Ala Gly Ala Thr Val His Asp Phe Phe Asn Trp
215 220 225

Leu Ser Val Leu Val Leu Leu Pro Val Glu Val Ala Thr His Tyr
230 235 240

Leu Glu Ile Ile Thr Gln Leu Ile Val Glu Ser Phe His Phe Lys
245 250 255

Asn Gly Glu Asp Ala Pro Asp Leu Leu Lys Val Ile Thr Lys Pro
260 265 270

Phe Thr Lys Leu Ile Val Gln Leu Asp Lys Lys Val Ile Ser Gln
275 280 285

Ile Ala Met Asn Asp Glu Lys Ala Lys Asn Lys Ser Leu Val Lys
290 295 300

Ile Trp Cys Lys Thr Phe Thr Asn Lys Thr Gln Ile Asn Val Thr
305 310 315

1232
1230

Val Pro Ser Thr Ala Asn Cys Thr Ser Pro Ser Leu Cys Trp Thr
320 325 330

Asp Gly Ile Gln Asn Trp Thr Met Lys Asn Val Thr Tyr Lys Glu
335 340 345

Asn Ile Ala Lys Cys Gln His Ile Phe Val Asn Phe His Leu Pro
350 355 360

Asp Leu Ala Val Gly Thr Ile Leu Leu Ile Leu Ser Leu Leu Val
365 370 375

Leu Cys Gly Cys Leu Ile Met Ile Val Lys Ile Leu Gly Ser Val
380 385 390

Leu Lys Gly Gln Val Ala Thr Val Ile Lys Lys Thr Ile Asn Thr
395 400 405

Asp Phe Pro Phe Pro Phe Ala Trp Leu Thr Gly Tyr Leu Ala Ile
410 415 420

Leu Val Gly Ala Gly Met Thr Phe Ile Val Gln Ser Ser Ser Val
425 430 435

Phe Thr Ser Ala Leu Thr Pro Leu Ile Gly Ile Gly Val Ile Thr
440 445 450

Ile Glu Arg Ala Tyr Pro Leu Thr Leu Gly Ser Asn Ile Gly Thr
455 460 465

Thr Thr Thr Ala Ile Leu Ala Ala Leu Ala Ser Pro Gly Asn Ala
470 475 480

Leu Arg Ser Ser Leu Gln Ile Ala Leu Cys His Phe Phe Phe Asn
485 490 495

Ile Ser Gly Ile Leu Leu Trp Tyr Pro Ile Pro Phe Thr Arg Leu
500 505 510

Pro Ile Arg Met Ala Lys Gly Leu Gly Asn Ile Ser Ala Lys Tyr
515 520 525

Arg Trp Phe Ala Val Phe Tyr Leu Ile Ile Phe Phe Phe Leu Ile
530 535 540

Pro Leu Thr Val Phe Gly Leu Ser Leu Ala Gly Trp Arg Val Leu
545 550 555

Val Gly Val Gly Val Pro Val Val Phe Ile Ile Ile Leu Val Leu
560 565 570

Cys Leu Arg Leu Leu Gln Ser Arg Cys Pro Arg Val Leu Pro Lys
575 580 585

Lys Leu Gln Asn Trp Asn Phe Leu Pro Leu Trp Met Arg Ser Leu
590 595 600

Lys Pro Trp Asp Ala Val Val Ser Lys Phe Thr Gly Cys Phe Gln
605 610 615

Met Arg Cys Cys Cys Cys Cys Arg Val Cys Cys Arg Ala Cys Cys
620 625 630

Leu Leu Cys Gly Cys Pro Lys Cys Cys Arg Cys Ser Lys Cys Cys
635 640 645

Glu Asp Leu Glu Glu Ala Gln Glu Gly Gln Asp Val Pro Val Lys
650 655 660

Ala Pro Glu Thr Phe Asp Asn Ile Thr Ile Ser Arg Glu Ala Gln
665 670 675

Gly Glu Val Pro Ala Ser Asp Ser Lys Thr Glu Cys Thr Ala Leu
680 685 690

<210> 108

<211> 411
<212> PRT
<213> Homo sapiens

<400> 108

```
Met Trp Ser Gly Trp Trp Leu Trp Pro Leu Val Ala Val Cys Thr
  1               5                  10                  15

Ala Asp Phe Phe Arg Asp Glu Ala Glu Arg Ile Met Arg Asp Ser
             20                  25                  30

Pro Val Ile Asp Gly His Asn Asp Leu Pro Trp Gln Leu Leu Asp
             35                  40                  45

Met Phe Asn Asn Arg Leu Gln Asp Glu Arg Ala Asn Leu Thr Thr
             50                  55                  60

Leu Ala Gly Thr His Thr Asn Ile Pro Lys Leu Arg Ala Gly Phe
             65                  70                  75

Val Gly Gly Gln Phe Trp Ser Val Tyr Thr Pro Cys Asp Thr Gln
             80                  85                  90

Asn Lys Asp Ala Val Arg Arg Thr Leu Glu Gln Met Asp Val Val
             95                 100                 105

His Arg Met Cys Arg Met Tyr Pro Glu Thr Phe Leu Tyr Val Thr
            110                 115                 120

Ser Ser Ala Gly Ile Arg Gln Ala Phe Arg Glu Gly Lys Val Ala
            125                 130                 135

Ser Leu Ile Gly Val Glu Gly Gly His Ser Ile Asp Ser Ser Leu
            140                 145                 150

Gly Val Leu Arg Ala Leu Tyr Gln Leu Gly Met Arg Tyr Leu Thr
            155                 160                 165

Leu Thr His Ser Cys Asn Thr Pro Trp Ala Asp Asn Trp Leu Val
            170                 175                 180

Asp Thr Gly Asp Ser Glu Pro Gln Ser Gln Gly Leu Ser Pro Phe
            185                 190                 195

Gly Gln Arg Val Val Lys Glu Leu Asn Arg Leu Gly Val Leu Ile
            200                 205                 210

Asp Leu Ala His Val Ser Val Ala Thr Met Lys Ala Thr Leu Gln
            215                 220                 225

Leu Ser Arg Ala Pro Val Ile Phe Ser His Ser Ser Ala Tyr Ser
            230                 235                 240

Val Cys Ala Ser Arg Arg Asn Val Pro Asp Asp Val Leu Arg Leu
            245                 250                 255

Val Lys Gln Thr Asp Ser Leu Val Met Val Asn Phe Tyr Asn Asn
            260                 265                 270

Tyr Ile Ser Cys Thr Asn Lys Ala Asn Leu Ser Gln Val Ala Asp
            275                 280                 285

His Leu Asp His Ile Lys Glu Val Ala Gly Ala Arg Ala Val Gly
            290                 295                 300

Phe Gly Gly Asp Phe Asp Gly Val Pro Arg Val Pro Glu Gly Leu
            305                 310                 315

Glu Asp Val Ser Lys Tyr Pro Asp Leu Ile Ala Glu Leu Leu Arg
            320                 325                 330

Arg Asn Trp Thr Glu Ala Glu Val Lys Gly Ala Leu Ala Asp Asn
            335                 340                 345

Leu Leu Arg Val Phe Gln Ala Val Glu Gln Ala Ser Asn Leu Thr
            350                 355                 360
```

1234
1232

Gln Ala Pro Glu Glu Glu Pro Ile Pro Leu Asp Gln Leu Gly Gly
365 370 375

Ser Cys Arg Thr His Tyr Gly Tyr Ser Ser Gly Ala Ser Ser Leu
380 385 390

His Arg His Trp Gly Leu Leu Leu Ala Ser Leu Ala Pro Leu Val
395 400 405

Leu Cys Leu Ser Leu Leu
410

<210> 109
<211> 320
<212> PRT
<213> Homo sapiens

<400> 109
Met Ser Ser Cys Asn Phe Thr His Ala Thr Phe Val Leu Ile Gly
1 5 10 15

Ile Pro Gly Leu Glu Lys Ala His Phe Trp Val Gly Phe Pro Leu
20 25 30

Leu Ser Met Tyr Val Val Ala Met Phe Gly Asn Cys Ile Val Val
35 40 45

Phe Ile Val Arg Thr Glu Arg Ser Leu His Ala Pro Met Tyr Leu
50 55 60

Phe Leu Cys Met Leu Ala Ala Ile Asp Leu Ala Leu Ser Thr Ser
65 70 75

Thr Met Pro Lys Ile Leu Ala Leu Phe Trp Phe Asp Ser Arg Glu
80 85 90

Ile Ser Phe Glu Ala Cys Leu Thr Gln Met Phe Phe Ile His Ala
95 100 105

Leu Ser Ala Ile Glu Ser Thr Ile Leu Leu Ala Met Ala Phe Asp
110 115 120

Arg Tyr Val Ala Ile Cys His Pro Leu Arg His Ala Ala Val Leu
125 130 135

Asn Asn Thr Val Thr Ala Gln Ile Gly Ile Val Ala Val Val Arg
140 145 150

Gly Ser Leu Phe Phe Phe Pro Leu Pro Leu Leu Ile Lys Arg Leu
155 160 165

Ala Phe Cys His Ser Asn Val Leu Ser His Ser Tyr Cys Val His
170 175 180

Gln Asp Val Met Lys Leu Ala Tyr Ala Asp Thr Leu Pro Asn Val
185 190 195

Val Tyr Gly Leu Thr Ala Ile Leu Leu Val Met Gly Val Asp Val
200 205 210

Met Phe Ile Ser Leu Ser Tyr Phe Leu Ile Ile Arg Thr Val Leu
215 220 225

Gln Leu Pro Ser Lys Ser Glu Arg Ala Lys Ala Phe Gly Thr Cys
230 235 240

Val Ser His Ile Gly Val Val Leu Ala Phe Tyr Val Pro Leu Ile
245 250 255

Gly Leu Ser Val Val His Arg Phe Gly Asn Ser Leu His Pro Ile
260 265 270

Val Arg Val Val Met Gly Asp Ile Tyr Leu Leu Leu Pro Pro Val
275 280 285

1235
1233

Ile Asn Pro Ile Ile Tyr Gly Ala Lys Thr Lys Gln Ile Arg Thr
290 295 300

Arg Val Leu Ala Met Phe Lys Ile Ser Cys Asp Lys Asp Leu Gln
305 310 315

Ala Val Gly Gly Lys
320

<210> 110
<211> 2314
<212> PRT
<213> Homo sapiens

<400> 110
Met Arg Ile Leu Lys Arg Phe Leu Ala Cys Ile Gln Leu Leu Cys
1 5 10 15

Val Cys Arg Leu Asp Trp Ala Asn Gly Tyr Tyr Arg Gln Gln Arg
20 25 30

Lys Leu Val Glu Glu Ile Gly Trp Ser Tyr Thr Gly Ala Leu Asn
35 40 45

Gln Lys Asn Trp Gly Lys Lys Tyr Pro Thr Cys Asn Ser Pro Lys
50 55 60

Gln Ser Pro Ile Asn Ile Asp Glu Asp Leu Thr Gln Val Asn Val
65 70 75

Asn Leu Lys Lys Leu Lys Phe Gln Gly Trp Asp Lys Thr Ser Leu
80 85 90

Glu Asn Thr Phe Ile His Asn Thr Gly Lys Thr Val Glu Ile Asn
95 100 105

Leu Thr Asn Asp Tyr Arg Val Ser Gly Gly Val Ser Glu Met Val
110 115 120

Phe Lys Ala Ser Lys Ile Thr Phe His Trp Gly Lys Cys Asn Met
125 130 135

Ser Ser Asp Gly Ser Glu His Ser Leu Glu Gly Gln Lys Phe Pro
140 145 150

Leu Glu Met Gln Ile Tyr Cys Phe Asp Ala Asp Arg Phe Ser Ser
155 160 165

Phe Glu Glu Ala Val Lys Gly Lys Gly Lys Leu Arg Ala Leu Ser
170 175 180

Ile Leu Phe Glu Val Gly Thr Glu Glu Asn Leu Asp Phe Lys Ala
185 190 195

Ile Ile Asp Gly Val Glu Ser Val Ser Arg Phe Gly Lys Gln Ala
200 205 210

Ala Leu Asp Pro Phe Ile Leu Leu Asn Leu Leu Pro Asn Ser Thr
215 220 225

Asp Lys Tyr Tyr Ile Tyr Asn Gly Ser Leu Thr Ser Pro Pro Cys
230 235 240

Thr Asp Thr Val Asp Trp Ile Val Phe Lys Asp Thr Val Ser Ile
245 250 255

Ser Glu Ser Gln Leu Ala Val Phe Cys Glu Val Leu Thr Met Gln
260 265 270

Gln Ser Gly Tyr Val Met Leu Met Asp Tyr Leu Gln Asn Asn Phe
275 280 285

Arg Glu Gln Gln Tyr Lys Phe Ser Arg Gln Val Phe Ser Ser Tyr
290 295 300

Thr Gly Lys Glu Glu Ile His Glu Ala Val Cys Ser Ser Glu Pro

305 310 315

Glu Asn Val Gln Ala Asp Pro Glu Asn Tyr Thr Ser Leu Leu Val
320 325 330

Thr Trp Glu Arg Pro Arg Val Val Tyr Asp Thr Met Ile Glu Lys
335 340 345

Phe Ala Val Leu Tyr Gln Gln Leu Asp Gly Glu Asp Gln Thr Lys
350 355 360

His Glu Phe Leu Thr Asp Gly Tyr Gln Asp Leu Gly Ala Ile Leu
365 370 375

Asn Asn Leu Leu Pro Asn Met Ser Tyr Val Leu Gln Ile Val Ala
380 385 390

Ile Cys Thr Asn Gly Leu Tyr Gly Lys Tyr Ser Asp Gln Leu Ile
395 400 405

Val Asp Met Pro Thr Asp Asn Pro Glu Leu Asp Leu Phe Pro Glu
410 415 420

Leu Ile Gly Thr Glu Glu Ile Ile Lys Glu Glu Glu Glu Gly Lys
425 430 435

Asp Ile Glu Glu Gly Ala Ile Val Asn Pro Gly Arg Asp Ser Ala
440 445 450

Thr Asn Gln Ile Arg Lys Lys Glu Pro Gln Ile Ser Thr Thr Thr
455 460 465

His Tyr Asn Arg Ile Gly Thr Lys Tyr Asn Glu Ala Lys Thr Asn
470 475 480

Arg Ser Pro Thr Arg Gly Ser Glu Phe Ser Gly Lys Gly Asp Val
485 490 495

Pro Asn Thr Ser Leu Asn Ser Thr Ser Gln Pro Val Thr Lys Leu
500 505 510

Ala Thr Glu Lys Asp Ile Ser Leu Thr Ser Gln Thr Val Thr Glu
515 520 525

Leu Pro Pro His Thr Val Glu Gly Thr Ser Ala Ser Leu Asn Asp
530 535 540

Gly Ser Lys Thr Val Leu Arg Ser Pro His Met Asn Leu Ser Gly
545 550 555

Thr Ala Glu Ser Leu Asn Thr Val Ser Ile Thr Glu Tyr Glu Glu
560 565 570

Glu Ser Leu Leu Thr Ser Phe Lys Leu Asp Thr Gly Ala Glu Asp
575 580 585

Ser Ser Gly Ser Ser Pro Ala Thr Ser Ala Ile Pro Phe Ile Ser
590 595 600

Glu Asn Ile Ser Gln Gly Tyr Ile Phe Ser Ser Glu Asn Pro Glu
605 610 615

Thr Ile Thr Tyr Asp Val Leu Ile Pro Glu Ser Ala Arg Asn Ala
620 625 630

Ser Glu Asp Ser Thr Ser Ser Gly Ser Glu Glu Ser Leu Lys Asp
635 640 645

Pro Ser Met Glu Gly Asn Val Trp Phe Pro Ser Ser Thr Asp Ile
650 655 660

Thr Ala Gln Pro Asp Val Gly Ser Gly Arg Glu Ser Phe Leu Gln
665 670 675

Thr Asn Tyr Thr Glu Ile Arg Val Asp Glu Ser Glu Lys Thr Thr
680 685 690

1257
1235

Lys Ser Phe Ser Ala Gly Pro Val Met Ser Gln Gly Pro Ser Val
695 700 705

Thr Asp Leu Glu Met Pro His Tyr Ser Thr Phe Ala Tyr Phe Pro
710 715 720

Thr Glu Val Thr Pro His Ala Phe Thr Pro Ser Ser Arg Gln Gln
725 730 735

Asp Leu Val Ser Thr Val Asn Val Val Tyr Ser Gln Thr Thr Gln
740 745 750

Pro Val Tyr Asn Gly Glu Thr Pro Leu Gln Pro Ser Tyr Ser Ser
755 760 765

Glu Val Phe Pro Leu Val Thr Pro Leu Leu Leu Asp Asn Gln Ile
770 775 780

Leu Asn Thr Thr Pro Ala Ala Ser Ser Ser Asp Ser Ala Leu His
785 790 795

Ala Thr Pro Val Phe Pro Ser Val Asp Val Ser Phe Glu Ser Ile
800 805 810

Leu Ser Ser Tyr Asp Gly Ala Pro Leu Leu Pro Phe Ser Ser Ala
815 820 825

Ser Phe Ser Ser Glu Leu Phe Arg His Leu His Thr Val Ser Gln
830 835 840

Ile Leu Pro Gln Val Thr Ser Ala Thr Glu Ser Asp Lys Val Pro
845 850 855

Leu His Ala Ser Leu Pro Val Ala Gly Gly Asp Leu Leu Leu Glu
860 865 870

Pro Ser Leu Ala Gln Tyr Ser Asp Val Leu Ser Thr Thr His Ala
875 880 885

Ala Ser Glu Thr Leu Glu Phe Gly Ser Glu Ser Gly Val Leu Tyr
890 895 900

Lys Thr Leu Met Phe Ser Gln Val Glu Pro Pro Ser Ser Asp Ala
905 910 915

Met Met His Ala Arg Ser Ser Gly Pro Glu Pro Ser Tyr Ala Leu
920 925 930

Ser Asp Asn Glu Gly Ser Gln His Ile Phe Thr Val Ser Tyr Ser
935 940 945

Ser Ala Ile Pro Val His Asp Ser Val Gly Val Thr Tyr Gln Gly
950 955 960

Ser Leu Phe Ser Gly Pro Ser His Ile Pro Ile Pro Lys Ser Ser
965 970 975

Leu Ile Thr Pro Thr Ala Ser Leu Leu Gln Pro Thr His Ala Leu
980 985 990

Ser Gly Asp Gly Glu Trp Ser Gly Ala Ser Ser Asp Ser Glu Phe
995 1000 1005

Leu Leu Pro Asp Thr Asp Gly Leu Thr Ala Leu Asn Ile Ser Ser
1010 1015 1020

Pro Val Ser Val Ala Glu Phe Thr Tyr Thr Thr Ser Val Phe Gly
1025 1030 1035

Asp Asp Asn Lys Ala Leu Ser Lys Ser Glu Ile Ile Tyr Gly Asn
1040 1045 1050

Glu Thr Glu Leu Gln Ile Pro Ser Phe Asn Glu Met Val Tyr Pro
1055 1060 1065

1238
1236

Ser Glu Ser Thr Val Met Pro Asn Met Tyr Asp Asn Val Asn Lys
1070 1075 1080

Leu Asn Ala Ser Leu Gln Glu Thr Ser Val Ser Ile Ser Ser Thr
1085 1090 1095

Lys Gly Met Phe Pro Gly Ser Leu Ala His Thr Thr Thr Lys Val
1100 1105 1110

Phe Asp His Glu Ile Ser Gln Val Pro Glu Asn Asn Phe Ser Val
1115 1120 1125

Gln Pro Thr His Thr Val Ser Gln Ala Ser Gly Asp Thr Ser Leu
1130 1135 1140

Lys Pro Val Leu Ser Ala Asn Ser Glu Pro Ala Ser Ser Asp Pro
1145 1150 1155

Ala Ser Ser Glu Met Leu Ser Pro Ser Thr Gln Leu Leu Phe Tyr
1160 1165 1170

Glu Thr Ser Ala Ser Phe Ser Thr Glu Val Leu Leu Gln Pro Ser
1175 1180 1185

Phe Gln Ala Ser Asp Val Asp Thr Leu Leu Lys Thr Val Leu Pro
1190 1195 1200

Ala Val Pro Ser Asp Pro Ile Leu Val Glu Thr Pro Lys Val Asp
1205 1210 1215

Lys Ile Ser Ser Thr Met Leu His Leu Ile Val Ser Asn Ser Ala
1220 1225 1230

Ser Ser Glu Asn Met Leu His Ser Thr Ser Val Pro Val Phe Asp
1235 1240 1245

Val Ser Pro Thr Ser His Met His Ser Ala Ser Leu Gln Gly Leu
1250 1255 1260

Thr Ile Ser Tyr Ala Ser Glu Lys Tyr Glu Pro Val Leu Leu Lys
1265 1270 1275

Ser Glu Ser Ser His Gln Val Val Pro Ser Leu Tyr Ser Asn Asp
1280 1285 1290

Glu Leu Phe Gln Thr Ala Asn Leu Glu Ile Asn Gln Ala His Pro
1295 1300 1305

Pro Lys Gly Arg His Val Phe Ala Thr Pro Val Leu Ser Ile Asp
1310 1315 1320

Glu Pro Leu Asn Thr Leu Ile Asn Lys Leu Ile His Ser Asp Glu
1325 1330 1335

Ile Leu Thr Ser Thr Lys Ser Ser Val Thr Gly Lys Val Phe Ala
1340 1345 1350

Gly Ile Pro Thr Val Ala Ser Asp Thr Phe Val Ser Thr Asp His
1355 1360 1365

Ser Val Pro Ile Gly Asn Gly His Val Ala Ile Thr Ala Val Ser
1370 1375 1380

Pro His Arg Asp Gly Ser Val Thr Ser Thr Lys Leu Leu Phe Pro
1385 1390 1395

Ser Lys Ala Thr Ser Glu Leu Ser His Ser Ala Lys Ser Asp Ala
1400 1405 1410

Gly Leu Val Gly Gly Gly Glu Asp Gly Asp Thr Asp Asp Asp Gly
1415 1420 1425

Asp Asp Asp Asp Asp Arg Asp Ser Asp Gly Leu Ser Ile His Lys
1430 1435 1440

Cys Met Ser Cys Ser Ser Tyr Arg Glu Ser Gln Glu Lys Val Met

1239
1237

1445 1450 1455

Asn Asp Ser Asp Thr His Glu Asn Ser Leu Met Asp Gln Asn Asn
1460 1465 1470

Pro Ile Ser Tyr Ser Leu Ser Glu Asn Ser Glu Glu Asp Asn Arg
1475 1480 1485

Val Thr Ser Val Ser Ser Asp Ser Gln Thr Gly Met Asp Arg Ser
1490 1495 1500

Pro Gly Lys Ser Pro Ser Ala Asn Gly Leu Ser Gln Lys His Asn
1505 1510 1515

Asp Gly Lys Glu Glu Asn Asp Ile Gln Thr Gly Ser Ala Leu Leu
1520 1525 1530

Pro Leu Ser Pro Glu Ser Lys Ala Trp Ala Val Leu Thr Ser Asp
1535 1540 1545

Glu Glu Ser Gly Ser Gly Gln Gly Thr Ser Asp Ser Leu Asn Glu
1550 1555 1560

Asn Glu Thr Ser Thr Asp Phe Ser Phe Ala Asp Thr Asn Glu Lys
1565 1570 1575

Asp Ala Asp Gly Ile Leu Ala Ala Gly Asp Ser Glu Ile Thr Pro
1580 1585 1590

Gly Phe Pro Gln Ser Pro Thr Ser Ser Val Thr Ser Glu Asn Ser
1595 1600 1605

Glu Val Phe His Val Ser Glu Ala Glu Ala Ser Asn Ser Ser His
1610 1615 1620

Glu Ser Arg Ile Gly Leu Ala Glu Gly Leu Glu Ser Glu Lys Lys
1625 1630 1635

Ala Val Ile Pro Leu Val Ile Val Ser Ala Leu Thr Phe Ile Cys
1640 1645 1650

Leu Val Val Leu Val Gly Ile Leu Ile Tyr Trp Arg Lys Cys Phe
1655 1660 1665

Gln Thr Ala His Phe Tyr Leu Glu Asp Ser Thr Ser Pro Arg Val
1670 1675 1680

Ile Ser Thr Pro Pro Thr Pro Ile Phe Pro Ile Ser Asp Asp Val
1685 1690 1695

Gly Ala Ile Pro Ile Lys His Phe Pro Lys His Val Ala Asp Leu
1700 1705 1710

His Ala Ser Ser Gly Phe Thr Glu Glu Phe Glu Thr Leu Lys Glu
1715 1720 1725

Phe Tyr Gln Glu Val Gln Ser Cys Thr Val Asp Leu Gly Ile Thr
1730 1735 1740

Ala Asp Ser Ser Asn His Pro Asp Asn Lys His Lys Asn Arg Tyr
1745 1750 1755

Ile Asn Ile Val Ala Tyr Asp His Ser Arg Val Lys Leu Ala Gln
1760 1765 1770

Leu Ala Glu Lys Asp Gly Lys Leu Thr Asp Tyr Ile Asn Ala Asn
1775 1780 1785

Tyr Val Asp Gly Tyr Asn Arg Pro Lys Ala Tyr Ile Ala Ala Gln
1790 1795 1800

Gly Pro Leu Lys Ser Thr Ala Glu Asp Phe Trp Arg Met Ile Trp
1805 1810 1815

Glu His Asn Val Glu Val Ile Val Met Ile Thr Asn Leu Val Glu
1820 1825 1830

1240
1238

Lys Gly Arg Arg Lys Cys Asp Gln Tyr Trp Pro Ala Asp Gly Ser
1835 1840 1845

Glu Glu Tyr Gly Asn Phe Leu Val Thr Gln Lys Ser Val Gln Val
1850 1855 1860

Leu Ala Tyr Tyr Thr Val Arg Asn Phe Thr Leu Arg Asn Thr Lys
1865 1870 1875

Ile Lys Lys Gly Ser Gln Lys Gly Arg Pro Ser Gly Arg Val Val
1880 1885 1890

Thr Gln Tyr His Tyr Thr Gln Trp Pro Asp Met Gly Val Pro Glu
1895 1900 1905

Tyr Ser Leu Pro Val Leu Thr Phe Val Arg Lys Ala Ala Tyr Ala
1910 1915 1920

Lys Arg His Ala Val Gly Pro Val Val Val His Cys Ser Ala Gly
1925 1930 1935

Val Gly Arg Thr Gly Thr Tyr Ile Val Leu Asp Ser Met Leu Gln
1940 1945 1950

Gln Ile Gln His Glu Gly Thr Val Asn Ile Phe Gly Phe Leu Lys
1955 1960 1965

His Ile Arg Ser Gln Arg Asn Tyr Leu Val Gln Thr Glu Glu Gln
1970 1975 1980

Tyr Val Phe Ile His Asp Thr Leu Val Glu Ala Ile Leu Ser Lys
1985 1990 1995

Glu Thr Glu Val Leu Asp Ser His Ile His Ala Tyr Val Asn Ala
2000 2005 2010

Leu Leu Ile Pro Gly Pro Ala Gly Lys Thr Lys Leu Glu Lys Gln
2015 2020 2025

Phe Gln Leu Leu Ser Gln Ser Asn Ile Gln Gln Ser Asp Tyr Ser
2030 2035 2040

Ala Ala Leu Lys Gln Cys Asn Arg Glu Lys Asn Arg Thr Ser Ser
2045 2050 2055

Ile Ile Pro Val Glu Arg Ser Arg Val Gly Ile Ser Ser Leu Ser
2060 2065 2070

Gly Glu Gly Thr Asp Tyr Ile Asn Ala Ser Tyr Ile Met Gly Tyr
2075 2080 2085

Tyr Gln Ser Asn Glu Phe Ile Ile Thr Gln His Pro Leu Leu His
2090 2095 2100

Thr Ile Lys Asp Phe Trp Arg Met Ile Trp Asp His Asn Ala Gln
2105 2110 2115

Leu Val Val Met Ile Pro Asp Gly Gln Asn Met Ala Glu Asp Glu
2120 2125 2130

Phe Val Tyr Trp Pro Asn Lys Asp Glu Pro Ile Asn Cys Glu Ser
2135 2140 2145

Phe Lys Val Thr Leu Met Ala Glu Glu His Lys Cys Leu Ser Asn
2150 2155 2160

Glu Glu Lys Leu Ile Ile Gln Asp Phe Ile Leu Glu Ala Thr Gln
2165 2170 2175

Asp Asp Tyr Val Leu Glu Val Arg His Phe Gln Cys Pro Lys Trp
2180 2185 2190

Pro Asn Pro Asp Ser Pro Ile Ser Lys Thr Phe Glu Leu Ile Ser
2195 2200 2205

Val Ile Lys Glu Glu Ala Ala Asn Arg Asp Gly Pro Met Ile Val
2210 2215 2220

His Asp Glu His Gly Gly Val Thr Ala Gly Thr Phe Cys Ala Leu
2225 2230 2235

Thr Thr Leu Met His Gln Leu Glu Lys Glu Asn Ser Val Asp Val
2240 2245 2250

Tyr Gln Val Ala Lys Met Ile Asn Leu Met Arg Pro Gly Val Phe
2255 2260 2265

Ala Asp Ile Glu Gln Tyr Gln Phe Leu Tyr Lys Val Ile Leu Ser
2270 2275 2280

Leu Val Ser Thr Arg Gln Glu Glu Asn Pro Ser Thr Ser Leu Asp
2285 2290 2295

Ser Asn Gly Ala Ala Leu Pro Asp Gly Asn Ile Ala Glu Ser Leu
2300 2305 2310

Glu Ser Leu Val

<210> 111
<211> 774
<212> PRT
<213> Homo sapiens

<400> 111
Met Cys Ala Arg Met Ala Gly Arg Thr Arg Ala Ala Pro Arg Gly
1 5 10 15

Pro Tyr Gly Pro Trp Leu Cys Leu Leu Val Ala Leu Ala Leu Asp
20 25 30

Val Val Arg Val Asp Cys Gly Gln Ala Pro Leu Asp Pro Val Tyr
35 40 45

Leu His Val Thr Ala Ala Arg Pro Ala Gln Pro Thr Leu Trp Thr
50 55 60

Ala Lys Leu Asp Arg Phe Lys Gly Ser Arg His His Thr Thr Leu
65 70 75

Ile Thr Cys His Arg Ala Gly Leu Thr Glu Pro Asp Ser Ser Ser
80 85 90

Pro Leu Glu Leu Ser Glu Phe Leu Trp Val Asp Phe Val Val Glu
95 100 105

Asn Ser Thr Gly Gly Gly Val Ala Val Thr Arg Pro Val Thr Trp
110 115 120

Gln Leu Glu Tyr Pro Gly Gln Ala Pro Glu Ala Glu Lys Asp Lys
125 130 135

Met Val Trp Glu Ile Leu Val Ser Glu Arg Asp Ile Arg Ala Leu
140 145 150

Ile Pro Leu Ala Lys Ala Glu Glu Leu Val Asn Thr Ala Pro Leu
155 160 165

Thr Gly Val Pro Gln His Val Pro Val Arg Leu Val Thr Val Asp
170 175 180

Gly Gly Gly Ala Leu Val Glu Val Thr Glu His Val Gly Cys Glu
185 190 195

Ser Ala Asn Thr Gln Val Leu Gln Val Ser Glu Ala Cys Asp Ala
200 205 210

Val Phe Val Ala Gly Lys Glu Ser Arg Gly Ala Arg Gly Val Arg
215 220 225

Val Asp Phe Trp Trp Arg Arg Leu Arg Ala Ser Leu Arg Leu Thr

1242
1240

230 235 240

Val Trp Ala Pro Leu Leu Pro Leu Arg Ile Glu Leu Thr Asp Thr
245 250 255

Thr Leu Glu Gln Val Arg Gly Trp Arg Val Pro Gly Pro Ala Glu
260 265 270

Gly Pro Ala Glu Pro Ala Ala Glu Ala Ser Asp Glu Ala Glu Arg
275 280 285

Arg Ala Arg Gly Cys His Leu Gln Tyr Gln Arg Ala Gly Val Arg
290 295 300

Phe Leu Ala Pro Phe Ala Ala His Pro Leu Asp Gly Gly Arg Arg
305 310 315

Leu Thr His Leu Leu Gly Pro Asp Trp Leu Leu Asp Val Ser His
320 325 330

Leu Val Ala Pro His Ala Arg Val Leu Asp Ser Arg Val Ala Ser
335 340 345

Leu Glu Gly Gly Arg Val Val Val Gly Arg Glu Pro Gly Val Thr
350 355 360

Ser Ile Glu Val Arg Ser Pro Leu Ser Asp Ser Ile Leu Gly Glu
365 370 375

Gln Ala Leu Ala Val Thr Asp Asp Lys Val Ser Val Leu Glu Leu
380 385 390

Arg Val Gln Pro Val Met Gly Ile Ser Leu Thr Leu Ser Arg Gly
395 400 405

Thr Ala His Pro Gly Glu Val Thr Ala Thr Cys Trp Ala Gln Ser
410 415 420

Ala Leu Pro Ala Pro Lys Gln Glu Val Ala Leu Ser Leu Trp Leu
425 430 435

Ser Phe Ser Asp His Thr Val Ala Pro Ala Glu Leu Tyr Asp Arg
440 445 450

Arg Asp Leu Gly Leu Ser Val Ser Ala Glu Glu Pro Gly Ala Ile
455 460 465

Leu Pro Ala Glu Glu Gln Gly Ala Gln Leu Gly Val Val Val Ser
470 475 480

Gly Ala Gly Ala Glu Gly Leu Pro Leu His Val Ala Leu His Pro
485 490 495

Pro Glu Pro Cys Arg Arg Gly Arg His Arg Val Pro Leu Ala Ser
500 505 510

Gly Thr Ala Trp Leu Gly Leu Pro Pro Ala Ser Thr Pro Ala Pro
515 520 525

Ala Leu Pro Ser Ser Pro Ala Trp Ser Pro Pro Ala Thr Glu Ala
530 535 540

Thr Met Gly Gly Lys Arg Gln Val Ala Gly Ser Val Gly Gly Asn
545 550 555

Thr Gly Val Arg Gly Lys Phe Glu Arg Ala Glu Glu Glu Ala Arg
560 565 570

Lys Glu Glu Thr Glu Ala Arg Glu Glu Glu Glu Glu Glu Glu Glu
575 580 585

Glu Met Val Pro Ala Pro Gln His Val Thr Glu Leu Glu Leu Gly
590 595 600

Met Tyr Ala Leu Leu Gly Val Phe Cys Val Ala Ile Phe Ile Phe
605 610 615

1245
1241

Leu Val Asn Gly Val Val Phe Val Leu Arg Tyr Gln Arg Lys Glu
620 625 630

Pro Pro Asp Ser Ala Thr Asp Pro Thr Ser Pro Gln Pro His Asn
635 640 645

Trp Val Trp Leu Gly Thr Asp Gln Glu Glu Leu Ser Arg Gln Leu
650 655 660

Asp Arg Gln Ser Pro Gly Pro Pro Lys Gly Glu Gly Ser Cys Pro
665 670 675

Cys Glu Ser Gly Gly Gly Gly Glu Ala Pro Thr Leu Ala Pro Gly
680 685 690

Pro Pro Gly Gly Thr Thr Ser Ser Ser Ser Thr Leu Ala Arg Lys
695 700 705

Glu Ala Gly Gly Arg Arg Lys Arg Val Glu Phe Val Thr Phe Val
710 715 720

Pro Ala Pro Pro Ala Gln Ser Pro Glu Glu Pro Val Gly Ala Pro
725 730 735

Ala Val Gln Ser Ile Leu Val Ala Gly Glu Glu Asp Ile Arg Trp
740 745 750

Val Cys Glu Asp Met Gly Leu Lys Asp Pro Glu Glu Leu Arg Asn
755 760 765

Tyr Met Glu Arg Ile Arg Gly Ser Ser
770

<210> 112
<211> 420
<212> PRT
<213> Homo sapiens

<400> 112
Met Ala Thr Lys Thr Glu Leu Ser Pro Thr Ala Arg Glu Ser Lys
1 5 10 15

Asn Ala Gln Asp Met Gln Val Asp Glu Thr Leu Ile Pro Arg Lys
20 25 30

Gly Pro Ser Leu Cys Ser Ala Arg Tyr Gly Ile Ala Leu Val Leu
35 40 45

His Phe Cys Asn Phe Thr Thr Ile Ala Gln Asn Val Ile Met Asn
50 55 60

Ile Thr Met Val Ala Met Val Asn Ser Thr Ser Pro Gln Ser Gln
65 70 75

Leu Asn Asp Ser Ser Glu Val Leu Pro Val Asp Ser Phe Gly Gly
80 85 90

Leu Ser Lys Ala Pro Lys Ser Leu Pro Ala Lys Ser Ser Ile Leu
95 100 105

Gly Gly Gln Phe Ala Ile Trp Glu Lys Trp Gly Pro Pro Gln Glu
110 115 120

Arg Ser Arg Leu Cys Ser Ile Ala Leu Ser Gly Met Leu Leu Gly
125 130 135

Cys Phe Thr Ala Ile Leu Ile Gly Gly Phe Ile Ser Glu Thr Leu
140 145 150

Gly Trp Pro Phe Val Phe Tyr Ile Phe Gly Gly Val Gly Cys Val
155 160 165

Cys Cys Leu Leu Trp Phe Val Val Ile Tyr Asp Asp Pro Phe Ser
170 175 180

Tyr Pro Trp Ile Ser Thr Ser Glu Lys Glu Tyr Ile Ile Ser Ser
185 190 195

Leu Lys Gln Gln Val Gly Ser Ser Lys Gln Pro Leu Pro Ile Lys
200 205 210

Ala Met Leu Arg Ser Leu Pro Ile Trp Ser Ile Cys Leu Gly Cys
215 220 225

Phe Ser His Gln Trp Leu Val Ser Thr Met Val Val Tyr Ile Pro
230 235 240

Thr Tyr Ile Ser Ser Val Tyr His Val Asn Ile Arg Asp Asn Gly
245 250 255

Leu Leu Ser Ala Leu Pro Phe Ile Val Ala Trp Val Ile Gly Met
260 265 270

Val Gly Gly Tyr Leu Ala Asp Phe Leu Leu Thr Lys Lys Phe Arg
275 280 285

Leu Ile Thr Val Arg Lys Ile Ala Thr Ile Leu Gly Ser Leu Pro
290 295 300

Ser Ser Ala Leu Ile Val Ser Leu Pro Tyr Leu Asn Ser Gly Tyr
305 310 315

Ile Thr Ala Thr Ala Leu Leu Thr Leu Ser Cys Gly Leu Ser Thr
320 325 330

Leu Cys Gln Ser Gly Ile Tyr Ile Asn Val Leu Asp Ile Ala Pro
335 340 345

Arg Tyr Ser Ser Phe Leu Met Gly Ala Ser Arg Gly Phe Ser Ser
350 355 360

Ile Ala Pro Val Ile Val Pro Thr Val Ser Gly Phe Leu Leu Ser
365 370 375

Gln Asp Pro Glu Phe Gly Trp Arg Asn Val Phe Phe Leu Leu Phe
380 385 390

Ala Val Asn Leu Leu Gly Leu Leu Phe Tyr Leu Ile Phe Gly Glu
395 400 405

Ala Asp Val Gln Glu Trp Ala Lys Glu Arg Lys Leu Thr Arg Leu
410 415 420

<210> 113
<211> 401
<212> PRT
<213> Homo sapiens

<400> 113
Met Gln Val Asp Glu Thr Leu Ile Pro Arg Lys Val Pro Ser Leu
1 5 10 15

Cys Ser Ala Arg Tyr Gly Ile Ala Leu Val Leu His Phe Cys Asn
20 25 30

Phe Thr Thr Ile Ala Gln Asn Val Ile Met Asn Ile Thr Met Val
35 40 45

Ala Met Val Asn Ser Thr Ser Pro Gln Ser Gln Leu Asn Asp Ser
50 55 60

Ser Glu Val Leu Pro Val Asp Ser Phe Gly Gly Leu Ser Lys Ala
65 70 75

Pro Lys Ser Leu Pro Thr Lys Ser Ser Ile Leu Gly Gly Gln Phe
80 85 90

Ala Ile Trp Glu Lys Trp Gly Pro Pro Gln Glu Arg Ser Arg Leu
95 100 105

Cys Ser Ile Ala Leu Ser Gly Met Leu Leu Gly Cys Phe Thr Ala

1243

110 115 120

Ile Leu Ile Gly Gly Phe Ile Ser Glu Thr Leu Gly Trp Pro Phe
125 130 135

Val Phe Tyr Ile Phe Gly Gly Val Gly Cys Val Cys Cys Leu Leu
140 145 150

Trp Phe Val Val Ile Tyr Asp Asp Pro Val Ser Tyr Pro Trp Ile
155 160 165

Ser Thr Ser Glu Lys Glu Tyr Ile Ile Ser Ser Leu Lys Gln Gln
170 175 180

Val Arg Ser Ser Lys Gln Pro Leu Pro Ile Lys Ala Met Leu Arg
185 190 195

Ser Leu Pro Ile Trp Ser Ile Cys Leu Gly Cys Phe Ser His Gln
200 205 210

Trp Leu Val Ser Thr Met Val Val Tyr Ile Pro Thr Tyr Ile Ser
215 220 225

Ser Val Tyr His Val Asn Ile Arg Asp Asn Gly Leu Leu Ser Ala
230 235 240

Leu Pro Phe Ile Val Ala Trp Val Ile Gly Met Val Gly Gly Tyr
245 250 255

Leu Ala Asp Phe Leu Leu Thr Lys Lys Phe Arg Leu Ile Thr Val
260 265 270

Arg Lys Ile Ala Thr Ile Leu Gly Ser Leu Pro Ser Ser Ala Leu
275 280 285

Ile Val Ser Leu Pro Tyr Leu Asn Ser Gly Tyr Ile Thr Ala Thr
290 295 300

Ala Leu Leu Thr Leu Ser Cys Gly Leu Ser Thr Leu Cys Gln Ser
305 310 315

Gly Ile Tyr Ile Asn Val Leu Asp Ile Ala Pro Arg Tyr Ser Ser
320 325 330

Phe Leu Met Gly Ala Ser Arg Gly Phe Ser Ser Ile Ala Pro Val
335 340 345

Ile Val Pro Thr Val Ser Gly Phe Leu Leu Ser Gln Asp Pro Glu
350 355 360

Phe Gly Trp Arg Asn Val Phe Phe Leu Leu Phe Ala Val Asn Leu
365 370 375

Leu Gly Leu Leu Phe Tyr Leu Ile Phe Gly Glu Ala Asp Val Gln
380 385 390

Glu Trp Ala Lys Glu Arg Lys Leu Thr Arg Leu
395 400

<210> 114
<211> 507
<212> PRT
<213> Homo sapiens

<400> 114
Met Ala Gly Ala Gly Pro Lys Arg Arg Ala Leu Ala Ala Pro Ala
1 5 10 15

Ala Glu Glu Lys Glu Glu Ala Arg Glu Lys Met Leu Ala Ala Lys
20 25 30

Ser Ala Asp Gly Ser Ala Pro Ala Gly Glu Gly Glu Gly Val Thr
35 40 45

Leu Gln Arg Asn Ile Thr Leu Leu Asn Gly Val Ala Ile Ile Val
50 55 60

PCT/US 1244

```
Gly Thr Ile Ile Gly Ser Gly Ile Phe Val Thr Pro Thr Gly Val
                65                  70                  75

Leu Lys Glu Ala Gly Ser Pro Gly Leu Ala Leu Val Val Trp Ala
                80                  85                  90

Ala Cys Gly Val Phe Ser Ile Val Gly Ala Leu Cys Tyr Ala Glu
                95                 100                 105

Leu Gly Thr Thr Ile Ser Lys Ser Gly Gly Asp Tyr Ala Tyr Met
               110                 115                 120

Leu Glu Val Tyr Gly Ser Leu Pro Ala Phe Leu Lys Leu Trp Ile
               125                 130                 135

Glu Leu Leu Ile Ile Arg Pro Ser Ser Gln Tyr Ile Val Ala Leu
               140                 145                 150

Val Phe Ala Thr Tyr Leu Leu Lys Pro Leu Phe Pro Thr Cys Pro
               155                 160                 165

Val Pro Glu Glu Ala Ala Lys Leu Val Ala Cys Leu Cys Val Leu
               170                 175                 180

Leu Leu Thr Ala Val Asn Cys Tyr Ser Val Lys Ala Ala Thr Arg
               185                 190                 195

Val Gln Asp Ala Phe Ala Ala Ala Lys Leu Leu Ala Leu Ala Leu
               200                 205                 210

Ile Ile Leu Leu Gly Phe Val Gln Ile Gly Lys Gly Asp Val Ser
               215                 220                 225

Asn Leu Asp Pro Lys Phe Ser Phe Glu Gly Thr Lys Leu Asp Val
               230                 235                 240

Gly Asn Ile Val Leu Ala Leu Tyr Ser Gly Leu Phe Ala Tyr Gly
               245                 250                 255

Gly Trp Asn Tyr Leu Asn Phe Val Thr Glu Glu Met Ile Asn Pro
               260                 265                 270

Tyr Arg Asn Leu Pro Leu Ala Ile Ile Ile Ser Leu Pro Ile Val
               275                 280                 285

Thr Leu Val Tyr Val Leu Thr Asn Leu Ala Tyr Phe Thr Thr Leu
               290                 295                 300

Ser Thr Glu Gln Met Leu Ser Ser Glu Ala Val Ala Val Asp Phe
               305                 310                 315

Gly Asn Tyr His Leu Gly Val Met Ser Trp Ile Ile Pro Val Phe
               320                 325                 330

Val Gly Leu Ser Cys Phe Gly Ser Val Asn Gly Ser Leu Phe Thr
               335                 340                 345

Ser Ser Arg Leu Phe Phe Val Gly Ser Arg Glu Gly His Leu Pro
               350                 355                 360

Ser Ile Leu Ser Met Ile His Pro Gln Leu Leu Thr Pro Val Pro
               365                 370                 375

Ser Leu Val Phe Thr Cys Val Met Thr Leu Leu Tyr Ala Phe Ser
               380                 385                 390

Lys Asp Ile Phe Ser Val Ile Asn Phe Phe Ser Phe Phe Asn Trp
               395                 400                 405

Leu Cys Val Ala Leu Ala Ile Ile Gly Met Ile Trp Leu Arg His
               410                 415                 420

Arg Lys Pro Glu Leu Glu Arg Pro Ile Lys Val Asn Leu Ala Leu
               425                 430                 435
```

```
Pro Val Phe Phe Ile Leu Ala Cys Leu Phe Leu Ile Ala Val Ser
                440                 445                 450

Phe Trp Lys Thr Pro Val Glu Cys Gly Ile Gly Phe Thr Ile Ile
                455                 460                 465

Leu Ser Gly Leu Pro Val Tyr Phe Phe Gly Val Trp Trp Lys Asn
                470                 475                 480

Lys Pro Lys Trp Leu Leu Gln Gly Ile Phe Ser Thr Thr Val Leu
                485                 490                 495

Cys Gln Lys Leu Met Gln Val Val Pro Gln Glu Thr
                500                 505
```

<210> 115
<211> 507
<212> PRT
<213> Homo sapiens

<400> 115

```
Met Ala Gly Ala Gly Pro Lys Arg Arg Ala Leu Ala Ala Pro Ala
  1               5                  10                  15

Ala Glu Glu Lys Glu Glu Ala Arg Glu Lys Met Leu Ala Ala Lys
                20                  25                  30

Ser Ala Asp Gly Ser Ala Pro Ala Gly Glu Gly Glu Gly Val Thr
                35                  40                  45

Leu Gln Arg Asn Ile Thr Leu Leu Asn Gly Val Ala Ile Ile Val
                50                  55                  60

Gly Thr Ile Ile Gly Ser Gly Ile Phe Val Thr Pro Thr Gly Val
                65                  70                  75

Leu Lys Glu Ala Gly Ser Pro Gly Leu Ala Leu Val Val Trp Ala
                80                  85                  90

Ala Cys Gly Val Phe Ser Ile Val Gly Ala Leu Cys Tyr Ala Glu
                95                 100                 105

Leu Gly Thr Thr Ile Ser Lys Ser Gly Gly Asp Tyr Ala Tyr Met
               110                 115                 120

Leu Glu Val Tyr Gly Ser Leu Pro Ala Phe Leu Lys Leu Trp Ile
               125                 130                 135

Glu Leu Leu Ile Ile Arg Pro Ser Ser Gln Tyr Ile Val Ala Leu
               140                 145                 150

Val Phe Ala Thr Tyr Leu Leu Lys Pro Leu Phe Pro Thr Cys Pro
               155                 160                 165

Val Pro Glu Glu Ala Ala Lys Leu Val Ala Cys Leu Cys Val Leu
               170                 175                 180

Leu Leu Thr Ala Val Asn Cys Tyr Ser Val Lys Ala Ala Thr Arg
               185                 190                 195

Val Gln Asp Ala Phe Ala Ala Ala Lys Leu Leu Ala Leu Ala Leu
               200                 205                 210

Ile Ile Leu Leu Gly Phe Val Gln Ile Gly Lys Gly Asp Val Ser
               215                 220                 225

Asn Leu Asp Pro Asn Phe Ser Phe Glu Gly Thr Lys Leu Asp Val
               230                 235                 240

Gly Asn Ile Val Leu Ala Leu Tyr Ser Gly Leu Phe Ala Tyr Gly
               245                 250                 255

Gly Trp Asn Tyr Leu Asn Phe Val Thr Glu Glu Met Ile Asn Pro
               260                 265                 270

Tyr Arg Asn Leu Pro Leu Ala Ile Ile Ile Ser Leu Pro Ile Val
```

275 280 285

Thr Leu Val Tyr Val Leu Thr Asn Leu Ala Tyr Phe Thr Thr Leu
290 295 300

Ser Thr Glu Gln Met Leu Ser Ser Glu Ala Val Ala Val Asp Phe
305 310 315

Gly Asn Tyr His Leu Gly Val Met Ser Trp Ile Ile Pro Val Phe
320 325 330

Val Gly Leu Ser Cys Phe Gly Ser Val Asn Gly Ser Leu Phe Thr
335 340 345

Ser Ser Arg Leu Phe Phe Val Gly Ser Arg Glu Gly His Leu Pro
350 355 360

Ser Ile Leu Ser Met Ile His Pro Gln Leu Leu Thr Pro Val Pro
365 370 375

Ser Leu Val Phe Thr Cys Val Met Thr Leu Leu Tyr Ala Phe Ser
380 385 390

Lys Asp Ile Phe Ser Val Ile Asn Phe Phe Ser Phe Phe Asn Trp
395 400 405

Leu Cys Val Ala Leu Ala Ile Ile Gly Met Ile Trp Leu Arg His
410 415 420

Arg Lys Pro Glu Leu Glu Arg Pro Ile Lys Val Asn Leu Ala Leu
425 430 435

Pro Val Phe Phe Ile Leu Ala Cys Leu Phe Leu Ile Ala Val Ser
440 445 450

Phe Trp Lys Thr Pro Val Glu Cys Gly Ile Gly Phe Thr Ile Ile
455 460 465

Leu Ser Gly Leu Pro Val Tyr Phe Phe Gly Val Trp Trp Lys Asn
470 475 480

Lys Pro Lys Trp Leu Leu Gln Gly Ile Phe Ser Thr Thr Val Leu
485 490 495

Cys Gln Lys Leu Met Gln Val Val Pro Gln Glu Thr
500 505

<210> 116
<211> 622
<212> PRT
<213> Homo sapiens

<400> 116
Met Ala Leu Pro Thr Ala Arg Pro Leu Leu Gly Ser Cys Gly Thr
1 5 10 15

Pro Ala Leu Gly Ser Leu Leu Phe Leu Leu Phe Ser Leu Gly Trp
20 25 30

Val Gln Pro Ser Arg Thr Leu Ala Gly Glu Thr Gly Gln Glu Ala
35 40 45

Ala Pro Leu Asp Gly Val Leu Ala Asn Pro Pro Asn Ile Ser Ser
50 55 60

Leu Ser Pro Arg Gln Leu Leu Gly Phe Pro Cys Ala Glu Val Ser
65 70 75

Gly Leu Ser Thr Glu Arg Val Arg Glu Leu Ala Val Ala Leu Ala
80 85 90

Gln Lys Asn Val Lys Leu Ser Thr Glu Gln Leu Arg Cys Leu Ala
95 100 105

His Arg Leu Ser Glu Pro Pro Glu Asp Leu Asp Ala Leu Pro Leu
110 115 120

```
Asp Leu Leu Leu Phe Leu Asn Pro Asp Ala Phe Ser Gly Pro Gln
                125                 130                 135

Ala Cys Thr Arg Phe Phe Ser Arg Ile Thr Lys Ala Asn Val Asp
                140                 145                 150

Leu Leu Pro Arg Gly Ala Pro Glu Arg Gln Arg Leu Leu Pro Ala
                155                 160                 165

Ala Leu Ala Cys Trp Gly Val Arg Gly Ser Leu Leu Ser Glu Ala
                170                 175                 180

Asp Val Arg Ala Leu Gly Gly Leu Ala Cys Asp Leu Pro Gly Arg
                185                 190                 195

Phe Val Ala Glu Ser Ala Glu Val Leu Leu Pro Arg Leu Val Ser
                200                 205                 210

Cys Pro Gly Pro Leu Asp Gln Asp Gln Gln Glu Ala Ala Arg Ala
                215                 220                 225

Ala Leu Gln Gly Gly Gly Pro Pro Tyr Gly Pro Pro Ser Thr Trp
                230                 235                 240

Ser Val Ser Thr Met Asp Ala Leu Arg Gly Leu Leu Pro Val Leu
                245                 250                 255

Gly Gln Pro Ile Ile Arg Ser Ile Pro Gln Gly Ile Val Ala Ala
                260                 265                 270

Trp Arg Gln Arg Ser Ser Arg Asp Pro Ser Trp Arg Gln Pro Glu
                275                 280                 285

Arg Thr Ile Leu Arg Pro Arg Phe Arg Arg Glu Val Glu Lys Thr
                290                 295                 300

Ala Cys Pro Ser Gly Lys Lys Ala Arg Glu Ile Asp Glu Ser Leu
                305                 310                 315

Ile Phe Tyr Lys Lys Trp Glu Leu Glu Ala Cys Val Asp Ala Ala
                320                 325                 330

Leu Leu Ala Thr Gln Met Asp Arg Val Asn Ala Ile Pro Phe Thr
                335                 340                 345

Tyr Glu Gln Leu Asp Val Leu Lys His Lys Leu Asp Glu Leu Tyr
                350                 355                 360

Pro Gln Gly Tyr Pro Glu Ser Val Ile Gln His Leu Gly Tyr Leu
                365                 370                 375

Phe Leu Lys Met Ser Pro Glu Asp Ile Arg Lys Trp Asn Val Thr
                380                 385                 390

Ser Leu Glu Thr Leu Lys Ala Leu Leu Glu Val Asn Lys Gly His
                395                 400                 405

Glu Met Ser Pro Gln Val Ala Thr Leu Ile Asp Arg Phe Val Lys
                410                 415                 420

Gly Arg Gly Gln Leu Asp Lys Asp Thr Leu Asp Thr Leu Thr Ala
                425                 430                 435

Phe Tyr Pro Gly Tyr Leu Cys Ser Leu Ser Pro Glu Glu Leu Ser
                440                 445                 450

Ser Val Pro Pro Ser Ser Ile Trp Ala Val Arg Pro Gln Asp Leu
                455                 460                 465

Asp Thr Cys Asp Pro Arg Gln Leu Asp Val Leu Tyr Pro Lys Ala
                470                 475                 480

Arg Leu Ala Phe Gln Asn Met Asn Gly Ser Glu Tyr Phe Val Lys
                485                 490                 495
```

```
Ile Gln Ser Phe Leu Gly Gly Ala Pro Thr Glu Asp Leu Lys Ala
                500                 505                 510

Leu Ser Gln Gln Asn Val Ser Met Asp Leu Ala Thr Phe Met Lys
                515                 520                 525

Leu Arg Thr Asp Ala Val Leu Pro Leu Thr Val Ala Glu Val Gln
                530                 535                 540

Lys Leu Leu Gly Pro His Val Glu Gly Leu Lys Ala Glu Glu Arg
                545                 550                 555

His Arg Pro Val Arg Asp Trp Ile Leu Arg Gln Arg Gln Asp Asp
                560                 565                 570

Leu Asp Thr Leu Gly Leu Gly Leu Gln Gly Gly Ile Pro Asn Gly
                575                 580                 585

Tyr Leu Val Leu Asp Leu Ser Val Gln Glu Ala Leu Ser Gly Thr
                590                 595                 600

Pro Cys Leu Leu Gly Pro Gly Pro Val Leu Thr Val Leu Ala Leu
                605                 610                 615

Leu Leu Ala Ser Thr Leu Ala
                620
```

<210> 117
<211> 621
<212> PRT
<213> Homo sapiens

<400> 117

```
Ala Leu Pro Thr Ala Arg Pro Leu Leu Gly Ser Cys Gly Thr Pro
  1               5                  10                  15

Ala Leu Gly Ser Leu Leu Phe Leu Leu Phe Ser Leu Gly Trp Val
                20                  25                  30

Gln Pro Ser Arg Thr Leu Ala Gly Glu Thr Gly Gln Glu Ala Ala
                35                  40                  45

Pro Leu Asp Gly Val Leu Ala Asn Pro Pro Asn Ile Ser Ser Leu
                50                  55                  60

Ser Pro Arg Gln Leu Leu Gly Phe Pro Cys Ala Glu Val Ser Gly
                65                  70                  75

Leu Ser Thr Glu Arg Val Arg Glu Leu Ala Val Ala Leu Ala Gln
                80                  85                  90

Lys Asn Val Lys Leu Ser Thr Glu Gln Leu Arg Cys Leu Ala His
                95                  100                 105

Arg Leu Ser Glu Pro Pro Glu Asp Leu Asp Ala Leu Pro Leu Asp
                110                 115                 120

Leu Leu Leu Phe Leu Asn Pro Asp Ala Phe Ser Gly Pro Gln Ala
                125                 130                 135

Cys Thr Arg Phe Phe Ser Arg Ile Thr Lys Ala Asn Val Asp Leu
                140                 145                 150

Leu Pro Arg Gly Ala Pro Glu Arg Gln Arg Leu Leu Pro Ala Ala
                155                 160                 165

Leu Ala Cys Trp Gly Val Arg Gly Ser Leu Leu Ser Glu Ala Asp
                170                 175                 180

Val Arg Ala Leu Gly Gly Leu Ala Cys Asp Leu Pro Gly Arg Phe
                185                 190                 195

Val Ala Glu Ser Ala Glu Val Leu Leu Pro Arg Leu Val Ser Cys
                200                 205                 210

Pro Gly Pro Leu Asp Gln Asp Gln Gln Glu Ala Ala Arg Ala Ala
```

215 220 225

Leu Gln Gly Gly Gly Pro Pro Tyr Gly Pro Pro Ser Thr Trp Ser
230 235 240

Val Ser Thr Met Asp Ala Leu Arg Gly Leu Leu Pro Val Leu Gly
245 250 255

Gln Pro Ile Ile Arg Ser Ile Pro Gln Gly Ile Val Ala Ala Trp
260 265 270

Arg Gln Arg Ser Ser Arg Asp Pro Ser Trp Arg Gln Pro Glu Arg
275 280 285

Thr Ile Leu Arg Pro Arg Phe Arg Arg Glu Val Glu Lys Thr Ala
290 295 300

Cys Pro Ser Gly Lys Lys Ala Arg Glu Ile Asp Glu Ser Leu Ile
305 310 315

Phe Tyr Lys Lys Trp Glu Leu Glu Ala Cys Val Asp Ala Ala Leu
320 325 330

Leu Ala Thr Gln Met Asp Arg Val Asn Ala Ile Pro Phe Thr Tyr
335 340 345

Glu Gln Leu Asp Val Leu Lys His Lys Leu Asp Glu Leu Tyr Pro
350 355 360

Gln Gly Tyr Pro Glu Ser Val Ile Gln His Leu Gly Tyr Leu Phe
365 370 375

Leu Lys Met Ser Pro Glu Asp Ile Arg Lys Trp Asn Val Thr Ser
380 385 390

Leu Glu Thr Leu Lys Ala Leu Leu Glu Val Asn Lys Gly His Glu
395 400 405

Met Ser Pro Gln Val Ala Thr Leu Ile Asp Arg Phe Val Lys Gly
410 415 420

Arg Gly Gln Leu Asp Lys Asp Thr Leu Asp Thr Leu Thr Ala Phe
425 430 435

Tyr Pro Gly Tyr Leu Cys Ser Leu Ser Pro Glu Glu Leu Ser Ser
440 445 450

Val Pro Pro Ser Ser Ile Trp Ala Val Arg Pro Gln Asp Leu Asp
455 460 465

Thr Cys Asp Pro Arg Gln Leu Asp Val Leu Tyr Pro Lys Ala Arg
470 475 480

Leu Ala Phe Gln Asn Met Asn Gly Ser Glu Tyr Phe Val Lys Ile
485 490 495

Gln Ser Phe Leu Gly Gly Ala Pro Thr Glu Asp Leu Lys Ala Leu
500 505 510

Ser Gln Gln Asn Val Ser Met Asp Leu Ala Thr Phe Met Lys Leu
515 520 525

Arg Thr Asp Ala Val Leu Pro Leu Thr Val Ala Glu Val Gln Lys
530 535 540

Leu Leu Gly Pro His Val Glu Gly Leu Lys Ala Glu Glu Arg His
545 550 555

Arg Pro Val Arg Asp Trp Ile Leu Arg Gln Arg Gln Asp Asp Leu
560 565 570

Asp Thr Leu Gly Leu Gly Leu Gln Gly Gly Ile Pro Asn Gly Tyr
575 580 585

Leu Val Leu Asp Leu Ser Val Gln Glu Ala Leu Ser Gly Thr Pro
590 595 600

Cys Leu Leu Gly Pro Gly Pro Val Leu Thr Val Leu Ala Leu Leu
605 610 615

Leu Ala Ser Thr Leu Ala
620

<210> 118
<211> 357
<212> PRT
<213> Homo sapiens

<400> 118
Met Ala Thr Thr Val Pro Asp Gly Cys Arg Asn Gly Leu Lys Ser
1 5 10 15

Lys Tyr Tyr Arg Leu Cys Asp Lys Ala Glu Ala Trp Gly Ile Val
20 25 30

Leu Glu Thr Val Ala Thr Ala Gly Val Val Thr Ser Val Ala Phe
35 40 45

Met Leu Thr Leu Pro Ile Leu Val Cys Lys Val Gln Asp Ser Asn
50 55 60

Arg Arg Lys Met Leu Pro Thr Gln Phe Leu Phe Leu Leu Gly Val
65 70 75

Leu Gly Ile Phe Gly Leu Thr Phe Ala Phe Ile Ile Gly Leu Asp
80 85 90

Gly Ser Thr Gly Pro Thr Arg Phe Phe Leu Phe Gly Ile Leu Phe
95 100 105

Ser Ile Cys Phe Ser Cys Leu Leu Ala His Ala Val Ser Leu Thr
110 115 120

Lys Leu Val Arg Gly Arg Lys Pro Leu Ser Leu Leu Val Ile Leu
125 130 135

Gly Leu Ala Val Gly Phe Ser Leu Val Gln Asp Val Ile Ala Ile
140 145 150

Glu Tyr Ile Val Leu Thr Met Asn Arg Thr Asn Val Asn Val Phe
155 160 165

Ser Glu Leu Ser Ala Pro Arg Arg Asn Glu Asp Phe Val Leu Leu
170 175 180

Leu Thr Tyr Val Leu Phe Leu Met Ala Leu Thr Phe Leu Met Ser
185 190 195

Ser Phe Thr Phe Cys Gly Ser Phe Thr Gly Trp Lys Arg His Gly
200 205 210

Ala His Ile Tyr Leu Thr Met Leu Leu Ser Ile Ala Ile Trp Val
215 220 225

Ala Trp Ile Thr Leu Leu Met Leu Pro Asp Phe Asp Arg Arg Trp
230 235 240

Asp Asp Thr Ile Leu Ser Ser Ala Leu Ala Ala Asn Gly Trp Val
245 250 255

Phe Leu Leu Ala Tyr Val Ser Pro Glu Phe Trp Leu Leu Thr Lys
260 265 270

Gln Arg Asn Pro Met Asp Tyr Pro Val Glu Asp Ala Phe Cys Lys
275 280 285

Pro Gln Leu Val Lys Lys Ser Tyr Gly Val Glu Asn Arg Ala Tyr
290 295 300

Ser Gln Glu Glu Ile Thr Gln Gly Phe Glu Glu Thr Gly Asp Thr
305 310 315

Leu Tyr Ala Pro Tyr Ser Thr His Phe Gln Leu Gln Asn Gln Pro
320 325 330

Pro Gln Lys Glu Phe Ser Ile Pro Arg Ala His Ala Trp Pro Ser
335 340 345

Pro Tyr Lys Asp Tyr Glu Val Lys Lys Glu Gly Ser
350 355

<210> 119
<211> 836
<212> PRT
<213> Homo sapiens

<400> 119
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
125 130 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
140 145 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
155 160 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
170 175 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
185 190 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
275 280 285

Val Leu Glu Arg Phe Met Gly Asp Lys Val Ala Ser Glu Ala Leu
290 295 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile

305 310 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
320 325 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
335 340 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
350 355 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
365 370 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
380 385 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
395 400 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
410 415 420

Val Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
425 430 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
440 445 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
455 460 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
470 475 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
485 490 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
500 505 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
515 520 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
530 535 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
545 550 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
560 565 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
575 580 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
590 595 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
605 610 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
620 625 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
635 640 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
650 655 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Thr Thr Lys Ile Ile Thr
665 670 675

Lys Val Val Glu Pro Lys Ile Lys Val Ile Glu Gly Ser Leu Gln
680 685 690

Pro Ile Ile Lys Thr Glu Gly Pro Thr Leu Thr Lys Val Lys Ile
695 700 705

Glu Gly Glu Pro Glu Phe Arg Leu Ile Lys Glu Gly Glu Thr Ile
710 715 720

Thr Glu Val Ile His Gly Glu Pro Ile Ile Lys Lys Tyr Thr Lys
725 730 735

Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys Glu Thr Arg
740 745 750

Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr Thr Arg Ile
755 760 765

Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys Lys Leu Leu
770 775 780

Gln Glu Glu Val Thr Lys Val Thr Lys Phe Ile Glu Gly Gly Asp
785 790 795

Gly His Leu Phe Glu Asp Glu Glu Ile Lys Arg Leu Leu Gln Gly
800 805 810

Asp Thr Pro Val Arg Lys Leu Gln Ala Asn Lys Lys Val Gln Gly
815 820 825

Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser Gln
830 835

<210> 120
<211> 809
<212> PRT
<213> Homo sapiens

<400> 120
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
125 130 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
140 145 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
155 160 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
170 175 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
185 190 195

1255
1254

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
275 280 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
290 295 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
305 310 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
320 325 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
335 340 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
350 355 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
365 370 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
380 385 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
395 400 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
410 415 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
425 430 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
440 445 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
455 460 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
470 475 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
485 490 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
500 505 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
515 520 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
530 535 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
545 550 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
560 565 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu

575 580 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
590 595 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
605 610 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
620 625 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
635 640 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
650 655 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Arg Pro Thr Leu Thr Lys
665 670 675

Val Lys Ile Glu Gly Glu Pro Glu Phe Arg Leu Ile Lys Glu Gly
680 685 690

Glu Thr Ile Thr Glu Val Ile His Gly Glu Pro Ile Ile Lys Lys
695 700 705

Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
710 715 720

Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
725 730 735

Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
740 745 750

Lys Leu Leu Gln Glu Glu Val Thr Lys Val Thr Lys Phe Ile Glu
755 760 765

Gly Gly Asp Gly His Leu Phe Glu Asp Glu Glu Ile Lys Arg Leu
770 775 780

Leu Gln Gly Asp Thr Pro Val Arg Lys Leu Gln Ala Asn Lys Lys
785 790 795

Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser Gln
800 805

<210> 121
<211> 751
<212> PRT
<213> Homo sapiens

<400> 121
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

```
Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
                125                 130                 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
                140                 145                 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
                155                 160                 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
                170                 175                 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
                185                 190                 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
                200                 205                 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
                215                 220                 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
                230                 235                 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
                245                 250                 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
                260                 265                 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
                275                 280                 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
                290                 295                 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
                305                 310                 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
                320                 325                 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
                335                 340                 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
                350                 355                 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
                365                 370                 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
                380                 385                 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
                395                 400                 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
                410                 415                 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
                425                 430                 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
                440                 445                 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
                455                 460                 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
                470                 475                 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
                485                 490                 495
```

1280
1259

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
500 505 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
515 520 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
530 535 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
545 550 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
560 565 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
575 580 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
590 595 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
605 610 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
620 625 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
635 640 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
650 655 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Lys Pro Ile Ile Lys Lys
665 670 675

Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
680 685 690

Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
695 700 705

Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
710 715 720

Lys Leu Leu Gln Glu Asp Thr Pro Val Arg Lys Leu Gln Ala Asn
725 730 735

Lys Lys Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser
740 745 750

Gln

<210> 122
<211> 781
<212> PRT
<213> Homo sapiens

<400> 122
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys

110 115 120

Ile Leu Ile Gly Gly Phe Ile Ser Glu Thr Leu Gly Trp Pro Phe
125 130 135

Val Phe Tyr Ile Phe Gly Gly Val Gly Cys Val Cys Cys Leu Leu
140 145 150

Trp Phe Val Val Ile Tyr Asp Asp Pro Val Ser Tyr Pro Trp Ile
155 160 165

Ser Thr Ser Glu Lys Glu Tyr Ile Ile Ser Ser Leu Lys Gln Gln
170 175 180

Val Arg Ser Ser Lys Gln Pro Leu Pro Ile Lys Ala Met Leu Arg
185 190 195

Ser Leu Pro Ile Trp Ser Ile Cys Leu Gly Cys Phe Ser His Gln
200 205 210

Trp Leu Val Ser Thr Met Val Val Tyr Ile Pro Thr Tyr Ile Ser
215 220 225

Ser Val Tyr His Val Asn Ile Arg Asp Asn Gly Leu Leu Ser Ala
230 235 240

Leu Pro Phe Ile Val Ala Trp Val Ile Gly Met Val Gly Gly Tyr
245 250 255

Leu Ala Asp Phe Leu Leu Thr Lys Lys Phe Arg Leu Ile Thr Val
260 265 270

Arg Lys Ile Ala Thr Ile Leu Gly Ser Leu Pro Ser Ser Ala Leu
275 280 285

Ile Val Ser Leu Pro Tyr Leu Asn Ser Gly Tyr Ile Thr Ala Thr
290 295 300

Ala Leu Leu Thr Leu Ser Cys Gly Leu Ser Thr Leu Cys Gln Ser
305 310 315

Gly Ile Tyr Ile Asn Val Leu Asp Ile Ala Pro Arg Tyr Ser Ser
320 325 330

Phe Leu Met Gly Ala Ser Arg Gly Phe Ser Ser Ile Ala Pro Val
335 340 345

Ile Val Pro Thr Val Ser Gly Phe Leu Leu Ser Gln Asp Pro Glu
350 355 360

Phe Gly Trp Arg Asn Val Phe Phe Leu Leu Phe Ala Val Asn Leu
365 370 375

Leu Gly Leu Leu Phe Tyr Leu Ile Phe Gly Glu Ala Asp Val Gln
380 385 390

Glu Trp Ala Lys Glu Arg Lys Leu Thr Arg Leu
395 400

<210> 114
<211> 507
<212> PRT
<213> Homo sapiens

<400> 114
Met Ala Gly Ala Gly Pro Lys Arg Arg Ala Leu Ala Ala Pro Ala
1 5 10 15

Ala Glu Glu Lys Glu Glu Ala Arg Glu Lys Met Leu Ala Ala Lys
20 25 30

Ser Ala Asp Gly Ser Ala Pro Ala Gly Glu Gly Glu Gly Val Thr
35 40 45

Leu Gln Arg Asn Ile Thr Leu Leu Asn Gly Val Ala Ile Ile Val
50 55 60

```
Gly Thr Ile Ile Gly Ser Gly Ile Phe Val Thr Pro Thr Gly Val
                65                  70                  75

Leu Lys Glu Ala Gly Ser Pro Gly Leu Ala Leu Val Val Trp Ala
                80                  85                  90

Ala Cys Gly Val Phe Ser Ile Val Gly Ala Leu Cys Tyr Ala Glu
                95                 100                 105

Leu Gly Thr Thr Ile Ser Lys Ser Gly Gly Asp Tyr Ala Tyr Met
               110                 115                 120

Leu Glu Val Tyr Gly Ser Leu Pro Ala Phe Leu Lys Leu Trp Ile
               125                 130                 135

Glu Leu Leu Ile Ile Arg Pro Ser Ser Gln Tyr Ile Val Ala Leu
               140                 145                 150

Val Phe Ala Thr Tyr Leu Leu Lys Pro Leu Phe Pro Thr Cys Pro
               155                 160                 165

Val Pro Glu Glu Ala Ala Lys Leu Val Ala Cys Leu Cys Val Leu
               170                 175                 180

Leu Leu Thr Ala Val Asn Cys Tyr Ser Val Lys Ala Ala Thr Arg
               185                 190                 195

Val Gln Asp Ala Phe Ala Ala Ala Lys Leu Leu Ala Leu Ala Leu
               200                 205                 210

Ile Ile Leu Leu Gly Phe Val Gln Ile Gly Lys Gly Asp Val Ser
               215                 220                 225

Asn Leu Asp Pro Lys Phe Ser Phe Glu Gly Thr Lys Leu Asp Val
               230                 235                 240

Gly Asn Ile Val Leu Ala Leu Tyr Ser Gly Leu Phe Ala Tyr Gly
               245                 250                 255

Gly Trp Asn Tyr Leu Asn Phe Val Thr Glu Glu Met Ile Asn Pro
               260                 265                 270

Tyr Arg Asn Leu Pro Leu Ala Ile Ile Ile Ser Leu Pro Ile Val
               275                 280                 285

Thr Leu Val Tyr Val Leu Thr Asn Leu Ala Tyr Phe Thr Thr Leu
               290                 295                 300

Ser Thr Glu Gln Met Leu Ser Ser Glu Ala Val Ala Val Asp Phe
               305                 310                 315

Gly Asn Tyr His Leu Gly Val Met Ser Trp Ile Ile Pro Val Phe
               320                 325                 330

Val Gly Leu Ser Cys Phe Gly Ser Val Asn Gly Ser Leu Phe Thr
               335                 340                 345

Ser Ser Arg Leu Phe Phe Val Gly Ser Arg Glu Gly His Leu Pro
               350                 355                 360

Ser Ile Leu Ser Met Ile His Pro Gln Leu Leu Thr Pro Val Pro
               365                 370                 375

Ser Leu Val Phe Thr Cys Val Met Thr Leu Leu Tyr Ala Phe Ser
               380                 385                 390

Lys Asp Ile Phe Ser Val Ile Asn Phe Phe Ser Phe Phe Asn Trp
               395                 400                 405

Leu Cys Val Ala Leu Ala Ile Ile Gly Met Ile Trp Leu Arg His
               410                 415                 420

Arg Lys Pro Glu Leu Glu Arg Pro Ile Lys Val Asn Leu Ala Leu
               425                 430                 435
```

Pro Val Phe Phe Ile Leu Ala Cys Leu Phe Leu Ile Ala Val Ser
440 445 450

Phe Trp Lys Thr Pro Val Glu Cys Gly Ile Gly Phe Thr Ile Ile
455 460 465

Leu Ser Gly Leu Pro Val Tyr Phe Phe Gly Val Trp Trp Lys Asn
470 475 480

Lys Pro Lys Trp Leu Leu Gln Gly Ile Phe Ser Thr Thr Val Leu
485 490 495

Cys Gln Lys Leu Met Gln Val Val Pro Gln Glu Thr
500 505

<210> 115
<211> 507
<212> PRT
<213> Homo sapiens

<400> 115
Met Ala Gly Ala Gly Pro Lys Arg Arg Ala Leu Ala Ala Pro Ala
1 5 10 15

Ala Glu Glu Lys Glu Glu Ala Arg Glu Lys Met Leu Ala Ala Lys
20 25 30

Ser Ala Asp Gly Ser Ala Pro Ala Gly Glu Gly Glu Gly Val Thr
35 40 45

Leu Gln Arg Asn Ile Thr Leu Leu Asn Gly Val Ala Ile Ile Val
50 55 60

Gly Thr Ile Ile Gly Ser Gly Ile Phe Val Thr Pro Thr Gly Val
65 70 75

Leu Lys Glu Ala Gly Ser Pro Gly Leu Ala Leu Val Val Trp Ala
80 85 90

Ala Cys Gly Val Phe Ser Ile Val Gly Ala Leu Cys Tyr Ala Glu
95 100 105

Leu Gly Thr Thr Ile Ser Lys Ser Gly Gly Asp Tyr Ala Tyr Met
110 115 120

Leu Glu Val Tyr Gly Ser Leu Pro Ala Phe Leu Lys Leu Trp Ile
125 130 135

Glu Leu Leu Ile Ile Arg Pro Ser Ser Gln Tyr Ile Val Ala Leu
140 145 150

Val Phe Ala Thr Tyr Leu Leu Lys Pro Leu Phe Pro Thr Cys Pro
155 160 165

Val Pro Glu Glu Ala Ala Lys Leu Val Ala Cys Leu Cys Val Leu
170 175 180

Leu Leu Thr Ala Val Asn Cys Tyr Ser Val Lys Ala Ala Thr Arg
185 190 195

Val Gln Asp Ala Phe Ala Ala Ala Lys Leu Leu Ala Leu Ala Leu
200 205 210

Ile Ile Leu Leu Gly Phe Val Gln Ile Gly Lys Gly Asp Val Ser
215 220 225

Asn Leu Asp Pro Asn Phe Ser Phe Glu Gly Thr Lys Leu Asp Val
230 235 240

Gly Asn Ile Val Leu Ala Leu Tyr Ser Gly Leu Phe Ala Tyr Gly
245 250 255

Gly Trp Asn Tyr Leu Asn Phe Val Thr Glu Glu Met Ile Asn Pro
260 265 270

Tyr Arg Asn Leu Pro Leu Ala Ile Ile Ile Ser Leu Pro Ile Val

275 280 285

Thr Leu Val Tyr Val Leu Thr Asn Leu Ala Tyr Phe Thr Thr Leu
290 295 300

Ser Thr Glu Gln Met Leu Ser Ser Glu Ala Val Ala Val Asp Phe
305 310 315

Gly Asn Tyr His Leu Gly Val Met Ser Trp Ile Ile Pro Val Phe
320 325 330

Val Gly Leu Ser Cys Phe Gly Ser Val Asn Gly Ser Leu Phe Thr
335 340 345

Ser Ser Arg Leu Phe Phe Val Gly Ser Arg Glu Gly His Leu Pro
350 355 360

Ser Ile Leu Ser Met Ile His Pro Gln Leu Leu Thr Pro Val Pro
365 370 375

Ser Leu Val Phe Thr Cys Val Met Thr Leu Leu Tyr Ala Phe Ser
380 385 390

Lys Asp Ile Phe Ser Val Ile Asn Phe Phe Ser Phe Phe Asn Trp
395 400 405

Leu Cys Val Ala Leu Ala Ile Ile Gly Met Ile Trp Leu Arg His
410 415 420

Arg Lys Pro Glu Leu Glu Arg Pro Ile Lys Val Asn Leu Ala Leu
425 430 435

Pro Val Phe Phe Ile Leu Ala Cys Leu Phe Leu Ile Ala Val Ser
440 445 450

Phe Trp Lys Thr Pro Val Glu Cys Gly Ile Gly Phe Thr Ile Ile
455 460 465

Leu Ser Gly Leu Pro Val Tyr Phe Phe Gly Val Trp Trp Lys Asn
470 475 480

Lys Pro Lys Trp Leu Leu Gln Gly Ile Phe Ser Thr Thr Val Leu
485 490 495

Cys Gln Lys Leu Met Gln Val Val Pro Gln Glu Thr
500 505

<210> 116
<211> 622
<212> PRT
<213> Homo sapiens

<400> 116
Met Ala Leu Pro Thr Ala Arg Pro Leu Leu Gly Ser Cys Gly Thr
1 5 10 15

Pro Ala Leu Gly Ser Leu Leu Phe Leu Leu Phe Ser Leu Gly Trp
20 25 30

Val Gln Pro Ser Arg Thr Leu Ala Gly Glu Thr Gly Gln Glu Ala
35 40 45

Ala Pro Leu Asp Gly Val Leu Ala Asn Pro Pro Asn Ile Ser Ser
50 55 60

Leu Ser Pro Arg Gln Leu Leu Gly Phe Pro Cys Ala Glu Val Ser
65 70 75

Gly Leu Ser Thr Glu Arg Val Arg Glu Leu Ala Val Ala Leu Ala
80 85 90

Gln Lys Asn Val Lys Leu Ser Thr Glu Gln Leu Arg Cys Leu Ala
95 100 105

His Arg Leu Ser Glu Pro Pro Glu Asp Leu Asp Ala Leu Pro Leu
110 115 120

Asp Leu Leu Leu Phe Leu Asn Pro Asp Ala Phe Ser Gly Pro Gln
125 130 135

Ala Cys Thr Arg Phe Phe Ser Arg Ile Thr Lys Ala Asn Val Asp
140 145 150

Leu Leu Pro Arg Gly Ala Pro Glu Arg Gln Arg Leu Leu Pro Ala
155 160 165

Ala Leu Ala Cys Trp Gly Val Arg Gly Ser Leu Leu Ser Glu Ala
170 175 180

Asp Val Arg Ala Leu Gly Gly Leu Ala Cys Asp Leu Pro Gly Arg
185 190 195

Phe Val Ala Glu Ser Ala Glu Val Leu Leu Pro Arg Leu Val Ser
200 205 210

Cys Pro Gly Pro Leu Asp Gln Asp Gln Gln Glu Ala Ala Arg Ala
215 220 225

Ala Leu Gln Gly Gly Gly Pro Pro Tyr Gly Pro Pro Ser Thr Trp
230 235 240

Ser Val Ser Thr Met Asp Ala Leu Arg Gly Leu Leu Pro Val Leu
245 250 255

Gly Gln Pro Ile Ile Arg Ser Ile Pro Gln Gly Ile Val Ala Ala
260 265 270

Trp Arg Gln Arg Ser Ser Arg Asp Pro Ser Trp Arg Gln Pro Glu
275 280 285

Arg Thr Ile Leu Arg Pro Arg Phe Arg Arg Glu Val Glu Lys Thr
290 295 300

Ala Cys Pro Ser Gly Lys Lys Ala Arg Glu Ile Asp Glu Ser Leu
305 310 315

Ile Phe Tyr Lys Lys Trp Glu Leu Glu Ala Cys Val Asp Ala Ala
320 325 330

Leu Leu Ala Thr Gln Met Asp Arg Val Asn Ala Ile Pro Phe Thr
335 340 345

Tyr Glu Gln Leu Asp Val Leu Lys His Lys Leu Asp Glu Leu Tyr
350 355 360

Pro Gln Gly Tyr Pro Glu Ser Val Ile Gln His Leu Gly Tyr Leu
365 370 375

Phe Leu Lys Met Ser Pro Glu Asp Ile Arg Lys Trp Asn Val Thr
380 385 390

Ser Leu Glu Thr Leu Lys Ala Leu Leu Glu Val Asn Lys Gly His
395 400 405

Glu Met Ser Pro Gln Val Ala Thr Leu Ile Asp Arg Phe Val Lys
410 415 420

Gly Arg Gly Gln Leu Asp Lys Asp Thr Leu Asp Thr Leu Thr Ala
425 430 435

Phe Tyr Pro Gly Tyr Leu Cys Ser Leu Ser Pro Glu Glu Leu Ser
440 445 450

Ser Val Pro Pro Ser Ser Ile Trp Ala Val Arg Pro Gln Asp Leu
455 460 465

Asp Thr Cys Asp Pro Arg Gln Leu Asp Val Leu Tyr Pro Lys Ala
470 475 480

Arg Leu Ala Phe Gln Asn Met Asn Gly Ser Glu Tyr Phe Val Lys
485 490 495

1264
1263

Ile Gln Ser Phe Leu Gly Gly Ala Pro Thr Glu Asp Leu Lys Ala
500 505 510

Leu Ser Gln Gln Asn Val Ser Met Asp Leu Ala Thr Phe Met Lys
515 520 525

Leu Arg Thr Asp Ala Val Leu Pro Leu Thr Val Ala Glu Val Gln
530 535 540

Lys Leu Leu Gly Pro His Val Glu Gly Leu Lys Ala Glu Glu Arg
545 550 555

His Arg Pro Val Arg Asp Trp Ile Leu Arg Gln Arg Gln Asp Asp
560 565 570

Leu Asp Thr Leu Gly Leu Gly Leu Gln Gly Gly Ile Pro Asn Gly
575 580 585

Tyr Leu Val Leu Asp Leu Ser Val Gln Glu Ala Leu Ser Gly Thr
590 595 600

Pro Cys Leu Leu Gly Pro Gly Pro Val Leu Thr Val Leu Ala Leu
605 610 615

Leu Leu Ala Ser Thr Leu Ala
620

<210> 117
<211> 621
<212> PRT
<213> Homo sapiens

<400> 117
Ala Leu Pro Thr Ala Arg Pro Leu Leu Gly Ser Cys Gly Thr Pro
1 5 10 15

Ala Leu Gly Ser Leu Leu Phe Leu Leu Phe Ser Leu Gly Trp Val
20 25 30

Gln Pro Ser Arg Thr Leu Ala Gly Glu Thr Gly Gln Glu Ala Ala
35 40 45

Pro Leu Asp Gly Val Leu Ala Asn Pro Pro Asn Ile Ser Ser Leu
50 55 60

Ser Pro Arg Gln Leu Leu Gly Phe Pro Cys Ala Glu Val Ser Gly
65 70 75

Leu Ser Thr Glu Arg Val Arg Glu Leu Ala Val Ala Leu Ala Gln
80 85 90

Lys Asn Val Lys Leu Ser Thr Glu Gln Leu Arg Cys Leu Ala His
95 100 105

Arg Leu Ser Glu Pro Pro Glu Asp Leu Asp Ala Leu Pro Leu Asp
110 115 120

Leu Leu Leu Phe Leu Asn Pro Asp Ala Phe Ser Gly Pro Gln Ala
125 130 135

Cys Thr Arg Phe Phe Ser Arg Ile Thr Lys Ala Asn Val Asp Leu
140 145 150

Leu Pro Arg Gly Ala Pro Glu Arg Gln Arg Leu Leu Pro Ala Ala
155 160 165

Leu Ala Cys Trp Gly Val Arg Gly Ser Leu Leu Ser Glu Ala Asp
170 175 180

Val Arg Ala Leu Gly Gly Leu Ala Cys Asp Leu Pro Gly Arg Phe
185 190 195

Val Ala Glu Ser Ala Glu Val Leu Leu Pro Arg Leu Val Ser Cys
200 205 210

Pro Gly Pro Leu Asp Gln Asp Gln Gln Glu Ala Ala Arg Ala Ala

1265
1264

215 220 225

Leu Gln Gly Gly Gly Pro Pro Tyr Gly Pro Pro Ser Thr Trp Ser
230 235 240

Val Ser Thr Met Asp Ala Leu Arg Gly Leu Leu Pro Val Leu Gly
245 250 255

Gln Pro Ile Ile Arg Ser Ile Pro Gln Gly Ile Val Ala Ala Trp
260 265 270

Arg Gln Arg Ser Ser Arg Asp Pro Ser Trp Arg Gln Pro Glu Arg
275 280 285

Thr Ile Leu Arg Pro Arg Phe Arg Arg Glu Val Glu Lys Thr Ala
290 295 300

Cys Pro Ser Gly Lys Lys Ala Arg Glu Ile Asp Glu Ser Leu Ile
305 310 315

Phe Tyr Lys Lys Trp Glu Leu Glu Ala Cys Val Asp Ala Ala Leu
320 325 330

Leu Ala Thr Gln Met Asp Arg Val Asn Ala Ile Pro Phe Thr Tyr
335 340 345

Glu Gln Leu Asp Val Leu Lys His Lys Leu Asp Glu Leu Tyr Pro
350 355 360

Gln Gly Tyr Pro Glu Ser Val Ile Gln His Leu Gly Tyr Leu Phe
365 370 375

Leu Lys Met Ser Pro Glu Asp Ile Arg Lys Trp Asn Val Thr Ser
380 385 390

Leu Glu Thr Leu Lys Ala Leu Leu Glu Val Asn Lys Gly His Glu
395 400 405

Met Ser Pro Gln Val Ala Thr Leu Ile Asp Arg Phe Val Lys Gly
410 415 420

Arg Gly Gln Leu Asp Lys Asp Thr Leu Asp Thr Leu Thr Ala Phe
425 430 435

Tyr Pro Gly Tyr Leu Cys Ser Leu Ser Pro Glu Glu Leu Ser Ser
440 445 450

Val Pro Pro Ser Ser Ile Trp Ala Val Arg Pro Gln Asp Leu Asp
455 460 465

Thr Cys Asp Pro Arg Gln Leu Asp Val Leu Tyr Pro Lys Ala Arg
470 475 480

Leu Ala Phe Gln Asn Met Asn Gly Ser Glu Tyr Phe Val Lys Ile
485 490 495

Gln Ser Phe Leu Gly Gly Ala Pro Thr Glu Asp Leu Lys Ala Leu
500 505 510

Ser Gln Gln Asn Val Ser Met Asp Leu Ala Thr Phe Met Lys Leu
515 520 525

Arg Thr Asp Ala Val Leu Pro Leu Thr Val Ala Glu Val Gln Lys
530 535 540

Leu Leu Gly Pro His Val Glu Gly Leu Lys Ala Glu Glu Arg His
545 550 555

Arg Pro Val Arg Asp Trp Ile Leu Arg Gln Arg Gln Asp Asp Leu
560 565 570

Asp Thr Leu Gly Leu Gly Leu Gln Gly Gly Ile Pro Asn Gly Tyr
575 580 585

Leu Val Leu Asp Leu Ser Val Gln Glu Ala Leu Ser Gly Thr Pro
590 595 600

Cys Leu Leu Gly Pro Gly Pro Val Leu Thr Val Leu Ala Leu Leu
605 610 615

Leu Ala Ser Thr Leu Ala
620

<210> 118
<211> 357
<212> PRT
<213> Homo sapiens

<400> 118
Met Ala Thr Thr Val Pro Asp Gly Cys Arg Asn Gly Leu Lys Ser
1 5 10 15

Lys Tyr Tyr Arg Leu Cys Asp Lys Ala Glu Ala Trp Gly Ile Val
20 25 30

Leu Glu Thr Val Ala Thr Ala Gly Val Val Thr Ser Val Ala Phe
35 40 45

Met Leu Thr Leu Pro Ile Leu Val Cys Lys Val Gln Asp Ser Asn
50 55 60

Arg Arg Lys Met Leu Pro Thr Gln Phe Leu Phe Leu Leu Gly Val
65 70 75

Leu Gly Ile Phe Gly Leu Thr Phe Ala Phe Ile Ile Gly Leu Asp
80 85 90

Gly Ser Thr Gly Pro Thr Arg Phe Phe Leu Phe Gly Ile Leu Phe
95 100 105

Ser Ile Cys Phe Ser Cys Leu Leu Ala His Ala Val Ser Leu Thr
110 115 120

Lys Leu Val Arg Gly Arg Lys Pro Leu Ser Leu Leu Val Ile Leu
125 130 135

Gly Leu Ala Val Gly Phe Ser Leu Val Gln Asp Val Ile Ala Ile
140 145 150

Glu Tyr Ile Val Leu Thr Met Asn Arg Thr Asn Val Asn Val Phe
155 160 165

Ser Glu Leu Ser Ala Pro Arg Arg Asn Glu Asp Phe Val Leu Leu
170 175 180

Leu Thr Tyr Val Leu Phe Leu Met Ala Leu Thr Phe Leu Met Ser
185 190 195

Ser Phe Thr Phe Cys Gly Ser Phe Thr Gly Trp Lys Arg His Gly
200 205 210

Ala His Ile Tyr Leu Thr Met Leu Leu Ser Ile Ala Ile Trp Val
215 220 225

Ala Trp Ile Thr Leu Leu Met Leu Pro Asp Phe Asp Arg Arg Trp
230 235 240

Asp Asp Thr Ile Leu Ser Ser Ala Leu Ala Ala Asn Gly Trp Val
245 250 255

Phe Leu Leu Ala Tyr Val Ser Pro Glu Phe Trp Leu Leu Thr Lys
260 265 270

Gln Arg Asn Pro Met Asp Tyr Pro Val Glu Asp Ala Phe Cys Lys
275 280 285

Pro Gln Leu Val Lys Lys Ser Tyr Gly Val Glu Asn Arg Ala Tyr
290 295 300

Ser Gln Glu Glu Ile Thr Gln Gly Phe Glu Glu Thr Gly Asp Thr
305 310 315

Leu Tyr Ala Pro Tyr Ser Thr His Phe Gln Leu Gln Asn Gln Pro
320 325 330

Pro Gln Lys Glu Phe Ser Ile Pro Arg Ala His Ala Trp Pro Ser
335 340 345

Pro Tyr Lys Asp Tyr Glu Val Lys Lys Glu Gly Ser
350 355

<210> 119
<211> 836
<212> PRT
<213> Homo sapiens

<400> 119
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
125 130 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
140 145 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
155 160 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
170 175 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
185 190 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
275 280 285

Val Leu Glu Arg Phe Met Gly Asp Lys Val Ala Ser Glu Ala Leu
290 295 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile

1268
1767

305 310 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
320 325 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
335 340 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
350 355 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
365 370 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
380 385 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
395 400 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
410 415 420

Val Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
425 430 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
440 445 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
455 460 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
470 475 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
485 490 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
500 505 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
515 520 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
530 535 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
545 550 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
560 565 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
575 580 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
590 595 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
605 610 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
620 625 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
635 640 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
650 655 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Thr Thr Lys Ile Ile Thr
665 670 675

Lys Val Val Glu Pro Lys Ile Lys Val Ile Glu Gly Ser Leu Gln
680 685 690

Pro Ile Ile Lys Thr Glu Gly Pro Thr Leu Thr Lys Val Lys Ile
695 700 705

Glu Gly Glu Pro Glu Phe Arg Leu Ile Lys Glu Gly Glu Thr Ile
710 715 720

Thr Glu Val Ile His Gly Glu Pro Ile Ile Lys Lys Tyr Thr Lys
725 730 735

Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys Glu Thr Arg
740 745 750

Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr Thr Arg Ile
755 760 765

Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys Lys Leu Leu
770 775 780

Gln Glu Glu Val Thr Lys Val Thr Lys Phe Ile Glu Gly Gly Asp
785 790 795

Gly His Leu Phe Glu Asp Glu Glu Ile Lys Arg Leu Leu Gln Gly
800 805 810

Asp Thr Pro Val Arg Lys Leu Gln Ala Asn Lys Lys Val Gln Gly
815 820 825

Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser Gln
830 835

<210> 120
<211> 809
<212> PRT
<213> Homo sapiens

<400> 120
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
125 130 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
140 145 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
155 160 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
170 175 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
185 190 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
275 280 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
290 295 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
305 310 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
320 325 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
335 340 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
350 355 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
365 370 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
380 385 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
395 400 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
410 415 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
425 430 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
440 445 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
455 460 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
470 475 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
485 490 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
500 505 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
515 520 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
530 535 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
545 550 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
560 565 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu

575 580 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
590 595 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
605 610 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
620 625 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
635 640 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
650 655 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Arg Pro Thr Leu Thr Lys
665 670 675

Val Lys Ile Glu Gly Glu Pro Glu Phe Arg Leu Ile Lys Glu Gly
680 685 690

Glu Thr Ile Thr Glu Val Ile His Gly Glu Pro Ile Ile Lys Lys
695 700 705

Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
710 715 720

Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
725 730 735

Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
740 745 750

Lys Leu Leu Gln Glu Glu Val Thr Lys Val Thr Lys Phe Ile Glu
755 760 765

Gly Gly Asp Gly His Leu Phe Glu Asp Glu Glu Ile Lys Arg Leu
770 775 780

Leu Gln Gly Asp Thr Pro Val Arg Lys Leu Gln Ala Asn Lys Lys
785 790 795

Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser Gln
800 805

<210> 121
<211> 751
<212> PRT
<213> Homo sapiens

<400> 121
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
125 130 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
140 145 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
155 160 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
170 175 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
185 190 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
275 280 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
290 295 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
305 310 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
320 325 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
335 340 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
350 355 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
365 370 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
380 385 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
395 400 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
410 415 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
425 430 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
440 445 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
455 460 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
470 475 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
485 490 495

1273
1272

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
500 505 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
515 520 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
530 535 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
545 550 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
560 565 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
575 580 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
590 595 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
605 610 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
620 625 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
635 640 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
650 655 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Lys Pro Ile Ile Lys Lys
665 670 675

Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
680 685 690

Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
695 700 705

Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
710 715 720

Lys Leu Leu Gln Glu Asp Thr Pro Val Arg Lys Leu Gln Ala Asn
725 730 735

Lys Lys Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser
740 745 750

Gln

<210> 122
<211> 781
<212> PRT
<213> Homo sapiens

<400> 122
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys

1273

80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
125 130 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
140 145 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
155 160 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
170 175 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
185 190 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
275 280 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
290 295 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
305 310 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
320 325 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
335 340 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
350 355 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
365 370 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
380 385 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
395 400 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
410 415 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
425 430 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
440 445 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
455 460 465

```
Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
                470                 475                 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
                485                 490                 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
                500                 505                 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
                515                 520                 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
                530                 535                 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
                545                 550                 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
                560                 565                 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
                575                 580                 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
                590                 595                 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
                605                 610                 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
                620                 625                 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
                635                 640                 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
                650                 655                 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Arg Pro Thr Leu Thr Lys
                665                 670                 675

Val Lys Ile Glu Gly Glu Pro Glu Phe Arg Leu Ile Lys Glu Gly
                680                 685                 690

Glu Thr Ile Thr Glu Val Ile His Gly Glu Pro Ile Ile Lys Lys
                695                 700                 705

Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
                710                 715                 720

Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
                725                 730                 735

Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
                740                 745                 750

Lys Leu Leu Gln Glu Asp Thr Pro Val Arg Lys Leu Gln Ala Asn
                755                 760                 765

Lys Lys Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser
                770                 775                 780

Gln
```

<210> 123
<211> 779
<212> PRT
<213> Homo sapiens

<400> 123

```
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
  1               5                  10                  15
```

1275

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
125 130 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
140 145 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
155 160 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
170 175 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
185 190 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
275 280 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
290 295 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
305 310 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
320 325 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
335 340 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
350 355 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
365 370 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
380 385 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu

395 400 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
410 415 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
425 430 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
440 445 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
455 460 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
470 475 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
485 490 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
500 505 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
515 520 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
530 535 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
545 550 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
560 565 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
575 580 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
590 595 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
605 610 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
620 625 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
635 640 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
650 655 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Lys Pro Ile Ile Lys Lys
665 670 675

Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
680 685 690

Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
695 700 705

Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
710 715 720

Lys Leu Leu Gln Glu Glu Val Thr Lys Val Thr Lys Phe Ile Glu
725 730 735

Gly Gly Asp Gly His Leu Phe Glu Asp Glu Glu Ile Lys Arg Leu
740 745 750

Leu Gln Gly Asp Thr Pro Val Arg Lys Leu Gln Ala Asn Lys Lys
755 760 765

Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser Gln
770 775

<210> 124
<211> 283
<212> PRT
<213> Homo sapiens

<400> 124
Met Val Asn Tyr Ala Trp Ala Gly Arg Ser Gln Arg Lys Leu Trp
1 5 10 15

Trp Arg Ser Val Ala Val Leu Thr Cys Lys Ser Val Val Arg Pro
20 25 30

Gly Tyr Arg Gly Gly Leu Gln Ala Arg Arg Ser Thr Leu Leu Lys
35 40 45

Thr Cys Ala Arg Ala Arg Ala Thr Ala Pro Gly Ala Met Lys Met
50 55 60

Val Ala Pro Trp Thr Arg Phe Tyr Ser Asn Ser Cys Cys Leu Cys
65 70 75

Cys His Val Arg Thr Gly Thr Ile Leu Leu Gly Val Trp Tyr Leu
80 85 90

Ile Ile Asn Ala Val Val Leu Leu Ile Leu Leu Ser Ala Leu Ala
95 100 105

Asp Pro Asp Gln Tyr Asn Phe Ser Ser Ser Glu Leu Gly Gly Asp
110 115 120

Phe Glu Phe Met Asp Asp Ala Asn Met Cys Ile Ala Ile Ala Ile
125 130 135

Ser Leu Leu Met Ile Leu Ile Cys Ala Met Ala Thr Tyr Gly Ala
140 145 150

Tyr Lys Gln Arg Ala Ala Trp Ile Ile Pro Phe Phe Cys Tyr Gln
155 160 165

Ile Phe Asp Phe Ala Leu Asn Met Leu Val Ala Ile Thr Val Leu
170 175 180

Ile Tyr Pro Asn Ser Ile Gln Glu Tyr Ile Arg Gln Leu Pro Pro
185 190 195

Asn Phe Pro Tyr Arg Asp Asp Val Met Ser Val Asn Pro Thr Cys
200 205 210

Leu Val Leu Ile Ile Leu Leu Phe Ile Ser Ile Ile Leu Thr Phe
215 220 225

Lys Gly Tyr Leu Ile Ser Cys Val Trp Asn Cys Tyr Arg Tyr Ile
230 235 240

Asn Gly Arg Asn Ser Ser Asp Val Leu Val Tyr Val Thr Ser Asn
245 250 255

Asp Thr Thr Val Leu Leu Pro Pro Tyr Asp Asp Ala Thr Val Asn
260 265 270

Gly Ala Ala Lys Glu Pro Pro Pro Pro Tyr Val Ser Ala
275 280

<210> 125
<211> 254
<212> PRT
<213> Homo sapiens

<400> 125
Met Pro Gly Gln Gly Gly Leu Gln Ala Arg Arg Ser Thr Leu Leu
1 5 10 15

Lys Thr Cys Ala Arg Ala Arg Ala Thr Ala Pro Gly Ala Met Lys
20 25 30

Met Val Ala Pro Trp Thr Arg Phe Tyr Ser Asn Ser Cys Cys Leu
35 40 45

Cys Cys His Val Arg Thr Gly Thr Ile Leu Leu Gly Val Trp Tyr
50 55 60

Leu Ile Ile Asn Ala Val Val Leu Leu Ile Leu Leu Ser Ala Leu
65 70 75

Ala Asp Pro Asp Gln Tyr Asn Phe Ser Ser Ser Glu Leu Gly Gly
80 85 90

Asp Phe Glu Phe Met Asp Asp Ala Asn Met Cys Ile Ala Ile Ala
95 100 105

Ile Ser Leu Leu Met Ile Leu Ile Cys Ala Met Ala Thr Tyr Gly
110 115 120

Ala Tyr Lys Gln Arg Ala Ala Trp Ile Ile Pro Phe Phe Cys Tyr
125 130 135

Gln Ile Phe Asp Phe Ala Leu Asn Met Leu Val Ala Ile Thr Val
140 145 150

Leu Ile Tyr Pro Asn Ser Ile Gln Glu Tyr Ile Arg Gln Leu Pro
155 160 165

Pro Asn Phe Pro Tyr Arg Asp Asp Val Met Ser Val Asn Pro Thr
170 175 180

Cys Leu Val Leu Ile Ile Leu Leu Phe Ile Ser Ile Ile Leu Thr
185 190 195

Phe Lys Gly Tyr Leu Ile Ser Cys Val Trp Asn Cys Tyr Arg Tyr
200 205 210

Ile Asn Gly Arg Asn Ser Ser Asp Val Leu Val Tyr Val Thr Ser
215 220 225

Asn Asp Thr Thr Val Leu Leu Pro Pro Tyr Asp Asp Ala Thr Val
230 235 240

Asn Gly Ala Ala Lys Glu Pro Pro Pro Pro Tyr Val Ser Ala
245 250

<210> 126
<211> 539
<212> PRT
<213> Homo sapiens

<400> 126
Met Gly Arg Ala Gly Gly Gly Gly Pro Gly Arg Gly Pro Pro Pro
1 5 10 15

Leu Leu Leu Phe Leu Gly Ala Ala Leu Val Leu Ala Ser Gly Ala
20 25 30

Val Pro Ala Arg Glu Ala Gly Ser Ala Val Glu Ala Glu Glu Leu
35 40 45

Val Lys Gly Ser Pro Ala Trp Glu Pro Pro Ala Asn Asp Thr Arg
50 55 60

Glu Glu Ala Gly Pro Pro Ala Ala Gly Glu Asp Glu Ala Ser Trp
65 70 75

Thr Ala Pro Gly Gly Glu Leu Ala Gly Pro Glu Glu Val Leu Gln
80 85 90

Glu Ser Ala Ala Val Thr Gly Thr Ala Trp Leu Glu Ala Asp Ser
95 100 105

Pro Gly Leu Gly Gly Val Thr Ala Glu Ala Gly Ser Gly Asp Ala
110 115 120

Gln Ala Leu Pro Ala Thr Leu Gln Ala Pro His Glu Val Leu Gly

```
                125                 130                 135

Gln Ser Ile Met Pro Pro Ala Ile Pro Glu Ala Thr Glu Ala Ser
                140                 145                 150

Gly Pro Pro Ser Pro Thr Pro Gly Asp Lys Leu Ser Pro Ala Ser
                155                 160                 165

Glu Leu Pro Lys Glu Ser Pro Leu Glu Val Trp Leu Asn Leu Gly
                170                 175                 180

Gly Ser Thr Pro Asp Pro Gln Val Pro Glu Leu Thr Tyr Pro Phe
                185                 190                 195

Gln Gly Thr Leu Glu Pro Gln Pro Ala Ser Asp Ile Ile Asp Ile
                200                 205                 210

Asp Tyr Phe Glu Gly Leu Asp Gly Glu Gly Arg Gly Ala Asp Leu
                215                 220                 225

Gly Ser Phe Pro Gly Ser Pro Gly Thr Ser Glu Asn His Pro Asp
                230                 235                 240

Thr Glu Gly Glu Thr Pro Ser Trp Ser Leu Leu Asp Leu Tyr Asp
                245                 250                 255

Asp Phe Thr Pro Phe Asp Glu Ser Asp Phe Tyr Pro Thr Thr Ser
                260                 265                 270

Phe Tyr Asp Asp Leu Asp Glu Glu Glu Glu Glu Glu Glu Asp Asp
                275                 280                 285

Lys Asp Ala Val Gly Gly Gly Asp Leu Glu Asp Glu Asn Glu Leu
                290                 295                 300

Leu Val Pro Thr Gly Lys Pro Gly Leu Gly Pro Gly Thr Gly Gln
                305                 310                 315

Pro Thr Ser Arg Trp His Ala Val Pro Pro Gln His Thr Leu Gly
                320                 325                 330

Ser Val Pro Gly Ser Ser Ile Ala Leu Arg Pro Arg Pro Gly Glu
                335                 340                 345

Pro Gly Arg Asp Leu Ala Ser Ser Glu Asn Gly Thr Glu Cys Arg
                350                 355                 360

Ser Gly Phe Val Arg His Asn Gly Ser Cys Arg Ser Val Cys Asp
                365                 370                 375

Leu Phe Pro Ser Tyr Cys His Asn Gly Gly Gln Cys Tyr Leu Val
                380                 385                 390

Glu Asn Ile Gly Ala Phe Cys Arg Cys Asn Thr Gln Asp Tyr Ile
                395                 400                 405

Trp His Lys Gly Met Arg Cys Glu Ser Ile Ile Thr Asp Phe Gln
                410                 415                 420

Val Met Cys Val Ala Val Gly Ser Ala Ala Leu Val Leu Leu Leu
                425                 430                 435

Leu Phe Met Met Thr Val Phe Phe Ala Lys Lys Leu Tyr Leu Leu
                440                 445                 450

Lys Thr Glu Asn Thr Lys Leu Arg Arg Thr Asn Lys Phe Arg Thr
                455                 460                 465

Pro Ser Glu Leu His Asn Asp Asn Phe Ser Leu Ser Thr Ile Ala
                470                 475                 480

Glu Gly Ser His Pro Asn Asp Asp Pro Ser Ala Pro His Lys Ile
                485                 490                 495

Gln Glu Val Leu Lys Ser Cys Leu Lys Glu Glu Glu Ser Phe Asn
                500                 505                 510
```

1280

Ile Gln Asn Ser Met Ser Pro Lys Leu Glu Gly Gly Lys Gly Asp
515 520 525

Gln Ala Asp Leu Asp Val Asn Cys Leu Gln Asn Asn Leu Thr
530 535

<210> 127
<211> 1325
<212> PRT
<213> Homo sapiens

<400> 127
Met Leu Pro Val Tyr Gln Glu Val Lys Pro Asn Pro Leu Gln Asp
1 5 10 15

Ala Asn Ile Cys Ser Arg Val Phe Phe Trp Trp Leu Asn Pro Leu
20 25 30

Phe Lys Ile Gly His Lys Arg Arg Leu Glu Glu Asp Asp Met Tyr
35 40 45

Ser Val Leu Pro Glu Asp Arg Ser Gln His Leu Gly Glu Glu Leu
50 55 60

Gln Gly Phe Trp Asp Lys Glu Val Leu Arg Ala Glu Asn Asp Ala
65 70 75

Gln Lys Pro Ser Leu Thr Arg Ala Ile Ile Lys Cys Tyr Trp Lys
80 85 90

Ser Tyr Leu Val Leu Gly Ile Phe Thr Leu Ile Glu Glu Ser Ala
95 100 105

Lys Val Ile Gln Pro Ile Phe Leu Gly Lys Ile Ile Asn Tyr Phe
110 115 120

Glu Asn Tyr Asp Pro Met Asp Ser Val Ala Leu Asn Thr Ala Tyr
125 130 135

Ala Tyr Ala Thr Val Leu Thr Phe Cys Thr Leu Ile Leu Ala Ile
140 145 150

Leu His His Leu Tyr Phe Tyr His Val Gln Cys Ala Gly Met Arg
155 160 165

Leu Arg Val Ala Met Cys His Met Ile Tyr Arg Lys Ala Leu Arg
170 175 180

Leu Ser Asn Met Ala Met Gly Lys Thr Thr Thr Gly Gln Ile Val
185 190 195

Asn Leu Leu Ser Asn Asp Val Asn Lys Phe Asp Gln Val Thr Val
200 205 210

Phe Leu His Phe Leu Trp Ala Gly Pro Leu Gln Ala Ile Ala Val
215 220 225

Thr Ala Leu Leu Trp Met Glu Ile Gly Ile Ser Cys Leu Ala Gly
230 235 240

Met Ala Val Leu Ile Ile Leu Leu Pro Leu Gln Ser Cys Phe Gly
245 250 255

Lys Leu Phe Ser Ser Leu Arg Ser Lys Thr Ala Thr Phe Thr Asp
260 265 270

Ala Arg Ile Arg Thr Met Asn Glu Val Ile Thr Gly Ile Arg Ile
275 280 285

Ile Lys Met Tyr Ala Trp Glu Lys Ser Phe Ser Asn Leu Ile Thr
290 295 300

Asn Leu Arg Lys Lys Glu Ile Ser Lys Ile Leu Arg Ser Ser Cys
305 310 315

1280
1281

Leu Arg Gly Met Asn Leu Ala Ser Phe Phe Ser Ala Ser Lys Ile
320 325 330

Ile Val Phe Val Thr Phe Thr Thr Tyr Val Leu Leu Gly Ser Val
335 340 345

Ile Thr Ala Ser Arg Val Phe Val Ala Val Thr Leu Tyr Gly Ala
350 355 360

Val Arg Leu Thr Val Thr Leu Phe Phe Pro Ser Ala Ile Glu Arg
365 370 375

Val Ser Glu Ala Ile Val Ser Ile Arg Arg Ile Gln Thr Phe Leu
380 385 390

Leu Leu Asp Glu Ile Ser Gln Arg Asn Arg Gln Leu Pro Ser Asp
395 400 405

Gly Lys Lys Met Val His Val Gln Asp Phe Thr Ala Phe Trp Asp
410 415 420

Lys Ala Ser Glu Thr Pro Thr Leu Gln Gly Leu Ser Phe Thr Val
425 430 435

Arg Pro Gly Glu Leu Leu Ala Val Val Gly Pro Val Gly Ala Gly
440 445 450

Lys Ser Ser Leu Leu Ser Ala Val Leu Gly Glu Leu Ala Pro Ser
455 460 465

His Gly Leu Val Ser Val His Gly Arg Ile Ala Tyr Val Ser Gln
470 475 480

Gln Pro Trp Val Phe Ser Gly Thr Leu Arg Ser Asn Ile Leu Phe
485 490 495

Gly Lys Lys Tyr Glu Lys Glu Arg Tyr Glu Lys Val Ile Lys Ala
500 505 510

Cys Ala Leu Lys Lys Asp Leu Gln Leu Leu Glu Asp Gly Asp Leu
515 520 525

Thr Val Ile Gly Asp Arg Gly Thr Thr Leu Ser Gly Gly Gln Lys
530 535 540

Ala Arg Val Asn Leu Ala Arg Ala Val Tyr Gln Asp Ala Asp Ile
545 550 555

Tyr Leu Leu Asp Asp Pro Leu Ser Ala Val Asp Ala Glu Val Ser
560 565 570

Arg His Leu Phe Glu Leu Cys Ile Cys Gln Ile Leu His Glu Lys
575 580 585

Ile Thr Ile Leu Val Thr His Gln Leu Gln Tyr Leu Lys Ala Ala
590 595 600

Ser Gln Ile Leu Ile Leu Lys Asp Gly Lys Met Val Gln Lys Gly
605 610 615

Thr Tyr Thr Glu Phe Leu Lys Ser Gly Ile Asp Phe Gly Ser Leu
620 625 630

Leu Lys Lys Asp Asn Glu Glu Ser Glu Gln Pro Pro Val Pro Gly
635 640 645

Thr Pro Thr Leu Arg Asn Arg Thr Phe Ser Glu Ser Ser Val Trp
650 655 660

Ser Gln Gln Ser Ser Arg Pro Ser Leu Lys Asp Gly Ala Leu Glu
665 670 675

Ser Gln Asp Thr Glu Asn Val Pro Val Thr Leu Ser Glu Glu Asn
680 685 690

Arg Ser Glu Gly Lys Val Gly Phe Gln Ala Tyr Lys Asn Tyr Phe

695 700 705

Arg Ala Gly Ala His Trp Ile Val Phe Ile Phe Leu Ile Leu Leu
710 715 720

Asn Thr Ala Ala Gln Val Ala Tyr Val Leu Gln Asp Trp Trp Leu
725 730 735

Ser Tyr Trp Ala Asn Lys Gln Ser Met Leu Asn Val Thr Val Asn
740 745 750

Gly Gly Gly Asn Val Thr Glu Lys Leu Asp Leu Asn Trp Tyr Leu
755 760 765

Gly Ile Tyr Ser Gly Leu Thr Val Ala Thr Val Leu Phe Gly Ile
770 775 780

Ala Arg Ser Leu Leu Val Phe Tyr Val Leu Val Asn Ser Ser Gln
785 790 795

Thr Leu His Asn Lys Met Phe Glu Ser Ile Leu Lys Ala Pro Val
800 805 810

Leu Phe Phe Asp Arg Asn Pro Ile Gly Arg Ile Leu Asn Arg Phe
815 820 825

Ser Lys Asp Ile Gly His Leu Asp Asp Leu Leu Pro Leu Thr Phe
830 835 840

Leu Asp Phe Ile Gln Thr Leu Leu Gln Val Val Gly Val Val Ser
845 850 855

Val Ala Val Ala Val Ile Pro Trp Ile Ala Ile Pro Leu Val Pro
860 865 870

Leu Gly Ile Ile Phe Ile Phe Leu Arg Arg Tyr Phe Leu Glu Thr
875 880 885

Ser Arg Asp Val Lys Arg Leu Glu Ser Thr Thr Arg Ser Pro Val
890 895 900

Phe Ser His Leu Ser Ser Ser Leu Gln Gly Leu Trp Thr Ile Arg
905 910 915

Ala Tyr Lys Ala Glu Glu Arg Cys Gln Glu Leu Phe Asp Ala His
920 925 930

Gln Asp Leu His Ser Glu Ala Trp Phe Leu Phe Leu Thr Thr Ser
935 940 945

Arg Trp Phe Ala Val Arg Leu Asp Ala Ile Cys Ala Met Phe Val
950 955 960

Ile Ile Val Ala Phe Gly Ser Leu Ile Leu Ala Lys Thr Leu Asp
965 970 975

Ala Gly Gln Val Gly Leu Ala Leu Ser Tyr Ala Leu Thr Leu Met
980 985 990

Gly Met Phe Gln Trp Cys Val Arg Gln Ser Ala Glu Val Glu Asn
995 1000 1005

Met Met Ile Ser Val Glu Arg Val Ile Glu Tyr Thr Asp Leu Glu
1010 1015 1020

Lys Glu Ala Pro Trp Glu Tyr Gln Lys Arg Pro Pro Pro Ala Trp
1025 1030 1035

Pro His Glu Gly Val Ile Ile Phe Asp Asn Val Asn Phe Met Tyr
1040 1045 1050

Ser Pro Gly Gly Pro Leu Val Leu Lys His Leu Thr Ala Leu Ile
1055 1060 1065

Lys Ser Gln Glu Lys Val Gly Ile Val Gly Arg Thr Gly Ala Gly
1070 1075 1080

1287
1283

Lys Ser Ser Leu Ile Ser Ala Leu Phe Arg Leu Ser Glu Pro Glu
1085 1090 1095

Gly Lys Ile Trp Ile Asp Lys Ile Leu Thr Thr Glu Ile Gly Leu
1100 1105 1110

His Asp Leu Arg Lys Lys Met Ser Ile Ile Pro Gln Glu Pro Val
1115 1120 1125

Leu Phe Thr Gly Thr Met Arg Lys Asn Leu Asp Pro Phe Lys Glu
1130 1135 1140

His Thr Asp Glu Glu Leu Trp Asn Ala Leu Gln Glu Val Gln Leu
1145 1150 1155

Lys Glu Thr Ile Glu Asp Leu Pro Gly Lys Met Asp Thr Glu Leu
1160 1165 1170

Ala Glu Ser Gly Ser Asn Phe Ser Val Gly Gln Arg Gln Leu Val
1175 1180 1185

Cys Leu Ala Arg Ala Ile Leu Arg Lys Asn Gln Ile Leu Ile Ile
1190 1195 1200

Asp Glu Ala Thr Ala Asn Val Asp Pro Arg Thr Asp Glu Leu Ile
1205 1210 1215

Gln Lys Lys Ile Arg Glu Lys Phe Ala His Cys Thr Val Leu Thr
1220 1225 1230

Ile Ala His Arg Leu Asn Thr Ile Ile Asp Ser Asp Lys Ile Met
1235 1240 1245

Val Leu Asp Ser Gly Arg Leu Lys Glu Tyr Asp Glu Pro Tyr Val
1250 1255 1260

Leu Leu Gln Asn Lys Glu Ser Leu Phe Tyr Lys Met Val Gln Gln
1265 1270 1275

Leu Gly Lys Ala Glu Ala Ala Ala Leu Thr Glu Thr Ala Lys Gln
1280 1285 1290

Val Tyr Phe Lys Arg Asn Tyr Pro His Ile Gly His Thr Asp His
1295 1300 1305

Met Val Thr Asn Thr Ser Asn Gly Gln Pro Ser Thr Leu Thr Ile
1310 1315 1320

Phe Glu Thr Ala Leu
1325

<210> 128
<211> 501
<212> PRT
<213> Homo sapiens

<400> 128
Met Ser Ala Leu Arg Arg Lys Phe Gly Asp Asp Tyr Gln Val Val
1 5 10 15

Thr Thr Ser Ser Ser Gly Ser Gly Leu Gln Pro Gln Gly Pro Gly
20 25 30

Gln Asp Pro Gln Gln Gln Leu Val Pro Lys Lys Lys Arg Gln Arg
35 40 45

Phe Val Asp Lys Asn Gly Arg Cys Asn Val Gln His Gly Asn Leu
50 55 60

Gly Ser Glu Thr Ser Arg Tyr Leu Ser Asp Leu Phe Thr Thr Leu
65 70 75

Val Asp Leu Lys Trp Arg Trp Asn Leu Phe Ile Phe Ile Leu Thr
80 85 90

1285
1284

Tyr Thr Val Ala Trp Leu Phe Met Ala Ser Met Trp Trp Val Ile
95 100 105

Ala Tyr Thr Arg Gly Asp Leu Asn Lys Ala His Val Gly Asn Tyr
110 115 120

Thr Pro Cys Val Ala Asn Val Tyr Asn Phe Pro Ser Ala Phe Leu
125 130 135

Phe Phe Ile Glu Thr Glu Ala Thr Ile Gly Tyr Gly Tyr Arg Tyr
140 145 150

Ile Thr Asp Lys Cys Pro Glu Gly Ile Ile Leu Phe Leu Phe Gln
155 160 165

Ser Ile Leu Gly Ser Ile Val Asp Ala Phe Leu Ile Gly Cys Met
170 175 180

Phe Ile Lys Met Ser Gln Pro Lys Lys Arg Ala Glu Thr Leu Met
185 190 195

Phe Ser Glu His Ala Val Ile Ser Met Arg Asp Gly Lys Leu Thr
200 205 210

Leu Met Phe Arg Val Gly Asn Leu Arg Asn Ser His Met Val Ser
215 220 225

Ala Gln Ile Arg Cys Lys Leu Leu Lys Ser Arg Gln Thr Pro Glu
230 235 240

Gly Glu Phe Leu Pro Leu Asp Gln Leu Glu Leu Asp Val Gly Phe
245 250 255

Ser Thr Gly Ala Asp Gln Leu Phe Leu Val Ser Pro Leu Thr Ile
260 265 270

Cys His Val Ile Asp Ala Lys Ser Pro Phe Tyr Asp Leu Ser Gln
275 280 285

Arg Ser Met Gln Thr Glu Gln Phe Glu Ile Val Val Ile Leu Glu
290 295 300

Gly Ile Val Glu Thr Thr Gly Met Thr Cys Gln Ala Arg Thr Ser
305 310 315

Tyr Thr Glu Asp Glu Val Leu Trp Gly His Arg Phe Phe Pro Val
320 325 330

Ile Ser Leu Glu Glu Gly Phe Phe Lys Val Asp Tyr Ser Gln Phe
335 340 345

His Ala Thr Phe Glu Val Pro Thr Pro Pro Tyr Ser Val Lys Glu
350 355 360

Gln Glu Glu Met Leu Leu Met Ser Ser Pro Leu Ile Ala Pro Ala
365 370 375

Ile Thr Asn Ser Lys Glu Arg His Asn Ser Val Glu Cys Leu Asp
380 385 390

Gly Leu Asp Asp Ile Thr Thr Lys Leu Pro Ser Lys Leu Gln Lys
395 400 405

Ile Thr Gly Arg Glu Asp Phe Pro Lys Lys Leu Leu Arg Met Ser
410 415 420

Ser Thr Thr Ser Glu Lys Ala Tyr Ser Leu Gly Asp Leu Pro Met
425 430 435

Lys Leu Gln Arg Ile Ser Ser Val Pro Gly Asn Ser Glu Glu Lys
440 445 450

Leu Val Ser Lys Thr Thr Lys Met Leu Ser Asp Pro Met Ser Gln
455 460 465

Ser Val Ala Asp Leu Pro Pro Lys Leu Gln Lys Met Ala Gly Gly

470 475 480

Ala Ala Arg Met Glu Gly Asn Leu Pro Ala Lys Leu Arg Lys Met
485 490 495

Asn Ser Asp Arg Phe Thr
500

<210> 129
<211> 1148
<212> PRT
<213> Homo sapiens

<400> 129
Met Pro Leu Phe Lys Asn Thr Ser Val Ser Ser Leu Tyr Ser Gly
1 5 10 15

Cys Arg Leu Thr Leu Leu Arg Pro Glu Lys Asp Gly Ala Ala Thr
20 25 30

Arg Val Asp Ala Val Cys Thr His Arg Pro Asp Pro Lys Ser Pro
35 40 45

Gly Leu Asp Arg Glu Arg Leu Tyr Trp Lys Leu Ser Gln Leu Thr
50 55 60

His Gly Ile Thr Glu Leu Gly Pro Tyr Thr Leu Asp Arg His Ser
65 70 75

Leu Tyr Val Asn Gly Phe Thr His Gln Ser Ser Met Thr Thr Thr
80 85 90

Arg Thr Pro Asp Thr Ser Thr Met His Leu Ala Thr Ser Arg Thr
95 100 105

Pro Ala Ser Leu Ser Gly Pro Thr Thr Ala Ser Pro Leu Leu Val
110 115 120

Leu Phe Thr Ile Asn Phe Thr Ile Thr Asn Leu Arg Tyr Glu Glu
125 130 135

Asn Met His His Pro Gly Ser Arg Lys Phe Asn Thr Thr Glu Arg
140 145 150

Val Leu Gln Gly Leu Leu Arg Pro Val Phe Lys Asn Thr Ser Val
155 160 165

Gly Pro Leu Tyr Ser Gly Cys Arg Leu Thr Leu Leu Arg Pro Lys
170 175 180

Lys Asp Gly Ala Ala Thr Lys Val Asp Ala Ile Cys Thr Tyr Arg
185 190 195

Pro Asp Pro Lys Ser Pro Gly Leu Asp Arg Glu Gln Leu Tyr Trp
200 205 210

Glu Leu Ser Gln Leu Thr His Ser Ile Thr Glu Leu Gly Pro Tyr
215 220 225

Thr Leu Asp Arg Asp Ser Leu Tyr Val Asn Gly Phe Thr Gln Arg
230 235 240

Ser Ser Val Pro Thr Thr Ser Ile Pro Gly Thr Pro Thr Val Asp
245 250 255

Leu Gly Thr Ser Gly Thr Pro Val Ser Lys Pro Gly Pro Ser Ala
260 265 270

Ala Ser Pro Leu Leu Val Leu Phe Thr Leu Asn Phe Thr Ile Thr
275 280 285

Asn Leu Arg Tyr Glu Glu Asn Met Gln His Pro Gly Ser Arg Lys
290 295 300

Phe Asn Thr Thr Glu Arg Val Leu Gln Gly Leu Leu Arg Ser Leu
305 310 315

Phe Lys Ser Thr Ser Val Gly Pro Leu Tyr Ser Gly Cys Arg Leu
320 325 330

Thr Leu Leu Arg Pro Glu Lys Asp Gly Thr Ala Thr Gly Val Asp
335 340 345

Ala Ile Cys Thr His His Pro Asp Pro Lys Ser Pro Arg Leu Asp
350 355 360

Arg Glu Gln Leu Tyr Trp Glu Leu Ser Gln Leu Thr His Asn Ile
365 370 375

Thr Glu Leu Gly His Tyr Ala Leu Asp Asn Asp Ser Leu Phe Val
380 385 390

Asn Gly Phe Thr His Arg Ser Ser Val Ser Thr Thr Ser Thr Pro
395 400 405

Gly Thr Pro Thr Val Tyr Leu Gly Ala Ser Lys Thr Pro Ala Ser
410 415 420

Ile Phe Gly Pro Ser Ala Ala Ser His Leu Leu Ile Leu Phe Thr
425 430 435

Leu Asn Phe Thr Ile Thr Asn Leu Arg Tyr Glu Glu Asn Met Trp
440 445 450

Pro Gly Ser Arg Lys Phe Asn Thr Thr Glu Arg Val Leu Gln Gly
455 460 465

Leu Leu Arg Pro Leu Phe Lys Asn Thr Ser Val Gly Pro Leu Tyr
470 475 480

Ser Gly Ser Arg Leu Thr Leu Leu Arg Pro Glu Lys Asp Gly Glu
485 490 495

Ala Thr Gly Val Asp Ala Ile Cys Thr His Arg Pro Asp Pro Thr
500 505 510

Gly Pro Gly Leu Asp Arg Glu Gln Leu Tyr Leu Glu Leu Ser Gln
515 520 525

Leu Thr His Ser Ile Thr Glu Leu Gly Pro Tyr Thr Leu Asp Arg
530 535 540

Asp Ser Leu Tyr Val Asn Gly Phe Thr His Arg Ser Ser Val Pro
545 550 555

Thr Thr Ser Thr Gly Val Val Ser Glu Glu Pro Phe Thr Leu Asn
560 565 570

Phe Thr Ile Asn Asn Leu Arg Tyr Met Ala Asp Met Gly Gln Pro
575 580 585

Gly Ser Leu Lys Phe Asn Ile Thr Asp Asn Val Met Lys His Leu
590 595 600

Leu Ser Pro Leu Phe Gln Arg Ser Ser Leu Gly Ala Arg Tyr Thr
605 610 615

Gly Cys Arg Val Ile Ala Leu Arg Ser Val Lys Asn Gly Ala Glu
620 625 630

Thr Arg Val Asp Leu Leu Cys Thr Tyr Leu Gln Pro Leu Ser Gly
635 640 645

Pro Gly Leu Pro Ile Lys Gln Val Phe His Glu Leu Ser Gln Gln
650 655 660

Thr His Gly Ile Thr Arg Leu Gly Pro Tyr Ser Leu Asp Lys Asp
665 670 675

Ser Leu Tyr Leu Asn Gly Tyr Asn Glu Pro Gly Leu Asp Glu Pro
680 685 690

1288
1287

Pro Thr Thr Pro Lys Pro Ala Thr Thr Phe Leu Pro Pro Leu Ser
695 700 705

Glu Ala Thr Thr Ala Met Gly Tyr His Leu Lys Thr Leu Thr Leu
710 715 720

Asn Phe Thr Ile Ser Asn Leu Gln Tyr Ser Pro Asp Met Gly Lys
725 730 735

Gly Ser Ala Thr Phe Asn Ser Thr Glu Gly Val Leu Gln His Leu
740 745 750

Leu Arg Pro Leu Phe Gln Lys Ser Ser Met Gly Pro Phe Tyr Leu
755 760 765

Gly Cys Gln Leu Ile Ser Leu Arg Pro Glu Lys Asp Gly Ala Ala
770 775 780

Thr Gly Val Asp Thr Thr Cys Thr Tyr His Pro Asp Pro Val Gly
785 790 795

Pro Gly Leu Asp Ile Gln Gln Leu Tyr Trp Glu Leu Ser Gln Leu
800 805 810

Thr His Gly Val Thr Gln Leu Gly Phe Tyr Val Leu Asp Arg Asp
815 820 825

Ser Leu Phe Ile Asn Gly Tyr Ala Pro Gln Asn Leu Ser Ile Arg
830 835 840

Gly Glu Tyr Gln Ile Asn Phe His Ile Val Asn Trp Asn Leu Ser
845 850 855

Asn Pro Asp Pro Thr Ser Ser Glu Tyr Ile Thr Leu Leu Arg Asp
860 865 870

Ile Gln Asp Lys Val Thr Thr Leu Tyr Lys Gly Ser Gln Leu His
875 880 885

Asp Thr Phe Arg Phe Cys Leu Val Thr Asn Leu Thr Met Asp Ser
890 895 900

Val Leu Val Thr Val Lys Ala Leu Phe Ser Ser Asn Leu Asp Pro
905 910 915

Ser Leu Val Glu Gln Val Phe Leu Asp Lys Thr Leu Asn Ala Ser
920 925 930

Phe His Trp Leu Gly Ser Thr Tyr Gln Leu Val Asp Ile His Val
935 940 945

Thr Glu Met Glu Ser Ser Val Tyr Gln Pro Thr Ser Ser Ser Ser
950 955 960

Thr Gln His Phe Tyr Pro Asn Phe Thr Ile Thr Asn Leu Pro Tyr
965 970 975

Ser Gln Asp Lys Ala Gln Pro Gly Thr Thr Asn Tyr Gln Arg Asn
980 985 990

Lys Arg Asn Ile Glu Asp Ala Leu Asn Gln Leu Phe Arg Asn Ser
995 1000 1005

Ser Ile Lys Ser Tyr Phe Ser Asp Cys Gln Val Ser Thr Phe Arg
1010 1015 1020

Ser Val Pro Asn Arg His His Thr Gly Val Asp Ser Leu Cys Asn
1025 1030 1035

Phe Ser Pro Leu Ala Arg Arg Val Asp Arg Val Ala Ile Tyr Glu
1040 1045 1050

Glu Phe Leu Arg Met Thr Arg Asn Gly Thr Gln Leu Gln Asn Phe
1055 1060 1065

Thr Leu Asp Arg Ser Ser Val Leu Val Asp Gly Tyr Ser Pro Asn

1070 1075 1080

Arg Asn Glu Pro Leu Thr Gly Asn Ser Asp Leu Pro Phe Trp Ala
1085 1090 1095

Val Ile Phe Ile Gly Leu Ala Gly Leu Leu Gly Leu Ile Thr Cys
1100 1105 1110

Leu Ile Cys Gly Val Leu Val Thr Thr Arg Arg Arg Lys Lys Glu
1115 1120 1125

Gly Glu Tyr Asn Val Gln Gln Gln Cys Pro Gly Tyr Tyr Gln Ser
1130 1135 1140

His Leu Asp Leu Glu Asp Leu Gln
1145

<210> 130
<211> 564
<212> PRT
<213> Homo sapiens

<400> 130
Met Gly Asn Trp Val Val Asn His Trp Phe Ser Val Leu Phe Leu
1 5 10 15

Val Val Trp Leu Gly Leu Asn Val Phe Leu Phe Val Asp Ala Phe
20 25 30

Leu Lys Tyr Glu Lys Ala Asp Lys Tyr Tyr Tyr Thr Arg Lys Ile
35 40 45

Leu Gly Ser Thr Leu Ala Cys Ala Arg Ala Ser Ala Leu Cys Leu
50 55 60

Asn Phe Asn Ser Thr Leu Ile Leu Leu Pro Val Cys Arg Asn Leu
65 70 75

Leu Ser Phe Leu Arg Gly Thr Cys Ser Phe Cys Ser Arg Thr Leu
80 85 90

Arg Lys Gln Leu Asp His Asn Leu Thr Phe His Lys Leu Val Ala
95 100 105

Tyr Met Ile Cys Leu His Thr Ala Ile His Ile Ile Ala His Leu
110 115 120

Phe Asn Phe Asp Cys Tyr Ser Arg Ser Arg Gln Ala Thr Asp Gly
125 130 135

Ser Leu Ala Ser Ile Leu Ser Ser Leu Ser His Asp Glu Lys Lys
140 145 150

Gly Gly Ser Trp Leu Asn Pro Ile Gln Ser Arg Asn Thr Thr Val
155 160 165

Glu Tyr Val Thr Phe Thr Ser Val Ala Gly Leu Thr Gly Val Ile
170 175 180

Met Thr Ile Ala Leu Ile Leu Met Val Thr Ser Ala Thr Glu Phe
185 190 195

Ile Arg Arg Ser Tyr Phe Glu Val Phe Trp Tyr Thr His His Leu
200 205 210

Phe Ile Phe Tyr Ile Leu Gly Leu Gly Ile His Gly Ile Gly Gly
215 220 225

Ile Val Arg Gly Gln Thr Glu Glu Ser Met Asn Glu Ser His Pro
230 235 240

Arg Lys Cys Ala Glu Ser Phe Glu Met Trp Asp Asp Arg Asp Ser
245 250 255

His Cys Arg Arg Pro Lys Phe Glu Gly His Pro Pro Glu Ser Trp
260 265 270

Lys Trp Ile Leu Ala Pro Val Ile Leu Tyr Ile Cys Glu Arg Ile
275 280 285

Leu Arg Phe Tyr Arg Ser Gln Gln Lys Val Val Ile Thr Lys Val
290 295 300

Val Met His Pro Ser Lys Val Leu Glu Leu Gln Met Asn Lys Arg
305 310 315

Gly Phe Ser Met Glu Val Gly Gln Tyr Ile Phe Val Asn Cys Pro
320 325 330

Ser Ile Ser Leu Leu Glu Trp His Pro Phe Thr Leu Thr Ser Ala
335 340 345

Pro Glu Glu Asp Phe Phe Ser Ile His Ile Arg Ala Ala Gly Asp
350 355 360

Trp Thr Glu Asn Leu Ile Arg Ala Phe Glu Gln Gln Tyr Ser Pro
365 370 375

Ile Pro Arg Ile Glu Val Asp Gly Pro Phe Gly Thr Ala Ser Glu
380 385 390

Asp Val Phe Gln Tyr Glu Val Ala Val Leu Val Gly Ala Gly Ile
395 400 405

Gly Val Thr Pro Phe Ala Ser Ile Leu Lys Ser Ile Trp Tyr Lys
410 415 420

Phe Gln Cys Ala Asp His Asn Leu Lys Thr Lys Lys Ile Tyr Phe
425 430 435

Tyr Trp Ile Cys Arg Glu Thr Gly Ala Phe Ser Trp Phe Asn Asn
440 445 450

Leu Leu Thr Ser Leu Glu Gln Glu Met Glu Glu Leu Gly Lys Val
455 460 465

Gly Phe Leu Asn Tyr Arg Leu Phe Leu Thr Gly Trp Asp Ser Asn
470 475 480

Ile Val Gly His Ala Ala Leu Asn Phe Asp Lys Ala Thr Asp Ile
485 490 495

Val Thr Gly Leu Lys Gln Lys Thr Ser Phe Gly Arg Pro Met Trp
500 505 510

Asp Asn Glu Phe Ser Thr Ile Ala Thr Ser His Pro Lys Ser Val
515 520 525

Val Gly Val Phe Leu Cys Gly Pro Arg Thr Leu Ala Lys Ser Leu
530 535 540

Arg Lys Cys Cys His Arg Tyr Ser Ser Leu Asp Pro Arg Lys Val
545 550 555

Gln Phe Tyr Phe Asn Lys Glu Asn Phe
560

<210> 131
<211> 564
<212> PRT
<213> Homo sapiens

<400> 131
Met Gly Asn Trp Val Val Asn His Trp Phe Ser Val Leu Phe Leu
1 5 10 15

Val Val Trp Leu Gly Leu Asn Val Phe Leu Phe Val Asp Ala Phe
20 25 30

Leu Lys Tyr Glu Lys Ala Asp Lys Tyr Tyr Tyr Thr Arg Lys Ile
35 40 45

Leu Gly Ser Thr Leu Ala Cys Ala Arg Ala Ser Ala Leu Cys Leu
50 55 60

Asn Phe Asn Ser Thr Leu Ile Leu Leu Pro Val Cys Arg Asn Leu
65 70 75

Leu Ser Phe Leu Arg Gly Thr Cys Ser Phe Cys Ser Arg Thr Leu
80 85 90

Arg Lys Gln Leu Asp His Asn Leu Thr Phe His Lys Leu Val Ala
95 100 105

Tyr Met Ile Cys Leu His Thr Ala Ile His Ile Ile Ala His Leu
110 115 120

Phe Asn Phe Asp Cys Tyr Ser Arg Ser Arg Gln Ala Thr Asp Gly
125 130 135

Ser Leu Ala Ser Ile Leu Ser Ser Leu Ser His Asp Glu Lys Lys
140 145 150

Gly Gly Ser Trp Leu Asn Pro Ile Gln Ser Arg Asn Thr Thr Val
155 160 165

Glu Tyr Val Thr Phe Thr Ser Ile Ala Gly Leu Thr Gly Val Ile
170 175 180

Met Thr Ile Ala Leu Ile Leu Met Val Thr Ser Ala Thr Glu Phe
185 190 195

Ile Arg Arg Ser Tyr Phe Glu Val Phe Trp Tyr Thr His His Leu
200 205 210

Phe Ile Phe Tyr Ile Leu Gly Leu Gly Ile His Gly Ile Gly Gly
215 220 225

Ile Val Arg Gly Gln Thr Glu Glu Ser Met Asn Glu Ser His Pro
230 235 240

Arg Lys Cys Ala Glu Ser Phe Glu Met Trp Asp Asp Arg Asp Ser
245 250 255

His Cys Arg Arg Pro Lys Phe Glu Gly His Pro Pro Glu Ser Trp
260 265 270

Lys Trp Ile Leu Ala Pro Val Ile Leu Tyr Ile Cys Glu Arg Ile
275 280 285

Leu Arg Phe Tyr Arg Ser Gln Gln Lys Val Val Ile Thr Lys Val
290 295 300

Val Met His Pro Ser Lys Val Leu Glu Leu Gln Met Asn Lys Arg
305 310 315

Gly Phe Ser Met Glu Val Gly Gln Tyr Ile Phe Val Asn Cys Pro
320 325 330

Ser Ile Ser Leu Leu Glu Trp His Pro Phe Thr Leu Thr Ser Ala
335 340 345

Pro Glu Glu Asp Phe Phe Ser Ile His Ile Arg Ala Ala Gly Asp
350 355 360

Trp Thr Glu Asn Leu Ile Arg Ala Phe Glu Gln Gln Tyr Ser Pro
365 370 375

Ile Pro Arg Ile Glu Val Asp Gly Pro Phe Gly Thr Ala Ser Glu
380 385 390

Asp Val Phe Gln Tyr Glu Val Ala Val Leu Val Gly Ala Gly Ile
395 400 405

Gly Val Thr Pro Phe Ala Ser Ile Leu Lys Ser Ile Trp Tyr Lys
410 415 420

Phe Gln Cys Ala Asp His Asn Leu Lys Thr Lys Lys Ile Tyr Phe

425 430 435

Tyr Trp Ile Cys Arg Glu Thr Gly Ala Phe Ser Trp Phe Asn Asn
440 445 450

Leu Leu Thr Ser Leu Glu Gln Glu Met Glu Glu Leu Gly Lys Val
455 460 465

Gly Phe Leu Asn Tyr Arg Leu Phe Leu Thr Gly Trp Asp Ser Asn
470 475 480

Ile Val Gly His Ala Ala Leu Asn Phe Asp Lys Ala Thr Asp Ile
485 490 495

Val Thr Gly Leu Lys Gln Lys Thr Ser Phe Gly Arg Pro Met Trp
500 505 510

Asp Asn Glu Phe Ser Thr Ile Ala Thr Ser His Pro Lys Ser Val
515 520 525

Val Gly Val Phe Leu Cys Gly Pro Arg Thr Leu Ala Lys Ser Leu
530 535 540

Arg Lys Cys Cys His Arg Tyr Ser Ser Leu Asp Pro Arg Lys Val
545 550 555

Gln Phe Tyr Phe Asn Lys Glu Asn Phe
560

<210> 132
<211> 294
<212> PRT
<213> Homo sapiens

<400> 132
Met Gly Asn Trp Val Val Asn His Trp Phe Ser Val Leu Phe Leu
1 5 10 15

Val Val Trp Leu Gly Leu Asn Val Phe Leu Phe Val Asp Ala Phe
20 25 30

Leu Lys Tyr Glu Lys Ala Asp Lys Tyr Tyr Tyr Thr Arg Lys Ile
35 40 45

Leu Gly Ser Trp Lys Trp Ile Leu Ala Pro Val Ile Leu Tyr Ile
50 55 60

Cys Glu Arg Ile Leu Arg Phe Tyr Arg Ser Gln Gln Lys Val Val
65 70 75

Ile Thr Lys Val Val Met His Pro Ser Lys Val Leu Glu Leu Gln
80 85 90

Met Asn Lys Arg Gly Phe Ser Met Glu Val Gly Gln Tyr Ile Phe
95 100 105

Val Asn Cys Pro Ser Ile Ser Leu Leu Glu Trp His Pro Phe Thr
110 115 120

Leu Thr Ser Ala Pro Glu Glu Asp Phe Phe Ser Ile His Ile Arg
125 130 135

Ala Ala Gly Asp Trp Thr Glu Asn Leu Ile Arg Ala Phe Glu Gln
140 145 150

Gln Tyr Ser Pro Ile Pro Arg Ile Glu Val Asp Gly Pro Phe Gly
155 160 165

Thr Ala Ser Glu Asp Val Phe Gln Tyr Glu Val Ala Val Leu Val
170 175 180

Gly Ala Gly Ile Gly Val Thr Pro Phe Ala Ser Ile Leu Lys Ser
185 190 195

Ile Trp Tyr Lys Phe Gln Cys Ala Asp His Asn Leu Lys Thr Lys
200 205 210

Lys Val Gly His Ala Ala Leu Asn Phe Asp Lys Ala Thr Asp Ile
215 220 225

Val Thr Gly Leu Lys Gln Lys Thr Ser Phe Gly Arg Pro Met Trp
230 235 240

Asp Asn Glu Phe Ser Thr Ile Ala Thr Ser His Pro Lys Ser Val
245 250 255

Val Gly Val Phe Leu Cys Gly Pro Arg Thr Leu Ala Lys Ser Leu
260 265 270

Arg Lys Cys Cys His Arg Tyr Ser Ser Leu Asp Pro Arg Lys Val
275 280 285

Gln Phe Tyr Phe Asn Lys Glu Asn Phe
290

<210> 133
<211> 515
<212> PRT
<213> Homo sapiens

<400> 133
Met Gly Asn Trp Val Val Asn His Trp Phe Ser Val Leu Phe Leu
1 5 10 15

Val Val Trp Leu Gly Leu Asn Val Phe Leu Phe Val Asp Ala Phe
20 25 30

Leu Lys Tyr Glu Lys Ala Asp Lys Tyr Tyr Tyr Thr Arg Lys Ile
35 40 45

Leu Gly Ser Thr Leu Ala Cys Ala Arg Ala Ser Ala Leu Cys Leu
50 55 60

Asn Phe Asn Ser Thr Leu Ile Leu Leu Pro Val Cys Arg Asn Leu
65 70 75

Leu Ser Phe Leu Arg Gly Thr Cys Ser Phe Cys Ser Arg Thr Leu
80 85 90

Arg Lys Gln Leu Asp His Asn Leu Thr Phe His Lys Leu Val Ala
95 100 105

Tyr Met Ile Cys Leu His Thr Ala Ile His Ile Ile Ala His Leu
110 115 120

Phe Asn Phe Asp Cys Tyr Ser Arg Ser Arg Gln Ala Thr Asp Gly
125 130 135

Ser Leu Ala Ser Ile Leu Ser Ser Leu Ser His Asp Glu Lys Lys
140 145 150

Gly Gly Ser Trp Leu Asn Pro Ile Gln Ser Arg Asn Thr Thr Val
155 160 165

Glu Tyr Val Thr Phe Thr Ser Ile Ala Gly Leu Thr Gly Val Ile
170 175 180

Met Thr Ile Ala Leu Ile Leu Met Val Thr Ser Ala Thr Glu Phe
185 190 195

Ile Arg Arg Ser Tyr Phe Glu Val Phe Trp Tyr Thr His His Leu
200 205 210

Phe Ile Phe Tyr Ile Leu Gly Leu Gly Ile His Gly Ile Gly Gly
215 220 225

Ile Val Arg Gly Gln Thr Glu Glu Ser Met Asn Glu Ser His Pro
230 235 240

Arg Lys Cys Ala Glu Ser Phe Glu Met Trp Asp Asp Arg Asp Ser
245 250 255

```
His Cys Arg Arg Pro Lys Phe Glu Gly His Pro Pro Glu Ser Trp
                260                 265                 270

Lys Trp Ile Leu Ala Pro Val Ile Leu Tyr Ile Cys Glu Arg Ile
                275                 280                 285

Leu Arg Phe Tyr Arg Ser Gln Gln Lys Val Val Ile Thr Lys Val
                290                 295                 300

Val Met His Pro Ser Lys Val Leu Glu Leu Gln Met Asn Lys Arg
                305                 310                 315

Gly Phe Ser Met Glu Val Gly Gln Tyr Ile Phe Val Asn Cys Pro
                320                 325                 330

Ser Ile Ser Leu Leu Glu Trp His Pro Phe Thr Leu Thr Ser Ala
                335                 340                 345

Pro Glu Glu Asp Phe Phe Ser Ile His Ile Arg Ala Ala Gly Asp
                350                 355                 360

Trp Thr Glu Asn Leu Ile Arg Ala Phe Glu Gln Gln Tyr Ser Pro
                365                 370                 375

Ile Pro Arg Ile Glu Val Asp Gly Pro Phe Gly Thr Ala Ser Glu
                380                 385                 390

Asp Val Phe Gln Tyr Glu Val Ala Val Leu Val Gly Ala Gly Ile
                395                 400                 405

Gly Val Thr Pro Phe Ala Ser Ile Leu Lys Ser Ile Trp Tyr Lys
                410                 415                 420

Phe Gln Cys Ala Asp His Asn Leu Lys Thr Lys Lys Val Gly His
                425                 430                 435

Ala Ala Leu Asn Phe Asp Lys Ala Thr Asp Ile Val Thr Gly Leu
                440                 445                 450

Lys Gln Lys Thr Ser Phe Gly Arg Pro Met Trp Asp Asn Glu Phe
                455                 460                 465

Ser Thr Ile Ala Thr Ser His Pro Lys Ser Val Val Gly Val Phe
                470                 475                 480

Leu Cys Gly Pro Arg Thr Leu Ala Lys Ser Leu Arg Lys Cys Cys
                485                 490                 495

His Arg Tyr Ser Ser Leu Asp Pro Arg Lys Val Gln Phe Tyr Phe
                500                 505                 510

Asn Lys Glu Asn Phe
                515

<210> 134
<211> 757
<212> PRT
<213> Homo sapiens

<400> 134
 Met Thr Ser Gly Ser Val Phe Phe Tyr Ile Leu Ile Phe Gly Lys
   1               5                  10                  15

 Tyr Phe Ser His Gly Gly Gly Gln Asp Val Lys Cys Ser Leu Gly
                20                  25                  30

 Tyr Phe Pro Cys Gly Asn Ile Thr Lys Cys Leu Pro Gln Leu Leu
                35                  40                  45

 His Cys Asn Gly Val Asp Asp Cys Gly Asn Gln Ala Asp Glu Asp
                50                  55                  60

 Asn Cys Gly Asp Asn Asn Gly Trp Ser Met Gln Phe Asp Lys Tyr
                65                  70                  75

 Phe Ala Ser Tyr Tyr Lys Met Thr Ser Gln Tyr Pro Phe Glu Ala
```

1295
1294

```
                  80                  85                  90

Glu Thr Pro Glu Cys Leu Val Gly Ser Val Pro Val Gln Cys Leu
                  95                 100                 105

Cys Gln Gly Leu Glu Leu Asp Cys Asp Glu Thr Asn Leu Arg Ala
                 110                 115                 120

Val Pro Ser Val Ser Ser Asn Val Thr Ala Met Ser Leu Gln Trp
                 125                 130                 135

Asn Leu Ile Arg Lys Leu Pro Pro Asp Cys Phe Lys Asn Tyr His
                 140                 145                 150

Asp Leu Gln Lys Leu Tyr Leu Gln Asn Asn Lys Ile Thr Ser Ile
                 155                 160                 165

Ser Ile Tyr Ala Phe Arg Gly Leu Asn Ser Leu Thr Lys Leu Tyr
                 170                 175                 180

Leu Ser His Asn Arg Ile Thr Phe Leu Lys Pro Gly Val Phe Glu
                 185                 190                 195

Asp Leu His Arg Leu Glu Trp Leu Ile Ile Glu Asp Asn His Leu
                 200                 205                 210

Ser Arg Ile Ser Pro Pro Thr Phe Tyr Gly Leu Asn Ser Leu Ile
                 215                 220                 225

Leu Leu Val Leu Met Asn Asn Val Leu Thr Arg Leu Pro Asp Lys
                 230                 235                 240

Pro Leu Cys Gln His Met Pro Arg Leu His Trp Leu Asp Leu Glu
                 245                 250                 255

Gly Asn His Ile His Asn Leu Arg Asn Leu Thr Phe Ile Ser Cys
                 260                 265                 270

Ser Asn Leu Thr Val Leu Val Met Arg Lys Asn Lys Ile Asn His
                 275                 280                 285

Leu Asn Glu Asn Thr Phe Ala Pro Leu Gln Lys Leu Asp Glu Leu
                 290                 295                 300

Asp Leu Gly Ser Asn Lys Ile Glu Asn Leu Pro Pro Leu Ile Phe
                 305                 310                 315

Lys Asp Leu Lys Glu Leu Ser Gln Leu Asn Leu Ser Tyr Asn Pro
                 320                 325                 330

Ile Gln Lys Ile Gln Ala Asn Gln Phe Asp Tyr Leu Val Lys Leu
                 335                 340                 345

Lys Ser Leu Ser Leu Glu Gly Ile Glu Ile Ser Asn Ile Gln Gln
                 350                 355                 360

Arg Met Phe Arg Pro Leu Met Asn Leu Ser His Ile Tyr Phe Lys
                 365                 370                 375

Lys Phe Gln Tyr Cys Gly Tyr Ala Pro His Val Arg Ser Cys Lys
                 380                 385                 390

Pro Asn Thr Asp Gly Ile Ser Ser Leu Glu Asn Leu Leu Ala Ser
                 395                 400                 405

Ile Ile Gln Arg Val Phe Val Trp Val Val Ser Ala Val Thr Cys
                 410                 415                 420

Phe Gly Asn Ile Phe Val Ile Cys Met Arg Pro Tyr Ile Arg Ser
                 425                 430                 435

Glu Asn Lys Leu Tyr Ala Met Ser Ile Ile Ser Leu Cys Cys Ala
                 440                 445                 450

Asp Cys Leu Met Gly Ile Tyr Leu Phe Val Ile Gly Gly Phe Asp
                 455                 460                 465
```

Leu Lys Phe Arg Gly Glu Tyr Asn Lys His Ala Gln Leu Trp Met
470 475 480

Glu Ser Thr His Cys Gln Leu Val Gly Ser Leu Ala Ile Leu Ser
485 490 495

Thr Glu Val Ser Val Leu Leu Leu Thr Phe Leu Thr Leu Glu Lys
500 505 510

Tyr Ile Cys Ile Val Tyr Pro Phe Arg Cys Val Arg Pro Gly Lys
515 520 525

Cys Arg Thr Ile Thr Val Leu Ile Leu Ile Trp Ile Thr Gly Phe
530 535 540

Ile Val Ala Phe Ile Pro Leu Ser Asn Lys Glu Phe Phe Lys Asn
545 550 555

Tyr Tyr Gly Thr Asn Gly Val Cys Phe Pro Leu His Ser Glu Asp
560 565 570

Thr Glu Ser Ile Gly Ala Gln Ile Tyr Ser Val Ala Ile Phe Leu
575 580 585

Gly Ile Asn Leu Ala Ala Phe Ile Ile Ile Val Phe Ser Tyr Gly
590 595 600

Ser Met Phe Tyr Ser Val His Gln Ser Ala Ile Thr Ala Thr Glu
605 610 615

Ile Arg Asn Gln Val Lys Lys Glu Met Ile Leu Ala Lys Arg Phe
620 625 630

Phe Phe Ile Val Phe Thr Asp Ala Leu Cys Trp Ile Pro Ile Phe
635 640 645

Val Val Lys Phe Leu Ser Leu Leu Gln Val Glu Ile Pro Gly Thr
650 655 660

Ile Thr Ser Trp Val Val Ile Phe Ile Leu Pro Ile Asn Ser Ala
665 670 675

Leu Asn Pro Ile Leu Tyr Thr Leu Thr Thr Arg Pro Phe Lys Glu
680 685 690

Met Ile His Arg Phe Trp Tyr Asn Tyr Arg Gln Arg Lys Ser Met
695 700 705

Asp Ser Lys Gly Gln Lys Thr Tyr Ala Pro Ser Phe Ile Trp Val
710 715 720

Glu Met Trp Pro Leu Gln Glu Met Pro Pro Glu Leu Met Lys Pro
725 730 735

Asp Leu Phe Thr Tyr Pro Cys Glu Met Ser Leu Ile Ser Gln Ser
740 745 750

Thr Arg Leu Asn Ser Tyr Ser
755

<210> 135
<211> 757
<212> PRT
<213> Homo sapiens

<400> 135
Met Thr Ser Gly Ser Val Phe Phe Tyr Ile Leu Ile Phe Gly Lys
1 5 10 15

Tyr Phe Ser His Gly Gly Gly Gln Asp Val Lys Cys Ser Leu Gly
20 25 30

Tyr Phe Pro Cys Gly Asn Ile Thr Lys Cys Leu Pro Gln Leu Leu
35 40 45

His Cys Asn Gly Val Asp Asp Cys Gly Asn Gln Ala Asp Glu Asp
50 55 60

Asn Cys Gly Asp Asn Asn Gly Trp Ser Leu Gln Phe Asp Lys Tyr
65 70 75

Phe Ala Ser Tyr Tyr Lys Met Thr Ser Gln Tyr Pro Phe Glu Ala
80 85 90

Glu Thr Pro Glu Cys Leu Val Gly Ser Val Pro Val Gln Cys Leu
95 100 105

Cys Gln Gly Leu Glu Leu Asp Cys Asp Glu Thr Asn Leu Arg Ala
110 115 120

Val Pro Ser Val Ser Ser Asn Val Thr Ala Met Ser Leu Gln Trp
125 130 135

Asn Leu Ile Arg Lys Leu Pro Pro Asp Cys Phe Lys Asn Tyr His
140 145 150

Asp Leu Gln Lys Leu Tyr Leu Gln Asn Asn Lys Ile Thr Ser Ile
155 160 165

Ser Ile Tyr Ala Phe Arg Gly Leu Asn Ser Leu Thr Lys Leu Tyr
170 175 180

Leu Ser His Asn Arg Ile Thr Phe Leu Lys Pro Gly Val Phe Glu
185 190 195

Asp Leu His Arg Leu Glu Trp Leu Ile Ile Glu Asp Asn His Leu
200 205 210

Ser Arg Ile Ser Pro Pro Thr Phe Tyr Gly Leu Asn Ser Leu Ile
215 220 225

Leu Leu Val Leu Met Asn Asn Val Leu Thr Arg Leu Pro Asp Lys
230 235 240

Pro Leu Cys Gln His Met Pro Arg Leu His Trp Leu Asp Leu Glu
245 250 255

Gly Asn His Ile His Asn Leu Arg Asn Leu Thr Phe Ile Ser Cys
260 265 270

Ser Asn Leu Thr Val Leu Val Met Arg Lys Asn Lys Ile Asn His
275 280 285

Leu Asn Glu Asn Thr Phe Ala Pro Leu Gln Lys Leu Asp Glu Leu
290 295 300

Asp Leu Gly Ser Asn Lys Ile Glu Asn Leu Pro Pro Leu Ile Phe
305 310 315

Lys Asp Leu Lys Glu Leu Ser Gln Leu Asn Leu Ser Tyr Asn Pro
320 325 330

Ile Gln Lys Ile Gln Ala Asn Gln Phe Asp Tyr Leu Val Lys Leu
335 340 345

Lys Ser Leu Ser Leu Glu Gly Ile Glu Ile Ser Asn Ile Gln Gln
350 355 360

Arg Met Phe Arg Pro Leu Met Asn Leu Ser His Ile Tyr Phe Lys
365 370 375

Lys Phe Gln Tyr Cys Gly Tyr Ala Pro His Val Arg Ser Cys Lys
380 385 390

Pro Asn Thr Asp Gly Ile Ser Ser Leu Glu Asn Leu Leu Ala Ser
395 400 405

Ile Ile Gln Arg Val Phe Val Trp Val Val Ser Ala Val Thr Cys
410 415 420

Phe Gly Asn Ile Phe Val Ile Cys Met Arg Pro Tyr Ile Arg Ser

425 430 435

Glu Asn Lys Leu Tyr Ala Met Ser Ile Ile Ser Leu Cys Cys Ala
440 445 450

Asp Cys Leu Met Gly Ile Tyr Leu Phe Val Ile Gly Gly Phe Asp
455 460 465

Leu Lys Phe Arg Gly Glu Tyr Asn Lys His Ala Gln Leu Trp Met
470 475 480

Glu Ser Thr His Cys Gln Leu Val Gly Ser Leu Ala Ile Leu Ser
485 490 495

Thr Glu Val Ser Val Leu Leu Leu Thr Phe Leu Thr Leu Glu Lys
500 505 510

Tyr Ile Cys Ile Val Tyr Pro Phe Arg Cys Val Arg Pro Gly Lys
515 520 525

Cys Arg Thr Ile Thr Val Leu Ile Leu Ile Trp Ile Thr Gly Phe
530 535 540

Ile Val Ala Phe Ile Pro Leu Ser Asn Lys Glu Phe Phe Lys Asn
545 550 555

Tyr Tyr Gly Thr Asn Gly Val Cys Phe Pro Leu His Ser Glu Asp
560 565 570

Thr Glu Ser Ile Gly Ala Gln Ile Tyr Ser Val Ala Ile Phe Leu
575 580 585

Gly Ile Asn Leu Ala Ala Phe Ile Ile Ile Val Phe Ser Tyr Gly
590 595 600

Ser Met Phe Tyr Ser Val His Gln Ser Ala Ile Thr Ala Thr Glu
605 610 615

Ile Arg Asn Gln Val Lys Lys Glu Met Ile Leu Ala Lys Arg Phe
620 625 630

Phe Phe Ile Val Phe Thr Asp Ala Leu Cys Trp Ile Pro Ile Phe
635 640 645

Val Val Lys Phe Leu Ser Leu Leu Gln Val Glu Ile Pro Gly Thr
650 655 660

Ile Thr Ser Trp Val Val Ile Phe Ile Leu Pro Ile Asn Ser Ala
665 670 675

Leu Asn Pro Ile Leu Tyr Thr Leu Thr Thr Arg Pro Phe Lys Glu
680 685 690

Met Ile His Arg Phe Trp Tyr Asn Tyr Arg Gln Arg Lys Ser Met
695 700 705

Asp Ser Lys Gly Gln Lys Thr Tyr Ala Pro Ser Phe Ile Trp Val
710 715 720

Glu Met Trp Pro Leu Gln Glu Met Pro Pro Glu Leu Met Lys Pro
725 730 735

Asp Leu Phe Thr Tyr Pro Cys Glu Met Ser Leu Ile Ser Gln Ser
740 745 750

Thr Arg Leu Asn Ser Tyr Ser
755

<210> 136
<211> 879
<212> PRT
<213> Homo sapiens

<400> 136
Met Gly Arg Leu Ala Ser Arg Pro Leu Leu Leu Ala Leu Leu Ser
1 5 10 15

```
Leu Ala Leu Cys Arg Gly Arg Val Val Arg Val Pro Thr Ala Thr
                20                  25                  30

Leu Val Arg Val Val Gly Thr Glu Leu Val Ile Pro Cys Asn Val
                35                  40                  45

Ser Asp Tyr Asp Gly Pro Ser Glu Gln Asn Phe Asp Trp Ser Phe
                50                  55                  60

Ser Ser Leu Gly Ser Ser Phe Val Glu Leu Ala Ser Thr Trp Glu
                65                  70                  75

Val Gly Phe Pro Ala Gln Leu Tyr Gln Glu Arg Leu Gln Arg Gly
                80                  85                  90

Glu Ile Leu Leu Arg Arg Thr Ala Asn Asp Ala Val Glu Leu His
                95                  100                 105

Ile Lys Asn Val Gln Pro Ser Asp Gln Gly His Tyr Lys Cys Ser
                110                 115                 120

Thr Pro Ser Thr Asp Ala Thr Val Gln Gly Asn Tyr Glu Asp Thr
                125                 130                 135

Val Gln Val Lys Val Leu Ala Asp Ser Leu His Val Gly Pro Ser
                140                 145                 150

Ala Arg Pro Pro Pro Ser Leu Ser Leu Arg Glu Gly Glu Pro Phe
                155                 160                 165

Glu Leu Arg Cys Thr Ala Ala Ser Ala Ser Pro Leu His Thr His
                170                 175                 180

Leu Ala Leu Leu Trp Glu Val His Arg Gly Pro Ala Arg Arg Ser
                185                 190                 195

Val Leu Ala Leu Thr His Glu Gly Arg Phe His Pro Gly Leu Gly
                200                 205                 210

Tyr Glu Gln Arg Tyr His Ser Gly Asp Val Arg Leu Asp Thr Val
                215                 220                 225

Gly Ser Asp Ala Tyr Arg Leu Ser Val Ser Arg Ala Leu Ser Ala
                230                 235                 240

Asp Gln Gly Ser Tyr Arg Cys Ile Val Ser Glu Trp Ile Ala Glu
                245                 250                 255

Gln Gly Asn Trp Gln Glu Ile Gln Glu Lys Ala Val Glu Val Ala
                260                 265                 270

Thr Val Val Ile Gln Pro Thr Val Leu Arg Ala Ala Val Pro Lys
                275                 280                 285

Asn Val Ser Val Ala Glu Gly Lys Glu Leu Asp Leu Thr Cys Asn
                290                 295                 300

Ile Thr Thr Asp Arg Ala Asp Asp Val Arg Pro Glu Val Thr Trp
                305                 310                 315

Ser Phe Ser Arg Met Pro Asp Ser Thr Leu Pro Gly Ser Arg Val
                320                 325                 330

Leu Ala Arg Leu Asp Arg Asp Ser Leu Val His Ser Ser Pro His
                335                 340                 345

Val Ala Leu Ser His Val Asp Ala Arg Ser Tyr His Leu Leu Val
                350                 355                 360

Arg Asp Val Ser Lys Glu Asn Ser Gly Tyr Tyr Tyr Cys His Val
                365                 370                 375

Ser Leu Trp Ala Pro Gly His Asn Arg Ser Trp His Lys Val Ala
                380                 385                 390
```

Glu Ala Val Ser Ser Pro Ala Gly Val Gly Val Thr Trp Leu Glu
395 400 405

Pro Asp Tyr Gln Val Tyr Leu Asn Ala Ser Lys Val Pro Gly Phe
410 415 420

Ala Asp Asp Pro Thr Glu Leu Ala Cys Arg Val Val Asp Thr Lys
425 430 435

Ser Gly Glu Ala Asn Val Arg Phe Thr Val Ser Trp Tyr Tyr Arg
440 445 450

Met Asn Arg Arg Ser Asp Asn Val Val Thr Ser Glu Leu Leu Ala
455 460 465

Val Met Asp Gly Asp Trp Thr Leu Lys Tyr Gly Glu Arg Ser Lys
470 475 480

Gln Arg Ala Gln Asp Gly Asp Phe Ile Phe Ser Lys Glu His Thr
485 490 495

Asp Thr Phe Asn Phe Arg Ile Gln Arg Thr Thr Glu Glu Asp Arg
500 505 510

Gly Asn Tyr Tyr Cys Val Val Ser Ala Trp Thr Lys Gln Arg Asn
515 520 525

Asn Ser Trp Val Lys Ser Lys Asp Val Phe Ser Lys Pro Val Asn
530 535 540

Ile Phe Trp Ala Leu Glu Asp Ser Val Leu Val Val Lys Ala Arg
545 550 555

Gln Pro Lys Pro Phe Phe Ala Ala Gly Asn Thr Phe Glu Met Thr
560 565 570

Cys Lys Val Ser Ser Lys Asn Ile Lys Ser Pro Arg Tyr Ser Val
575 580 585

Leu Ile Met Ala Glu Lys Pro Val Gly Asp Leu Ser Ser Pro Asn
590 595 600

Glu Thr Lys Tyr Ile Ile Ser Leu Asp Gln Asp Ser Val Val Lys
605 610 615

Leu Glu Asn Trp Thr Asp Ala Ser Arg Val Asp Gly Val Val Leu
620 625 630

Glu Lys Val Gln Glu Asp Glu Phe Arg Tyr Arg Met Tyr Gln Thr
635 640 645

Gln Val Ser Asp Ala Gly Leu Tyr Arg Cys Met Val Thr Ala Trp
650 655 660

Ser Pro Val Arg Gly Ser Leu Trp Arg Glu Ala Ala Thr Ser Leu
665 670 675

Ser Asn Pro Ile Glu Ile Asp Phe Gln Thr Ser Gly Pro Ile Phe
680 685 690

Asn Ala Ser Val His Ser Asp Thr Pro Ser Val Ile Arg Gly Asp
695 700 705

Leu Ile Lys Leu Phe Cys Ile Ile Thr Val Glu Gly Ala Ala Leu
710 715 720

Asp Pro Asp Asp Met Ala Phe Asp Val Ser Trp Phe Ala Val His
725 730 735

Ser Phe Gly Leu Asp Lys Ala Pro Val Leu Leu Ser Ser Leu Asp
740 745 750

Arg Lys Gly Ile Val Thr Thr Ser Arg Arg Asp Trp Lys Ser Asp
755 760 765

Leu Ser Leu Glu Arg Val Ser Val Leu Glu Phe Leu Leu Gln Val

1304
1300

770 775 780

His Gly Ser Glu Asp Gln Asp Phe Gly Asn Tyr Tyr Cys Ser Val
785 790 795

Thr Pro Trp Val Lys Ser Pro Thr Gly Ser Trp Gln Lys Glu Ala
800 805 810

Glu Ile His Ser Lys Pro Val Phe Ile Thr Val Lys Met Asp Val
815 820 825

Leu Asn Ala Phe Lys Tyr Pro Leu Leu Ile Gly Ile Gly Leu Ser
830 835 840

Thr Val Ile Gly Leu Leu Ser Cys Leu Ile Gly Tyr Cys Ser Ser
845 850 855

His Trp Cys Cys Lys Lys Glu Val Gln Glu Thr Arg Arg Glu Arg
860 865 870

Arg Arg Leu Met Ser Met Glu Met Asp
875

<210> 137
<211> 434
<212> PRT
<213> Homo sapiens

<400> 137
Met Ser Arg Ser Arg His Leu Gly Lys Ile Arg Lys Arg Leu Glu
1 5 10 15

Asp Val Lys Ser Gln Trp Val Arg Pro Ala Arg Ala Asp Phe Ser
20 25 30

Asp Asn Glu Ser Ala Arg Leu Ala Thr Asp Ala Leu Leu Asp Gly
35 40 45

Gly Ser Glu Ala Tyr Trp Arg Val Leu Ser Gln Glu Gly Glu Val
50 55 60

Asp Phe Leu Ser Ser Val Glu Ala Gln Tyr Ile Gln Ala Gln Ala
65 70 75

Arg Glu Pro Pro Cys Pro Pro Asp Thr Leu Gly Gly Ala Glu Ala
80 85 90

Gly Pro Lys Gly Leu Asp Ser Ser Ser Leu Gln Ser Gly Thr Tyr
95 100 105

Phe Pro Val Ala Ser Glu Gly Ser Glu Pro Ala Leu Leu His Ser
110 115 120

Trp Ala Ser Ala Glu Lys Pro Tyr Leu Lys Glu Lys Ser Ser Ala
125 130 135

Thr Val Tyr Phe Gln Thr Val Lys His Asn Asn Ile Arg Asp Leu
140 145 150

Val Arg Arg Cys Ile Thr Arg Thr Ser Gln Val Leu Val Ile Leu
155 160 165

Met Asp Val Phe Thr Asp Val Glu Ile Phe Cys Asp Ile Leu Glu
170 175 180

Ala Ala Asn Lys Arg Gly Val Phe Val Cys Val Leu Leu Asp Gln
185 190 195

Gly Gly Val Lys Leu Phe Gln Glu Met Cys Asp Lys Val Gln Ile
200 205 210

Ser Asp Ser His Leu Lys Asn Ile Ser Ile Arg Ser Val Glu Gly
215 220 225

Glu Ile Tyr Cys Ala Lys Ser Gly Arg Lys Phe Ala Gly Gln Ile
230 235 240

Arg Glu Lys Phe Ile Ile Ser Asp Trp Arg Phe Val Leu Ser Gly
245 250 255

Ser Tyr Ser Phe Thr Trp Leu Cys Gly His Val His Arg Asn Ile
260 265 270

Leu Ser Lys Phe Thr Gly Gln Ala Val Glu Leu Phe Asp Glu Glu
275 280 285

Phe Arg His Leu Tyr Ala Ser Ser Lys Pro Val Met Gly Leu Lys
290 295 300

Ser Pro Arg Leu Val Ala Pro Val Pro Pro Gly Ala Ala Pro Ala
305 310 315

Asn Gly Arg Leu Ser Ser Ser Ser Gly Ser Ala Ser Asp Arg Thr
320 325 330

Ser Ser Asn Pro Phe Ser Gly Arg Ser Ala Gly Ser His Pro Gly
335 340 345

Thr Arg Ser Val Ser Ala Ser Ser Gly Pro Cys Ser Pro Ala Ala
350 355 360

Pro His Pro Pro Pro Pro Pro Arg Phe Gln Pro His Gln Gly Pro
365 370 375

Trp Gly Ala Pro Ser Pro Gln Ala His Leu Ser Pro Arg Pro His
380 385 390

Asp Gly Pro Pro Ala Ala Val Tyr Ser Asn Leu Gly Ala Tyr Arg
395 400 405

Pro Thr Arg Leu Gln Leu Glu Gln Leu Gly Leu Val Pro Arg Leu
410 415 420

Thr Pro Thr Trp Arg Pro Phe Leu Gln Ala Ser Pro His Phe
425 430

<210> 138
<211> 933
<212> PRT
<213> Homo sapiens

<400> 138
Met Ala Leu Pro Gln Pro Pro Pro Pro Ala Gly Ala Arg Asp Pro
1 5 10 15

Val Thr Pro Glu His Trp Ala Ser Gly Pro Pro Ser Gln Ala Gln
20 25 30

Pro Leu Leu Arg Gln Glu Ala Lys Glu Glu Glu Glu Gly Glu Glu
35 40 45

Thr Gly Val Gln Gly Ala Trp Gly Thr Gly Thr Ala Glu Gln Arg
50 55 60

Arg Arg Gly Trp Gly Glu Ala Ala Glu Ser Ala Ala Ala Glu Glu
65 70 75

Gly Gln Ala Glu Val Gly Gly Ala Ala Ala Ala Gly Ser Gly Ser
80 85 90

Pro Ala Gly Gly Ala Gly Gly Gly Leu Gly Ser Trp Arg Pro Leu
95 100 105

Leu Ala Trp Leu Gln Arg Arg Gln Pro Gln Cys Cys Pro Cys Ala
110 115 120

Ala Pro Leu Ser Arg Ser Ala Ala His Cys Cys His Gly Gly Thr
125 130 135

Lys Met Ala Ala Leu Ala Tyr Asn Leu Gly Lys Arg Glu Ile Asn
140 145 150

```
His Tyr Phe Ser Val Arg Ser Ala Lys Val Leu Ala Leu Val Ala
                155                 160                 165

Val Leu Leu Leu Ala Ala Cys His Leu Ala Ser Arg Arg Tyr Arg
                170                 175                 180

Gly Asn Asp Ser Cys Glu Tyr Leu Leu Ser Ser Gly Arg Phe Leu
                185                 190                 195

Gly Glu Lys Val Trp Gln Pro His Ser Cys Met Met His Lys Tyr
                200                 205                 210

Lys Ile Ser Glu Ala Lys Asn Cys Leu Val Asp Lys His Ile Ala
                215                 220                 225

Phe Ile Gly Asp Ser Arg Ile Arg Gln Leu Phe Tyr Ser Phe Val
                230                 235                 240

Lys Ile Ile Asn Pro Gln Phe Lys Glu Glu Gly Asn Lys His Glu
                245                 250                 255

Asn Ile Pro Phe Glu Asp Lys Thr Ala Ser Val Lys Val Asp Phe
                260                 265                 270

Leu Trp His Pro Glu Val Asn Gly Ser Met Lys Gln Cys Ile Lys
                275                 280                 285

Val Trp Thr Glu Asp Ser Ile Ala Lys Pro His Val Ile Val Ala
                290                 295                 300

Gly Ala Ala Thr Trp Ser Ile Lys Ile His Asn Gly Ser Ser Glu
                305                 310                 315

Ala Leu Ser Gln Tyr Lys Met Asn Ile Thr Ser Ile Ala Pro Leu
                320                 325                 330

Leu Glu Lys Leu Ala Lys Thr Ser Asp Val Tyr Trp Val Leu Gln
                335                 340                 345

Asp Pro Val Tyr Glu Asp Leu Leu Ser Glu Asn Arg Lys Met Ile
                350                 355                 360

Thr Asn Glu Lys Ile Asp Ala Tyr Asn Glu Ala Ala Val Ser Ile
                365                 370                 375

Leu Asn Ser Ser Thr Arg Asn Ser Lys Ser Asn Val Lys Met Phe
                380                 385                 390

Ser Val Ser Lys Leu Ile Ala Gln Glu Thr Ile Met Glu Ser Leu
                395                 400                 405

Asp Gly Leu His Leu Pro Glu Ser Ser Arg Glu Thr Thr Ala Met
                410                 415                 420

Ile Leu Met Asn Val Tyr Cys Asn Lys Ile Leu Lys Pro Val Asp
                425                 430                 435

Gly Ser Cys Cys Gln Pro Arg Pro Pro Val Thr Leu Ile Gln Lys
                440                 445                 450

Leu Ala Ala Cys Phe Phe Thr Leu Ser Ile Ile Gly Tyr Leu Ile
                455                 460                 465

Phe Tyr Ile Ile His Arg Asn Ala His Arg Lys Asn Lys Pro Cys
                470                 475                 480

Thr Asp Leu Glu Ser Gly Glu Glu Lys Lys Asn Ile Ile Asn Thr
                485                 490                 495

Pro Val Ser Ser Leu Glu Ile Leu Leu Gln Ser Phe Cys Lys Leu
                500                 505                 510

Gly Leu Ile Met Ala Tyr Phe Tyr Met Cys Asp Arg Ala Asn Leu
                515                 520                 525

Phe Met Lys Glu Asn Lys Phe Tyr Thr His Ser Ser Phe Phe Ile
```

530 535 540

Pro Ile Ile Tyr Ile Leu Val Leu Gly Val Phe Tyr Asn Glu Asn
545 550 555

Thr Lys Glu Thr Lys Val Leu Asn Arg Glu Gln Thr Asp Glu Trp
560 565 570

Lys Gly Trp Met Gln Leu Val Ile Leu Ile Tyr His Ile Ser Gly
575 580 585

Ala Ser Thr Phe Leu Pro Val Tyr Met His Ile Arg Val Leu Val
590 595 600

Ala Ala Tyr Leu Phe Gln Thr Gly Tyr Gly His Phe Ser Tyr Phe
605 610 615

Trp Ile Lys Gly Asp Phe Gly Ile Tyr Arg Val Cys Gln Val Leu
620 625 630

Phe Arg Leu Asn Phe Leu Val Val Val Leu Cys Ile Val Met Asp
635 640 645

Arg Pro Tyr Gln Phe Tyr Tyr Phe Val Pro Leu Val Thr Val Trp
650 655 660

Phe Met Val Ile Tyr Val Thr Leu Ala Leu Trp Pro Gln Ile Ile
665 670 675

Gln Lys Lys Ala Asn Gly Asn Cys Phe Trp His Phe Gly Leu Leu
680 685 690

Leu Lys Leu Gly Phe Leu Leu Leu Phe Ile Cys Phe Leu Ala Tyr
695 700 705

Ser Gln Gly Ala Phe Glu Lys Ile Phe Ser Leu Trp Pro Leu Ser
710 715 720

Lys Cys Phe Glu Leu Lys Gly Asn Val Tyr Glu Trp Trp Phe Arg
725 730 735

Trp Arg Leu Asp Arg Tyr Val Val Phe His Gly Met Leu Phe Ala
740 745 750

Phe Ile Tyr Leu Ala Leu Gln Lys Arg Gln Ile Leu Ser Glu Gly
755 760 765

Lys Gly Glu Pro Leu Phe Ser Asn Lys Ile Ser Asn Phe Leu Leu
770 775 780

Phe Ile Ser Val Val Ser Phe Leu Thr Tyr Ser Ile Trp Ala Ser
785 790 795

Ser Cys Lys Asn Lys Ala Glu Cys Asn Glu Leu His Pro Ser Val
800 805 810

Ser Val Val Gln Ile Leu Ala Phe Ile Leu Ile Arg Asn Ile Pro
815 820 825

Gly Tyr Ala Arg Ser Val Tyr Ser Ser Phe Phe Ala Trp Phe Gly
830 835 840

Lys Ile Ser Leu Glu Leu Phe Ile Cys Gln Tyr His Ile Trp Leu
845 850 855

Ala Ala Asp Thr Arg Gly Ile Leu Val Leu Ile Pro Gly Asn Pro
860 865 870

Met Leu Asn Ile Ile Val Ser Thr Phe Ile Phe Val Cys Val Ala
875 880 885

His Glu Ile Ser Gln Ile Thr Asn Asp Leu Ala Gln Ile Ile Ile
890 895 900

Pro Lys Asp Asn Ser Ser Leu Leu Lys Arg Leu Ala Cys Ile Ala
905 910 915

Ala Phe Phe Cys Gly Leu Leu Ile Leu Ser Ser Ile Gln Asp Lys
920 925 930

Ser Lys His

<210> 139
<211> 320
<212> PRT
<213> Homo sapiens

<400> 139
Met Tyr His Cys His Ser Gly Ser Lys Pro Thr Glu Lys Gly Ala
1 5 10 15

Asn Glu Tyr Ala Tyr Ala Lys Trp Lys Leu Cys Ser Ala Ser Ala
20 25 30

Ile Cys Phe Ile Phe Met Ile Ala Glu Val Val Gly Gly His Ile
35 40 45

Ala Gly Ser Leu Ala Val Val Thr Asp Ala Ala His Leu Leu Ile
50 55 60

Asp Leu Thr Ser Phe Leu Leu Ser Leu Phe Ser Leu Trp Leu Ser
65 70 75

Ser Lys Pro Pro Ser Lys Arg Leu Thr Phe Gly Trp His Arg Ala
80 85 90

Glu Ile Leu Gly Ala Leu Leu Ser Ile Leu Cys Ile Trp Val Val
95 100 105

Thr Gly Val Leu Val Tyr Leu Ala Cys Glu Arg Leu Leu Tyr Pro
110 115 120

Asp Tyr Gln Ile Gln Ala Thr Val Met Ile Ile Val Ser Ser Cys
125 130 135

Ala Val Ala Ala Asn Ile Val Leu Thr Val Val Leu His Gln Arg
140 145 150

Cys Leu Gly His Asn His Lys Glu Val Gln Ala Asn Ala Ser Val
155 160 165

Arg Ala Ala Phe Val His Ala Leu Gly Asp Leu Phe Gln Ser Ile
170 175 180

Ser Val Leu Ile Ser Ala Leu Ile Ile Tyr Phe Lys Pro Glu Tyr
185 190 195

Lys Ile Ala Asp Pro Ile Cys Thr Phe Ile Phe Ser Ile Leu Val
200 205 210

Leu Ala Ser Thr Ile Thr Ile Leu Lys Asp Phe Ser Ile Leu Leu
215 220 225

Met Glu Gly Val Pro Lys Ser Leu Asn Tyr Ser Gly Val Lys Glu
230 235 240

Leu Ile Leu Ala Val Asp Gly Val Leu Ser Val His Ser Leu His
245 250 255

Ile Trp Ser Leu Thr Met Asn Gln Val Ile Leu Ser Ala His Val
260 265 270

Ala Thr Ala Ala Ser Arg Asp Ser Gln Val Val Arg Arg Glu Ile
275 280 285

Ala Lys Ala Leu Ser Lys Ser Phe Thr Met His Ser Leu Thr Ile
290 295 300

Gln Met Glu Ser Pro Val Asp Gln Asp Pro Asp Cys Leu Phe Cys
305 310 315

Glu Asp Pro Cys Asp
320

<210> 140
<211> 426
<212> PRT
<213> Homo sapiens

<400> 140
Met Val Leu Ser Val Pro Val Ile Ala Leu Gly Ala Thr Leu Gly
1 5 10 15

Thr Ala Thr Ser Ile Leu Ala Leu Cys Gly Val Thr Cys Leu Cys
20 25 30

Arg His Met His Pro Lys Lys Gly Leu Leu Pro Arg Asp Gln Asp
35 40 45

Pro Asp Leu Glu Lys Ala Lys Pro Ser Leu Leu Gly Ser Ala Gln
50 55 60

Gln Phe Asn Val Lys Lys Ser Thr Glu Pro Val Gln Pro Arg Ala
65 70 75

Leu Leu Lys Phe Pro Asp Ile Tyr Gly Pro Arg Pro Ala Val Thr
80 85 90

Ala Pro Glu Val Ile Asn Tyr Ala Asp Tyr Ser Leu Arg Ser Thr
95 100 105

Glu Glu Pro Thr Ala Pro Ala Ser Pro Gln Pro Pro Asn Asp Ser
110 115 120

Arg Leu Lys Arg Gln Val Thr Glu Glu Leu Phe Ile Leu Pro Gln
125 130 135

Asn Gly Val Val Glu Asp Val Cys Val Met Glu Thr Trp Asn Pro
140 145 150

Glu Lys Ala Ala Ser Trp Asn Gln Ala Pro Lys Leu His Tyr Cys
155 160 165

Leu Asp Tyr Asp Cys Gln Lys Ala Glu Leu Phe Val Thr Arg Leu
170 175 180

Glu Ala Val Thr Ser Asn His Asp Gly Gly Cys Asp Cys Tyr Val
185 190 195

Gln Gly Ser Val Ala Asn Arg Thr Gly Ser Val Glu Ala Gln Thr
200 205 210

Ala Leu Lys Lys Arg Gln Leu His Thr Thr Trp Glu Glu Gly Leu
215 220 225

Val Leu Pro Leu Ala Glu Glu Glu Leu Pro Thr Ala Thr Leu Thr
230 235 240

Leu Thr Leu Arg Thr Cys Asp Arg Phe Ser Arg His Ser Val Ala
245 250 255

Gly Glu Leu Arg Leu Gly Leu Asp Gly Thr Ser Val Pro Leu Gly
260 265 270

Ala Ala Gln Trp Gly Glu Leu Lys Thr Ser Ala Lys Glu Pro Ser
275 280 285

Ala Gly Ala Gly Glu Val Leu Leu Ser Ile Ser Tyr Leu Pro Ala
290 295 300

Ala Asn Arg Leu Leu Val Val Leu Ile Lys Ala Lys Asn Leu His
305 310 315

Ser Asn Gln Ser Lys Glu Leu Leu Gly Lys Asp Val Ser Val Lys
320 325 330

Val Thr Leu Lys His Gln Ala Arg Lys Leu Lys Lys Lys Gln Thr

335 340 345

Lys Arg Ala Lys His Lys Ile Asn Pro Val Trp Asn Glu Met Ile
350 355 360

Met Phe Glu Leu Pro Asp Asp Leu Leu Gln Ala Ser Ser Val Glu
365 370 375

Leu Glu Val Leu Gly Gln Asp Asp Ser Gly Gln Ser Cys Ala Leu
380 385 390

Gly His Cys Ser Leu Gly Leu His Thr Ser Gly Ser Glu Arg Ser
395 400 405

His Trp Glu Glu Met Leu Lys Asn Pro Arg Arg Gln Ile Ala Met
410 415 420

Trp His Gln Leu His Leu
425

<210> 141
<211> 883
<212> PRT
<213> Homo sapiens

<400> 141
Met Gln Lys Ile Met His Ile Ser Val Leu Leu Ser Pro Val Leu
1 5 10 15

Trp Gly Leu Ile Phe Gly Val Ser Ser Asn Ser Ile Gln Ile Gly
20 25 30

Gly Leu Phe Pro Arg Gly Ala Asp Gln Glu Tyr Ser Ala Phe Arg
35 40 45

Val Gly Met Val Gln Phe Ser Thr Ser Glu Phe Arg Leu Thr Pro
50 55 60

His Ile Asp Asn Leu Glu Val Ala Asn Ser Phe Ala Val Thr Asn
65 70 75

Ala Phe Cys Ser Gln Phe Ser Arg Gly Val Tyr Ala Ile Phe Gly
80 85 90

Phe Tyr Asp Lys Lys Ser Val Asn Thr Ile Thr Ser Phe Cys Gly
95 100 105

Thr Leu His Val Ser Phe Ile Thr Pro Ser Phe Pro Thr Asp Gly
110 115 120

Thr His Pro Phe Val Ile Gln Met Arg Pro Asp Leu Lys Gly Ala
125 130 135

Leu Leu Ser Leu Ile Glu Tyr Tyr Gln Trp Asp Lys Phe Ala Tyr
140 145 150

Leu Tyr Asp Ser Asp Arg Gly Leu Ser Thr Leu Gln Ala Val Leu
155 160 165

Asp Ser Ala Ala Glu Lys Lys Trp Gln Val Thr Ala Ile Asn Val
170 175 180

Gly Asn Ile Asn Asn Asp Lys Lys Asp Glu Met Tyr Arg Ser Leu
185 190 195

Phe Gln Asp Leu Glu Leu Lys Lys Glu Arg Arg Val Ile Leu Asp
200 205 210

Cys Glu Arg Asp Lys Val Asn Asp Ile Val Asp Gln Val Ile Thr
215 220 225

Ile Gly Lys His Val Lys Gly Tyr His Tyr Ile Ile Ala Asn Leu
230 235 240

Glu Phe Thr Asp Gly Asp Leu Leu Lys Ile Gln Phe Gly Gly Ala
245 250 255

```
Asn Val Ser Gly Phe Gln Ile Val Asp Tyr Asp Asp Ser Leu Val
                260                 265                 270

Ser Lys Phe Ile Glu Arg Trp Ser Thr Leu Glu Glu Lys Glu Tyr
                275                 280                 285

Pro Gly Ala His Thr Thr Thr Ile Lys Tyr Thr Ser Ala Leu Thr
                290                 295                 300

Tyr Asp Ala Val Gln Val Met Thr Glu Ala Phe Arg Asn Leu Arg
                305                 310                 315

Lys Gln Arg Ile Glu Ile Ser Arg Arg Gly Asn Ala Gly Asp Cys
                320                 325                 330

Leu Ala Asn Pro Ala Val Pro Trp Gly Gln Gly Val Glu Ile Glu
                335                 340                 345

Arg Ala Leu Lys Gln Val Gln Val Glu Gly Leu Ser Gly Asn Ile
                350                 355                 360

Lys Phe Asp Gln Asn Gly Lys Arg Ile Asn Tyr Thr Ile Asn Ile
                365                 370                 375

Met Glu Leu Lys Thr Asn Gly Pro Arg Lys Ile Gly Tyr Trp Ser
                380                 385                 390

Glu Val Asp Lys Met Val Val Thr Leu Thr Glu Leu Pro Ser Gly
                395                 400                 405

Asn Asp Thr Ser Gly Leu Glu Asn Lys Thr Val Val Val Thr Thr
                410                 415                 420

Ile Leu Glu Ser Pro Tyr Val Met Met Lys Lys Asn His Glu Met
                425                 430                 435

Leu Glu Gly Asn Glu Arg Tyr Glu Gly Tyr Cys Val Asp Leu Ala
                440                 445                 450

Ala Glu Ile Ala Lys His Cys Gly Phe Lys Tyr Lys Leu Thr Ile
                455                 460                 465

Val Gly Asp Gly Lys Tyr Gly Ala Arg Asp Ala Asp Thr Lys Ile
                470                 475                 480

Trp Asn Gly Met Val Gly Glu Leu Val Tyr Gly Lys Ala Asp Ile
                485                 490                 495

Ala Ile Ala Pro Leu Thr Ile Thr Leu Val Arg Glu Glu Val Ile
                500                 505                 510

Asp Phe Ser Lys Pro Phe Met Ser Leu Gly Ile Ser Ile Met Ile
                515                 520                 525

Lys Lys Pro Gln Lys Ser Lys Pro Gly Val Phe Ser Phe Leu Asp
                530                 535                 540

Pro Leu Ala Tyr Glu Ile Trp Met Cys Ile Val Phe Ala Tyr Ile
                545                 550                 555

Gly Val Ser Val Val Leu Phe Leu Val Ser Arg Phe Ser Pro Tyr
                560                 565                 570

Glu Trp His Thr Glu Glu Phe Glu Asp Gly Arg Glu Thr Gln Ser
                575                 580                 585

Ser Glu Ser Thr Asn Glu Phe Gly Ile Phe Asn Ser Leu Trp Phe
                590                 595                 600

Ser Leu Gly Ala Phe Met Arg Gln Gly Cys Asp Ile Ser Pro Arg
                605                 610                 615

Ser Leu Ser Gly Arg Ile Val Gly Gly Val Trp Trp Phe Phe Thr
                620                 625                 630
```

Leu Ile Ile Ile Ser Ser Tyr Thr Ala Asn Leu Ala Ala Phe Leu
635 640 645

Thr Val Glu Arg Met Val Ser Pro Ile Glu Ser Ala Glu Asp Leu
650 655 660

Ser Lys Gln Thr Glu Ile Ala Tyr Gly Thr Leu Asp Ser Gly Ser
665 670 675

Thr Lys Glu Phe Phe Arg Arg Ser Lys Ile Ala Val Phe Asp Lys
680 685 690

Met Trp Thr Tyr Met Arg Ser Ala Glu Pro Ser Val Phe Val Arg
695 700 705

Thr Thr Ala Glu Gly Val Ala Arg Val Arg Lys Ser Lys Gly Lys
710 715 720

Tyr Ala Tyr Leu Leu Glu Ser Thr Met Asn Glu Tyr Ile Glu Gln
725 730 735

Arg Lys Pro Cys Asp Thr Met Lys Val Gly Gly Asn Leu Asp Ser
740 745 750

Lys Gly Tyr Gly Ile Ala Thr Pro Lys Gly Ser Ser Leu Gly Thr
755 760 765

Pro Val Asn Leu Ala Val Leu Lys Leu Ser Glu Gln Gly Val Leu
770 775 780

Asp Lys Leu Lys Asn Lys Trp Trp Tyr Asp Lys Gly Glu Cys Gly
785 790 795

Ala Lys Asp Ser Gly Ser Lys Glu Lys Thr Ser Ala Leu Ser Leu
800 805 810

Ser Asn Val Ala Gly Val Phe Tyr Ile Leu Val Gly Gly Leu Gly
815 820 825

Leu Ala Met Leu Val Ala Leu Ile Glu Phe Cys Tyr Lys Ser Arg
830 835 840

Ala Glu Ala Lys Arg Met Lys Val Ala Lys Asn Ala Gln Asn Ile
845 850 855

Asn Pro Ser Ser Ser Gln Asn Ser Gln Asn Phe Ala Thr Tyr Lys
860 865 870

Glu Gly Tyr Asn Val Tyr Gly Ile Glu Ser Val Lys Ile
875 880

<210> 142
<211> 357
<212> PRT
<213> Homo sapiens

<400> 142
Met Ala Thr Thr Val Pro Asp Gly Cys Arg Asn Gly Leu Lys Ser
1 5 10 15

Lys Tyr Tyr Arg Leu Cys Asp Lys Ala Glu Ala Trp Gly Ile Val
20 25 30

Leu Glu Thr Val Ala Thr Ala Gly Val Val Thr Ser Val Ala Phe
35 40 45

Met Leu Thr Leu Pro Ile Leu Val Cys Lys Val Gln Asp Ser Asn
50 55 60

Arg Arg Lys Met Leu Pro Thr Gln Phe Leu Phe Leu Leu Gly Val
65 70 75

Leu Gly Ile Phe Gly Leu Thr Phe Ala Phe Ile Ile Gly Leu Asp
80 85 90

Gly Ser Thr Gly Pro Thr Arg Phe Phe Leu Phe Gly Ile Leu Phe

95 100 105

Ser Ile Cys Phe Ser Cys Leu Leu Ala His Ala Val Ser Leu Thr
110 115 120

Lys Leu Val Arg Gly Arg Lys Pro Leu Ser Leu Leu Val Ile Leu
125 130 135

Gly Leu Ala Val Gly Phe Ser Leu Val Gln Asp Val Ile Ala Ile
140 145 150

Glu Tyr Ile Val Leu Thr Met Asn Arg Thr Asn Val Asn Val Phe
155 160 165

Ser Glu Leu Ser Ala Pro Arg Arg Asn Glu Asp Phe Val Leu Leu
170 175 180

Leu Thr Tyr Val Leu Phe Leu Met Ala Leu Thr Phe Leu Met Ser
185 190 195

Ser Phe Thr Phe Cys Gly Ser Phe Thr Gly Trp Lys Arg His Gly
200 205 210

Ala His Ile Tyr Leu Thr Met Leu Leu Ser Ile Ala Ile Trp Val
215 220 225

Ala Trp Ile Thr Leu Leu Met Leu Pro Asp Phe Asp Arg Arg Trp
230 235 240

Asp Asp Thr Ile Leu Ser Ser Ala Leu Ala Ala Asn Gly Trp Val
245 250 255

Phe Leu Leu Ala Tyr Val Ser Pro Glu Phe Trp Leu Leu Thr Lys
260 265 270

Gln Arg Asn Pro Met Asp Tyr Pro Val Glu Asp Ala Phe Cys Lys
275 280 285

Pro Gln Leu Val Lys Lys Ser Tyr Gly Val Glu Asn Arg Ala Tyr
290 295 300

Ser Gln Glu Glu Ile Thr Gln Gly Phe Glu Glu Thr Gly Asp Thr
305 310 315

Leu Tyr Ala Pro Tyr Ser Thr His Phe Gln Leu Gln Asn Gln Pro
320 325 330

Pro Gln Lys Glu Phe Ser Ile Pro Arg Ala His Ala Trp Pro Ser
335 340 345

Pro Tyr Lys Asp Tyr Glu Val Lys Lys Glu Gly Ser
350 355

<210> 143
<211> 222
<212> PRT
<213> Homo sapiens

<400> 143
Met Leu Pro Ala Gln Glu Ala Ala Lys Leu Tyr His Thr Asn Tyr
1 5 10 15

Val Arg Asn Ser Arg Ala Ile Gly Val Leu Trp Ala Ile Phe Thr
20 25 30

Ile Cys Phe Ala Ile Val Asn Val Val Cys Phe Ile Gln Pro Tyr
35 40 45

Trp Ile Gly Asp Gly Val Asp Thr Pro Gln Ala Gly Tyr Phe Gly
50 55 60

Leu Phe His Tyr Cys Ile Gly Asn Gly Phe Ser Arg Glu Leu Thr
65 70 75

Cys Arg Gly Ser Phe Thr Asp Phe Ser Thr Leu Pro Ser Gly Ala
80 85 90

Phe Lys Ala Ala Ser Phe Phe Ile Gly Leu Ser Met Met Leu Ile
95 100 105

Ile Ala Cys Ile Ile Cys Phe Thr Leu Phe Phe Phe Cys Asn Thr
110 115 120

Ala Thr Val Tyr Lys Ile Cys Ala Trp Met Gln Leu Thr Ser Ala
125 130 135

Ala Cys Leu Val Leu Gly Cys Met Ile Phe Pro Asp Gly Trp Asp
140 145 150

Ser Asp Glu Val Lys Arg Met Cys Gly Glu Lys Thr Asp Lys Tyr
155 160 165

Thr Leu Gly Ala Cys Ser Val Arg Trp Ala Tyr Ile Leu Ala Ile
170 175 180

Ile Gly Ile Leu Asp Ala Leu Ile Leu Ser Phe Leu Ala Phe Val
185 190 195

Leu Gly Asn Arg Gln Asp Ser Leu Met Ala Glu Glu Leu Lys Ala
200 205 210

Glu Asn Lys Val Leu Leu Ser Gln Tyr Ser Leu Glu
215 220

<210> 144
<211> 236
<212> PRT
<213> Homo sapiens

<400> 144
Met Pro Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Met
1 5 10 15

Leu Pro Ala Gln Glu Ala Ala Lys Leu Tyr His Thr Asn Tyr Val
20 25 30

Arg Asn Ser Arg Ala Ile Gly Val Leu Trp Ala Ile Phe Thr Ile
35 40 45

Cys Phe Ala Ile Val Asn Val Val Cys Phe Ile Gln Pro Tyr Trp
50 55 60

Ile Gly Asp Gly Val Asp Thr Pro Gln Ala Gly Tyr Phe Gly Leu
65 70 75

Phe His Tyr Cys Ile Gly Asn Gly Phe Ser Arg Glu Leu Thr Cys
80 85 90

Arg Gly Ser Phe Thr Asp Phe Ser Thr Leu Pro Ser Gly Ala Phe
95 100 105

Lys Ala Ala Ser Phe Phe Ile Gly Leu Ser Met Met Leu Ile Ile
110 115 120

Ala Cys Ile Ile Cys Phe Thr Leu Phe Phe Phe Cys Asn Thr Ala
125 130 135

Thr Val Tyr Lys Ile Cys Ala Trp Met Gln Leu Thr Ser Ala Ala
140 145 150

Cys Leu Val Leu Gly Cys Met Ile Phe Pro Asp Gly Trp Asp Ser
155 160 165

Asp Glu Val Lys Arg Met Cys Gly Glu Lys Thr Asp Lys Tyr Thr
170 175 180

Leu Gly Ala Cys Ser Val Arg Trp Ala Tyr Ile Leu Ala Ile Ile
185 190 195

Gly Ile Leu Asp Ala Leu Ile Leu Ser Phe Leu Ala Phe Val Leu
200 205 210

Gly Asn Arg Gln Asp Ser Leu Met Ala Glu Glu Leu Lys Ala Glu
215 220 225

Asn Lys Val Leu Leu Ser Gln Tyr Ser Leu Glu
230 235

<210> 145
<211> 236
<212> PRT
<213> Homo sapiens

<400> 145
Met Pro Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Met
1 5 10 15

Leu Pro Ala Gln Glu Ala Ala Lys Leu Tyr His Thr Asn Tyr Val
20 25 30

Arg Asn Ser Arg Ala Ile Gly Val Leu Trp Ala Ile Phe Thr Ile
35 40 45

Cys Phe Ala Ile Val Asn Val Val Cys Phe Ile Gln Pro Tyr Trp
50 55 60

Ile Gly Asp Gly Val Asp Thr Pro Gln Ala Gly Tyr Phe Gly Leu
65 70 75

Phe His Tyr Cys Ile Gly Asn Gly Phe Ser Arg Glu Leu Thr Cys
80 85 90

Arg Gly Ser Phe Thr Asp Phe Ser Thr Leu Pro Ser Gly Ala Phe
95 100 105

Lys Ala Ala Ser Phe Phe Ile Gly Leu Ser Met Met Leu Ile Ile
110 115 120

Ala Cys Ile Ile Cys Phe Thr Leu Phe Phe Phe Cys Asn Thr Ala
125 130 135

Thr Val Tyr Lys Ile Cys Ala Trp Met Gln Leu Thr Ser Ala Ala
140 145 150

Cys Leu Val Leu Gly Cys Met Ile Phe Pro Asp Gly Trp Asp Ser
155 160 165

Asp Glu Val Lys Arg Met Cys Gly Glu Lys Thr Asp Lys Tyr Thr
170 175 180

Leu Gly Ala Cys Ser Val Arg Trp Ala Tyr Ile Leu Ala Ile Ile
185 190 195

Gly Ile Leu Asp Ala Leu Ile Leu Ser Phe Leu Ala Phe Val Leu
200 205 210

Gly Asn Arg Gln Asp Ser Leu Met Ala Glu Glu Leu Lys Ala Glu
215 220 225

Asn Lys Val Leu Leu Ser Gln Tyr Ser Leu Glu
230 235

<210> 146
<211> 408
<212> PRT
<213> Homo sapiens

<400> 146
Arg Pro Leu Arg Leu Pro Cys Arg Gly Pro Arg Ala Ala Gly Thr
1 5 10 15

Val Ser Ser Glu Ser Leu Gly His Pro Ala Pro Gly Gln Val Thr
20 25 30

Asn Arg Ser Pro Gln Phe Pro Met Glu Leu Phe Gln Ser Ser Pro
35 40 45

Arg Pro Pro Leu Gln Pro Val Cys Val Cys Pro Gly Leu Glu Leu

LSTG
1312

50 55 60

Cys Asp Pro Pro His Arg Leu Ala Phe Ser Met Gly Ala Gly Leu
65 70 75

Phe Ser Val Val Gly Thr Leu Leu Leu Pro Gly Leu Ala Ala Leu
80 85 90

Val Gln Asp Trp Arg Leu Leu Gln Gly Leu Gly Ala Leu Met Ser
95 100 105

Gly Leu Leu Leu Leu Phe Trp Gly Phe Pro Ala Leu Phe Pro Glu
110 115 120

Ser Pro Cys Trp Leu Leu Ala Thr Gly Gln Val Ala Arg Ala Arg
125 130 135

Lys Ile Leu Trp Arg Phe Ala Glu Ala Ser Gly Val Gly Pro Gly
140 145 150

Asp Ser Ser Leu Glu Glu Asn Ser Leu Ala Thr Glu Leu Thr Met
155 160 165

Leu Ser Ala Arg Ser Pro Gln Pro Arg Tyr His Ser Pro Leu Gly
170 175 180

Leu Leu Arg Thr Arg Val Thr Trp Arg Asn Gly Leu Ile Leu Gly
185 190 195

Phe Ser Ser Leu Val Gly Gly Gly Ile Arg Ala Ser Phe Arg Arg
200 205 210

Ser Leu Ala Pro Gln Val Pro Thr Phe Tyr Leu Pro Tyr Phe Leu
215 220 225

Glu Ala Gly Leu Glu Ala Ala Ala Leu Val Phe Leu Leu Leu Thr
230 235 240

Ala Asp Cys Cys Gly Arg Arg Pro Val Leu Leu Leu Gly Thr Met
245 250 255

Val Thr Gly Leu Ala Ser Leu Leu Leu Leu Ala Gly Ala Gln Tyr
260 265 270

Leu Pro Gly Trp Thr Val Leu Phe Leu Ser Val Leu Gly Leu Leu
275 280 285

Ala Ser Arg Ala Val Ser Ala Leu Ser Ser Leu Phe Ala Ala Glu
290 295 300

Val Phe Pro Thr Val Ile Arg Gly Ala Gly Leu Gly Leu Val Leu
305 310 315

Gly Ala Gly Phe Leu Gly Gln Ala Ala Gly Pro Leu Asp Thr Leu
320 325 330

His Gly Arg Gln Gly Phe Phe Leu Gln Gln Val Val Phe Ala Ser
335 340 345

Leu Ala Val Leu Ala Leu Leu Cys Val Leu Leu Leu Pro Glu Ser
350 355 360

Arg Ser Arg Gly Leu Pro Gln Ser Leu Gln Asp Ala Asp Arg Leu
365 370 375

Arg Arg Ser Pro Leu Leu Arg Gly Arg Pro Arg Gln Asp His Leu
380 385 390

Pro Leu Leu Pro Pro Ser Asn Ser Tyr Trp Ala Gly His Thr Pro
395 400 405

Glu Gln His

<210> 147
<211> 1152

<212> PRT
<213> Homo sapiens

<220>
<221> unsure
<222> 382
<223> unknown amino acid

<400> 147

```
Met Pro Phe Glu Gln Cys His Leu Ser Trp Lys Asn Val Arg Arg
  1               5                  10                  15

Leu Tyr Asn Ala Val Ser Lys Glu Leu Val Gly Glu Phe Leu Gln
             20                  25                  30

Phe Val Gln Leu Asp Lys Glu Ala Ser Asp Pro Phe Ser Leu Asn
             35                  40                  45

Glu Leu Leu Asp Glu Leu Ser Arg Lys Gln Lys Glu Glu Leu Trp
             50                  55                  60

Gln Arg Leu Lys Asn Leu Leu Thr Asp Val Leu Leu Glu Ser Pro
             65                  70                  75

Val Asp Gly Trp Gln Val Val Glu Ala Gln Gly Glu Asp Asn Met
             80                  85                  90

Glu Thr Glu His Gly Ser Lys Met Arg Lys Ser Ile Glu Ile Ile
             95                 100                 105

Tyr Ala Ile Thr Ser Val Ile Leu Ala Ser Val Ser Val Ile Asn
            110                 115                 120

Glu Ser Glu Asn Tyr Glu Ala Leu Leu Glu Cys Val Ile Ile Leu
            125                 130                 135

Asn Gly Ile Leu Tyr Ala Leu Pro Glu Ser Glu Arg Lys Leu Gln
            140                 145                 150

Ser Ser Ile Gln Asp Leu Cys Val Thr Trp Trp Glu Lys Gly Leu
            155                 160                 165

Pro Ala Lys Glu Asp Thr Gly Lys Thr Ala Phe Val Met Leu Leu
            170                 175                 180

Arg Arg Ser Leu Glu Thr Lys Thr Gly Ala Asp Val Cys Arg Leu
            185                 190                 195

Trp Arg Ile His Gln Ala Leu Tyr Cys Phe Asp Tyr Asp Leu Glu
            200                 205                 210

Glu Ser Gly Glu Ile Lys Asp Met Leu Leu Glu Cys Phe Ile Asn
            215                 220                 225

Ile Asn Tyr Ile Lys Lys Glu Glu Gly Arg Arg Phe Leu Ser Cys
            230                 235                 240

Leu Phe Asn Trp Asn Ile Asn Phe Ile Lys Met Ile His Gly Thr
            245                 250                 255

Ile Lys Asn Gln Leu Gln Gly Leu Gln Lys Ser Leu Met Val Tyr
            260                 265                 270

Ile Ala Glu Ile Tyr Phe Arg Ala Trp Lys Lys Ala Ser Gly Lys
            275                 280                 285

Ile Leu Glu Ala Ile Glu Asn Asp Cys Ile Gln Asp Phe Met Phe
            290                 295                 300

His Gly Ile His Leu Pro Arg Arg Ser Pro Val His Ser Lys Val
            305                 310                 315

Arg Glu Val Leu Ser Tyr Phe His His Gln Lys Lys Val Arg Gln
            320                 325                 330

Gly Val Glu Glu Met Leu Tyr Arg Leu Tyr Lys Pro Ile Leu Trp
```

335 340 345

Arg Gly Leu Lys Ala Arg Asn Ser Glu Val Arg Ser Asn Ala Ala
350 355 360

Leu Leu Phe Val Glu Ala Phe Pro Ile Arg Asp Pro Asn Leu His
365 370 375

Ala Ile Glu Met Asp Ser Xaa Ile Gln Lys Gln Phe Glu Glu Leu
380 385 390

Tyr Ser Leu Leu Glu Asp Pro Tyr Pro Met Val Arg Ser Thr Gly
395 400 405

Ile Leu Gly Val Cys Lys Ile Thr Ser Lys Tyr Trp Glu Met Met
410 415 420

Pro Pro Thr Ile Leu Ile Asp Leu Leu Lys Lys Val Thr Gly Glu
425 430 435

Leu Ala Phe Asp Thr Ser Ser Ala Asp Val Arg Cys Ser Val Phe
440 445 450

Lys Cys Leu Pro Met Ile Leu Asp Asn Lys Leu Ser His Pro Leu
455 460 465

Leu Glu Gln Leu Leu Pro Ala Leu Arg Tyr Ser Leu His Asp Asn
470 475 480

Ser Glu Lys Val Arg Val Ala Phe Val Asp Met Leu Leu Lys Ile
485 490 495

Lys Ala Val Arg Ala Ala Lys Phe Trp Lys Ile Cys Pro Met Glu
500 505 510

His Ile Leu Val Arg Leu Glu Thr Asp Ser Arg Pro Val Ser Arg
515 520 525

Arg Leu Val Ser Leu Ile Phe Asn Ser Phe Leu Pro Val Asn Gln
530 535 540

Pro Glu Glu Val Trp Cys Glu Arg Cys Val Thr Leu Val Gln Met
545 550 555

Asn His Ala Ala Ala Arg Arg Phe Tyr Gln Tyr Ala His Glu His
560 565 570

Thr Ala Cys Thr Asn Ile Ala Lys Leu Ile His Val Ile Arg His
575 580 585

Cys Leu Asn Ala Cys Ile Gln Arg Ala Val Arg Glu Pro Pro Glu
590 595 600

Asp Glu Glu Glu Glu Asp Gly Arg Glu Lys Glu Asn Val Thr Val
605 610 615

Leu Asp Lys Thr Leu Ser Val Asn Asp Val Ala Cys Met Ala Gly
620 625 630

Leu Leu Glu Ile Ile Val Ile Leu Trp Lys Ser Ile Asp Arg Ser
635 640 645

Met Glu Asn Asn Lys Glu Ala Lys Leu Tyr Thr Ile Asn Lys Phe
650 655 660

Ala Ser Val Leu Pro Glu Tyr Leu Lys Val Phe Lys Asp Asp Arg
665 670 675

Cys Lys Ile Pro Leu Phe Met Leu Met Ser Phe Met Pro Ala Ser
680 685 690

Ala Val Pro Pro Phe Ser Cys Gly Val Ile Ser Thr Leu Arg Ser
695 700 705

Arg Glu Glu Gly Ala Val Asp Lys Ser Tyr Cys Thr Leu Leu Asp
710 715 720

```
Cys Leu Cys Ser Trp Gly Gln Val Gly His Ile Leu Glu Leu Val
                725                 730                 735

Asp Asn Trp Leu Pro Thr Glu His Ala Gln Ala Lys Ser Asn Thr
                740                 745                 750

Ala Ser Lys Gly Arg Val Gln Ile His Asp Thr Arg Pro Val Lys
                755                 760                 765

Pro Glu Leu Ala Leu Val Tyr Ile Glu Tyr Leu Leu Thr His Pro
                770                 775                 780

Lys Asn Arg Glu Cys Leu Leu Ser Ala Pro Arg Lys Lys Leu Asn
                785                 790                 795

His Leu Leu Lys Ala Leu Glu Thr Ser Lys Ala Asp Leu Glu Ser
                800                 805                 810

Leu Leu Gln Thr Pro Gly Gly Lys Pro Arg Gly Phe Ser Glu Ala
                815                 820                 825

Ala Ala Pro Arg Ala Phe Gly Leu His Cys Arg Leu Ser Ile His
                830                 835                 840

Leu Gln His Lys Phe Cys Ser Glu Gly Lys Val Tyr Leu Ser Met
                845                 850                 855

Leu Glu Asp Thr Gly Phe Trp Leu Glu Ser Lys Ile Leu Ser Phe
                860                 865                 870

Ile Gln Asp Gln Glu Glu Asp Tyr Leu Lys Leu His Arg Val Ile
                875                 880                 885

Tyr Gln Gln Ile Ile Gln Thr Tyr Leu Thr Val Cys Lys Asp Val
                890                 895                 900

Val Met Val Gly Leu Gly Asp His Gln Phe Gln Met Gln Leu Leu
                905                 910                 915

Gln Arg Ser Leu Gly Ile Met Gln Thr Val Lys Gly Phe Phe Tyr
                920                 925                 930

Val Ser Leu Leu Leu Asp Ile Leu Lys Glu Ile Thr Gly Ser Ser
                935                 940                 945

Leu Ile Gln Lys Thr Asp Ser Asp Glu Glu Val Ala Met Leu Leu
                950                 955                 960

Asp Thr Val Gln Lys Val Phe Gln Lys Met Leu Glu Cys Ile Ala
                965                 970                 975

Arg Ser Phe Arg Lys Gln Pro Glu Glu Gly Leu Arg Leu Leu Tyr
                980                 985                 990

Ser Val Gln Arg Pro Leu His Glu Phe Ile Thr Ala Val Gln Ser
                995                1000                1005

Arg His Thr Asp Thr Pro Val His Arg Gly Val Leu Ser Thr Leu
               1010                1015                1020

Ile Ala Gly Pro Val Val Glu Ile Ser His Gln Leu Arg Lys Val
               1025                1030                1035

Ser Asp Val Glu Glu Leu Thr Pro Pro Glu His Leu Ser Asp Leu
               1040                1045                1050

Pro Pro Phe Ser Arg Cys Leu Ile Gly Ile Ile Ile Lys Ser Ser
               1055                1060                1065

Asn Val Val Arg Ser Phe Leu Asp Glu Leu Lys Ala Cys Val Ala
               1070                1075                1080

Ser Asn Asp Ile Glu Gly Ile Val Cys Leu Thr Ala Ala Val His
               1085                1090                1095
```

```
Ile Ile Leu Val Ile Asn Ala Gly Lys His Lys Ser Ser Lys Val
                1100                1105                1110

Arg Glu Val Ala Ala Thr Val His Arg Lys Leu Lys Thr Phe Met
                1115                1120                1125

Glu Ile Thr Leu Glu Glu Asp Ser Ile Glu Arg Phe Leu Tyr Glu
                1130                1135                1140

Ser Ser Ser Arg Thr Leu Gly Glu Leu Leu Asn Ser
                1145                1150
```

<210> 148
<211> 1143
<212> PRT
<213> Homo sapiens

<400> 148

```
Met Glu Lys Arg Glu Thr Phe Val Gln Ala Val Ser Lys Glu Leu
  1               5                  10                  15

Val Gly Glu Phe Leu Gln Phe Val Gln Leu Asp Lys Glu Ala Ser
                20                  25                  30

Asp Pro Phe Ser Leu Asn Glu Leu Leu Asp Glu Leu Ser Arg Lys
                35                  40                  45

Gln Lys Glu Glu Leu Trp Gln Arg Leu Lys Asn Leu Leu Thr Asp
                50                  55                  60

Val Leu Leu Glu Ser Pro Val Asp Gly Trp Gln Val Val Glu Ala
                65                  70                  75

Gln Gly Glu Asp Asn Met Glu Thr Glu His Gly Ser Lys Met Arg
                80                  85                  90

Lys Ser Ile Glu Ile Ile Tyr Ala Ile Thr Ser Val Ile Leu Ala
                95                  100                 105

Ser Val Ser Val Ile Asn Glu Ser Glu Asn Tyr Glu Ala Leu Leu
                110                 115                 120

Glu Cys Val Ile Ile Leu Asn Gly Ile Leu Tyr Ala Leu Pro Glu
                125                 130                 135

Ser Glu Arg Lys Leu Gln Ser Ser Ile Gln Asp Leu Cys Val Thr
                140                 145                 150

Trp Trp Glu Lys Gly Leu Pro Ala Lys Glu Asp Thr Gly Lys Thr
                155                 160                 165

Ala Phe Val Met Leu Leu Arg Arg Ser Leu Glu Thr Lys Thr Gly
                170                 175                 180

Ala Asp Val Cys Arg Leu Trp Arg Ile His Gln Ala Leu Tyr Cys
                185                 190                 195

Phe Asp Tyr Asp Leu Glu Glu Ser Gly Glu Ile Lys Asp Met Leu
                200                 205                 210

Leu Glu Cys Phe Ile Asn Ile Asn Tyr Ile Lys Lys Glu Glu Gly
                215                 220                 225

Arg Arg Phe Leu Ser Cys Leu Phe Asn Trp Asn Ile Asn Phe Ile
                230                 235                 240

Lys Met Ile His Gly Thr Ile Lys Asn Gln Leu Gln Gly Leu Gln
                245                 250                 255

Lys Ser Leu Met Val Tyr Ile Ala Glu Ile Tyr Phe Arg Ala Trp
                260                 265                 270

Lys Lys Ala Ser Gly Lys Ile Leu Glu Ala Ile Glu Asn Asp Cys
                275                 280                 285

Ile Gln Asp Phe Met Phe His Gly Ile His Leu Pro Arg Arg Ser
```

290 295 300

Pro Val His Ser Lys Val Arg Glu Val Leu Ser Tyr Phe His His
305 310 315

Gln Lys Lys Val Arg Gln Gly Val Glu Glu Met Leu Tyr Arg Leu
320 325 330

Tyr Lys Pro Ile Leu Trp Arg Gly Leu Lys Ala Arg Asn Ser Glu
335 340 345

Val Arg Ser Asn Ala Ala Leu Leu Phe Val Glu Ala Phe Pro Ile
350 355 360

Arg Asp Pro Asn Leu His Ala Ile Glu Met Asp Ser Glu Ile Gln
365 370 375

Lys Gln Phe Glu Glu Leu Tyr Ser Leu Leu Glu Asp Pro Tyr Pro
380 385 390

Met Val Arg Ser Thr Gly Ile Leu Gly Val Cys Lys Ile Thr Ser
395 400 405

Lys Tyr Trp Glu Met Met Pro Pro Thr Ile Leu Ile Asp Leu Leu
410 415 420

Lys Lys Val Thr Gly Glu Leu Ala Phe Asp Thr Ser Ser Ala Asp
425 430 435

Val Arg Cys Ser Val Phe Lys Cys Leu Pro Met Ile Leu Asp Asn
440 445 450

Lys Leu Ser His Pro Leu Leu Glu Gln Leu Leu Pro Ala Leu Arg
455 460 465

Tyr Ser Leu His Asp Asn Ser Glu Lys Val Arg Val Ala Phe Val
470 475 480

Asp Met Leu Leu Lys Ile Lys Ala Val Arg Ala Ala Lys Phe Trp
485 490 495

Lys Ile Cys Pro Met Glu His Ile Leu Val Arg Leu Glu Thr Asp
500 505 510

Ser Arg Pro Val Ser Arg Arg Leu Val Ser Leu Ile Phe Asn Ser
515 520 525

Phe Leu Pro Val Asn Gln Pro Glu Glu Val Trp Cys Glu Arg Cys
530 535 540

Val Thr Leu Val Gln Met Asn His Ala Ala Ala Arg Arg Phe Tyr
545 550 555

Gln Tyr Ala His Glu His Thr Ala Cys Thr Asn Ile Ala Lys Leu
560 565 570

Ile His Val Ile Arg His Cys Leu Asn Ala Cys Ile Gln Arg Ala
575 580 585

Val Arg Glu Pro Pro Glu Asp Glu Glu Glu Glu Asp Gly Arg Glu
590 595 600

Lys Glu Asn Val Thr Val Leu Asp Lys Thr Leu Ser Val Asn Asp
605 610 615

Val Ala Cys Met Ala Gly Leu Leu Glu Ile Ile Val Ile Leu Trp
620 625 630

Lys Ser Ile Asp Arg Ser Met Glu Asn Asn Lys Glu Ala Lys Leu
635 640 645

Tyr Thr Ile Asn Lys Phe Ala Ser Val Leu Pro Glu Tyr Leu Lys
650 655 660

Val Phe Lys Asp Asp Arg Cys Lys Ile Pro Leu Phe Met Leu Met
665 670 675

```
Ser Phe Met Pro Ala Ser Ala Val Pro Pro Phe Ser Cys Gly Val
                680                 685                 690

Ile Ser Thr Leu Arg Ser Arg Glu Glu Gly Ala Val Asp Lys Ser
                695                 700                 705

Tyr Cys Thr Leu Leu Asp Cys Leu Cys Ser Trp Gly Gln Val Gly
                710                 715                 720

His Ile Leu Glu Leu Val Asp Asn Trp Leu Pro Thr Glu His Ala
                725                 730                 735

Gln Ala Lys Ser Asn Thr Ala Ser Lys Gly Arg Val Gln Ile His
                740                 745                 750

Asp Thr Arg Pro Val Lys Pro Glu Leu Ala Leu Val Tyr Ile Glu
                755                 760                 765

Tyr Leu Leu Thr His Pro Lys Asn Arg Glu Cys Leu Leu Ser Ala
                770                 775                 780

Pro Arg Lys Lys Leu Asn His Leu Leu Lys Ala Leu Glu Thr Ser
                785                 790                 795

Lys Ala Asp Leu Glu Ser Leu Leu Gln Thr Pro Gly Gly Lys Pro
                800                 805                 810

Arg Gly Phe Ser Glu Ala Ala Ala Pro Arg Ala Phe Gly Leu His
                815                 820                 825

Cys Arg Leu Ser Ile His Leu Gln His Lys Phe Cys Ser Glu Gly
                830                 835                 840

Lys Val Tyr Leu Ser Met Leu Glu Asp Thr Gly Phe Trp Leu Glu
                845                 850                 855

Ser Lys Ile Leu Ser Phe Ile Gln Asp Gln Glu Glu Asp Tyr Leu
                860                 865                 870

Lys Leu His Arg Val Ile Tyr Gln Gln Ile Ile Gln Thr Tyr Leu
                875                 880                 885

Thr Val Cys Lys Asp Val Val Met Val Gly Leu Gly Asp His Gln
                890                 895                 900

Phe Gln Met Gln Leu Leu Gln Arg Ser Leu Gly Ile Met Gln Thr
                905                 910                 915

Val Lys Gly Phe Phe Tyr Val Ser Leu Leu Leu Asp Ile Leu Lys
                920                 925                 930

Glu Ile Thr Gly Ser Ser Leu Ile Gln Lys Thr Asp Ser Asp Glu
                935                 940                 945

Glu Val Ala Met Leu Leu Asp Thr Val Gln Lys Val Phe Gln Lys
                950                 955                 960

Met Leu Glu Cys Ile Ala Arg Ser Phe Arg Lys Gln Pro Glu Glu
                965                 970                 975

Gly Leu Arg Leu Leu Tyr Ser Val Gln Arg Pro Leu His Glu Phe
                980                 985                 990

Ile Thr Ala Val Gln Ser Arg His Thr Asp Thr Pro Val His Arg
                995                1000                1005

Gly Val Leu Ser Thr Leu Ile Ala Gly Pro Val Val Glu Ile Ser
               1010                1015                1020

His Gln Leu Arg Lys Val Ser Asp Val Glu Glu Leu Thr Pro Pro
               1025                1030                1035

Glu His Leu Ser Asp Leu Pro Pro Phe Ser Arg Cys Leu Ile Gly
               1040                1045                1050
```

Ile Ile Ile Lys Ser Ser Asn Val Val Arg Ser Phe Leu Asp Glu
1055 1060 1065

Leu Lys Ala Cys Val Ala Ser Asn Asp Ile Glu Gly Ile Val Cys
1070 1075 1080

Leu Thr Ala Ala Val His Ile Ile Leu Val Ile Asn Ala Gly Lys
1085 1090 1095

His Lys Ser Ser Lys Val Arg Glu Val Ala Ala Thr Val His Arg
1100 1105 1110

Lys Leu Lys Thr Phe Met Glu Ile Thr Leu Glu Glu Asp Ser Ile
1115 1120 1125

Glu Arg Phe Leu Tyr Glu Ser Ser Ser Arg Thr Leu Gly Glu Leu
1130 1135 1140

Leu Asn Ser

<210> 149
<211> 384
<212> PRT
<213> Homo sapiens

<400> 149
Met Asn Ser Thr Leu Phe Ser Gln Val Glu Asn His Ser Val His
1 5 10 15

Ser Asn Phe Ser Glu Lys Asn Ala Gln Leu Leu Ala Phe Glu Asn
20 25 30

Asp Asp Cys His Leu Pro Leu Ala Met Ile Phe Thr Leu Ala Leu
35 40 45

Ala Tyr Gly Ala Val Ile Ile Leu Gly Val Ser Gly Asn Leu Ala
50 55 60

Leu Ile Ile Ile Ile Leu Lys Gln Lys Glu Met Arg Asn Val Thr
65 70 75

Asn Ile Leu Ile Val Asn Leu Ser Phe Ser Asp Leu Leu Val Ala
80 85 90

Ile Met Cys Leu Pro Phe Thr Phe Val Tyr Thr Leu Met Asp His
95 100 105

Trp Val Phe Gly Glu Ala Met Cys Lys Leu Asn Pro Phe Val Gln
110 115 120

Cys Val Ser Ile Thr Val Ser Ile Phe Ser Leu Val Leu Ile Ala
125 130 135

Val Glu Arg His Gln Leu Ile Ile Asn Pro Arg Gly Trp Arg Pro
140 145 150

Asn Asn Arg His Ala Tyr Val Gly Ile Ala Val Ile Trp Val Leu
155 160 165

Ala Val Ala Ser Ser Leu Pro Phe Leu Ile Tyr Gln Val Met Thr
170 175 180

Asp Glu Pro Phe Gln Asn Val Thr Leu Asp Ala Tyr Lys Asp Lys
185 190 195

Tyr Val Cys Phe Asp Gln Phe Pro Ser Asp Ser His Arg Leu Ser
200 205 210

Tyr Thr Thr Leu Leu Leu Val Leu Gln Tyr Phe Gly Pro Leu Cys
215 220 225

Phe Ile Phe Ile Cys Tyr Phe Lys Ile Tyr Ile Arg Leu Lys Arg
230 235 240

Arg Asn Asn Met Met Asp Lys Met Arg Asp Asn Lys Tyr Arg Ser

245 250 255

Ser Glu Thr Lys Arg Ile Asn Ile Met Leu Leu Ser Ile Val Val
260 265 270

Ala Phe Ala Val Cys Trp Leu Pro Leu Thr Ile Phe Asn Thr Val
275 280 285

Phe Asp Trp Asn His Gln Ile Ile Ala Thr Cys Asn His Asn Leu
290 295 300

Leu Phe Leu Leu Cys His Leu Thr Ala Met Ile Ser Thr Cys Val
305 310 315

Asn Pro Ile Phe Tyr Gly Phe Leu Asn Lys Asn Phe Gln Arg Asp
320 325 330

Leu Gln Phe Phe Phe Asn Phe Cys Asp Phe Arg Ser Arg Asp Asp
335 340 345

Asp Tyr Glu Thr Ile Ala Met Ser Thr Met His Thr Asp Val Ser
350 355 360

Lys Thr Ser Leu Lys Gln Ala Ser Pro Val Ala Phe Lys Lys Ile
365 370 375

Asn Asn Asn Asp Asp Asn Glu Lys Ile
380

<210> 150
<211> 327
<212> PRT
<213> Homo sapiens

<400> 150
Met Val Arg Cys Asp Arg Gly Leu Gln Met Leu Leu Thr Thr Ala
1 5 10 15

Gly Ala Phe Ala Ala Phe Ser Leu Met Ala Ile Ala Ile Gly Thr
20 25 30

Asp Tyr Trp Leu Tyr Ser Ser Ala His Ile Cys Asn Gly Thr Asn
35 40 45

Leu Thr Met Asp Asp Gly Pro Pro Pro Arg Arg Ala Arg Gly Asp
50 55 60

Leu Thr His Ser Gly Leu Trp Arg Val Cys Cys Ile Glu Gly Ile
65 70 75

Tyr Lys Gly His Cys Phe Arg Ile Asn His Phe Pro Glu Asp Asn
80 85 90

Asp Tyr Asp His Asp Ser Ser Glu Tyr Leu Leu Arg Ile Val Arg
95 100 105

Ala Ser Ser Val Phe Pro Ile Leu Ser Thr Ile Leu Leu Leu Leu
110 115 120

Gly Gly Leu Cys Ile Gly Ala Gly Arg Ile Tyr Ser Arg Lys Asn
125 130 135

Asn Ile Val Leu Ser Ala Gly Ile Leu Phe Val Ala Ala Gly Leu
140 145 150

Ser Asn Ile Ile Gly Ile Ile Val Tyr Ile Ser Ser Asn Thr Gly
155 160 165

Asp Pro Ser Asp Lys Arg Asp Glu Asp Lys Lys Asn His Tyr Asn
170 175 180

Tyr Gly Trp Ser Phe Tyr Phe Gly Ala Leu Ser Phe Ile Val Ala
185 190 195

Glu Thr Val Gly Val Leu Ala Val Asn Ile Tyr Ile Glu Lys Asn
200 205 210

Lys Glu Leu Arg Phe Lys Thr Lys Arg Glu Phe Leu Lys Ala Ser
215 220 225

Ser Ser Ser Pro Tyr Ala Arg Met Pro Ser Tyr Arg Tyr Arg Arg
230 235 240

Arg Arg Ser Arg Ser Ser Ser Arg Ser Thr Glu Ala Ser Pro Ser
245 250 255

Arg Asp Val Ser Pro Met Gly Leu Lys Ile Thr Gly Ala Ile Pro
260 265 270

Met Gly Glu Leu Ser Met Tyr Thr Leu Ser Arg Glu Pro Leu Lys
275 280 285

Val Thr Thr Ala Ala Ser Tyr Ser Pro Asp Gln Glu Ala Ser Phe
290 295 300

Leu Gln Val His Asp Phe Phe Gln Gln Asp Leu Lys Glu Gly Phe
305 310 315

His Val Ser Met Leu Asn Arg Arg Thr Thr Pro Val
320 325

<210> 151
<211> 323
<212> PRT
<213> Homo sapiens

<400> 151
Met Ala Arg Gly Pro Gly Leu Ala Pro Pro Pro Leu Arg Leu Pro
1 5 10 15

Leu Leu Leu Leu Val Leu Ala Ala Val Thr Gly His Thr Ala Ala
20 25 30

Gln Asp Asn Cys Thr Cys Pro Thr Asn Lys Met Thr Val Cys Ser
35 40 45

Pro Asp Gly Pro Gly Gly Arg Cys Gln Cys Arg Ala Leu Gly Ser
50 55 60

Gly Met Ala Val Asp Cys Ser Thr Leu Thr Ser Lys Cys Leu Leu
65 70 75

Leu Lys Ala Arg Met Ser Ala Pro Lys Asn Ala Arg Thr Leu Val
80 85 90

Arg Pro Ser Glu His Ala Leu Val Asp Asn Asp Gly Leu Tyr Asp
95 100 105

Pro Asp Cys Asp Pro Glu Gly Arg Phe Lys Ala Arg Gln Cys Asn
110 115 120

Gln Thr Ser Val Cys Trp Cys Val Asn Ser Val Gly Val Arg Arg
125 130 135

Thr Asp Lys Gly Asp Leu Ser Leu Arg Cys Asp Glu Leu Val Arg
140 145 150

Thr His His Ile Leu Ile Asp Leu Arg His Arg Pro Thr Ala Gly
155 160 165

Ala Phe Asn His Ser Asp Leu Asp Ala Glu Leu Arg Arg Leu Phe
170 175 180

Arg Glu Arg Tyr Arg Leu His Pro Lys Phe Val Ala Ala Val His
185 190 195

Tyr Glu Gln Pro Thr Ile Gln Ile Glu Leu Arg Gln Asn Thr Ser
200 205 210

Gln Lys Ala Ala Gly Glu Val Asp Ile Gly Asp Ala Ala Tyr Tyr
215 220 225

1329
1322

Phe Glu Arg Asp Ile Lys Gly Glu Ser Leu Phe Gln Gly Arg Gly
230 235 240

Gly Leu Asp Leu Arg Val Arg Gly Glu Pro Leu Gln Val Glu Arg
245 250 255

Thr Leu Ile Tyr Tyr Leu Asp Glu Ile Pro Pro Lys Phe Ser Met
260 265 270

Lys Arg Leu Thr Ala Gly Leu Ile Ala Val Ile Val Val Val Val
275 280 285

Val Ala Leu Val Ala Gly Met Ala Val Leu Val Ile Thr Asn Arg
290 295 300

Arg Lys Ser Gly Lys Tyr Lys Lys Val Glu Ile Lys Glu Leu Gly
305 310 315

Glu Leu Arg Lys Glu Pro Ser Leu
320

<210> 152
<211> 207
<212> PRT
<213> Homo sapiens

<400> 152
Met Gly Lys Cys Ser Gly Arg Cys Thr Leu Val Ala Phe Cys Cys
1 5 10 15

Leu Gln Leu Val Ala Ala Leu Glu Arg Gln Ile Phe Asp Phe Leu
20 25 30

Gly Tyr Gln Trp Ala Pro Ile Leu Ala Asn Phe Leu His Ile Met
35 40 45

Ala Val Ile Leu Gly Ile Phe Gly Thr Val Gln Tyr Arg Ser Arg
50 55 60

Tyr Leu Ile Leu Tyr Ala Ala Trp Leu Val Leu Trp Val Gly Trp
65 70 75

Asn Ala Phe Ile Ile Cys Phe Tyr Leu Glu Val Gly Gln Leu Ser
80 85 90

Gln Asp Arg Asp Phe Ile Met Thr Phe Asn Thr Ser Leu His Arg
95 100 105

Ser Trp Trp Met Glu Asn Gly Pro Gly Cys Leu Val Thr Pro Val
110 115 120

Leu Asn Ser Arg Leu Ala Leu Glu Asp His His Val Ile Ser Val
125 130 135

Thr Gly Cys Leu Leu Asp Tyr Pro Tyr Ile Glu Ala Leu Ser Ser
140 145 150

Ala Leu Gln Ile Phe Leu Ala Leu Phe Gly Phe Val Phe Ala Cys
155 160 165

Tyr Val Ser Lys Val Phe Leu Glu Glu Glu Asp Ser Phe Asp Phe
170 175 180

Ile Gly Gly Phe Asp Ser Tyr Gly Tyr Gln Ala Pro Gln Lys Thr
185 190 195

Ser His Leu Gln Leu Gln Pro Leu Tyr Thr Ser Gly
200 205

<210> 153
<211> 581
<212> PRT
<213> Homo sapiens

<400> 153
Met Pro Ala Pro Arg Ala Arg Glu Gln Pro Arg Val Pro Gly Glu

~~1325~~
1323

```
  1               5                  10                  15

Arg Gln Pro Leu Leu Pro Arg Gly Ala Arg Gly Pro Arg Arg Trp
                20                  25                  30

Arg Arg Ala Ala Gly Ala Ala Val Leu Leu Val Glu Met Leu Glu
                35                  40                  45

Arg Ala Ala Phe Phe Gly Val Thr Ala Asn Leu Val Leu Tyr Leu
                50                  55                  60

Asn Ser Thr Asn Phe Asn Trp Thr Gly Glu Gln Ala Thr Arg Ala
                65                  70                  75

Ala Leu Val Phe Leu Gly Ala Ser Tyr Leu Leu Ala Pro Val Gly
                80                  85                  90

Gly Trp Leu Ala Asp Val Tyr Leu Gly Arg Tyr Arg Ala Val Ala
                95                 100                 105

Leu Ser Leu Leu Leu Tyr Leu Ala Ala Ser Gly Leu Leu Pro Ala
               110                 115                 120

Thr Ala Phe Pro Asp Gly Arg Ser Ser Phe Cys Gly Glu Met Pro
               125                 130                 135

Ala Ser Pro Leu Gly Pro Ala Cys Pro Ser Ala Gly Cys Pro Arg
               140                 145                 150

Ser Ser Pro Ser Pro Tyr Cys Ala Pro Val Leu Tyr Ala Gly Leu
               155                 160                 165

Leu Leu Leu Gly Leu Ala Ala Ser Ser Val Arg Ser Asn Leu Thr
               170                 175                 180

Ser Phe Gly Ala Asp Gln Val Met Asp Leu Gly Arg Asp Ala Thr
               185                 190                 195

Arg Arg Phe Phe Asn Trp Phe Tyr Trp Ser Ile Asn Leu Gly Ala
               200                 205                 210

Val Leu Ser Leu Leu Val Val Ala Phe Ile Gln Gln Asn Ile Ser
               215                 220                 225

Phe Leu Leu Gly Tyr Ser Ile Pro Val Gly Cys Val Gly Leu Ala
               230                 235                 240

Phe Phe Ile Phe Leu Phe Ala Thr Pro Val Phe Ile Thr Lys Pro
               245                 250                 255

Pro Met Gly Ser Gln Val Ser Ser Met Leu Lys Leu Ala Leu Gln
               260                 265                 270

Asn Cys Cys Pro Gln Leu Trp Gln Arg His Ser Ala Arg Asp Arg
               275                 280                 285

Gln Cys Ala Arg Val Leu Ala Asp Glu Arg Ser Pro Gln Pro Gly
               290                 295                 300

Ala Ser Pro Gln Glu Asp Ile Ala Asn Phe Gln Val Leu Val Lys
               305                 310                 315

Ile Leu Pro Val Met Val Thr Leu Val Pro Tyr Trp Met Val Tyr
               320                 325                 330

Phe Gln Met Gln Ser Thr Tyr Val Leu Gln Gly Leu His Leu His
               335                 340                 345

Ile Pro Asn Ile Phe Pro Ala Asn Pro Ala Asn Ile Ser Val Ala
               350                 355                 360

Leu Arg Ala Gln Gly Ser Ser Tyr Thr Ile Pro Glu Ala Trp Leu
               365                 370                 375

Leu Leu Ala Asn Val Val Val Val Leu Ile Leu Val Pro Leu Lys
               380                 385                 390
```

Asp Arg Leu Ile Asp Pro Leu Leu Leu Arg Cys Lys Leu Leu Pro
395 400 405

Ser Ala Leu Gln Lys Met Ala Leu Gly Met Phe Phe Gly Phe Thr
410 415 420

Ser Val Ile Val Ala Gly Val Leu Glu Met Glu Arg Leu His Tyr
425 430 435

Ile His His Asn Glu Thr Val Ser Gln Gln Ile Gly Glu Val Leu
440 445 450

Tyr Asn Ala Ala Pro Leu Ser Ile Trp Trp Gln Ile Pro Gln Tyr
455 460 465

Leu Leu Ile Gly Ile Ser Glu Ile Phe Ala Ser Ile Pro Gly Leu
470 475 480

Glu Phe Ala Tyr Ser Glu Ala Pro Arg Ser Met Gln Gly Ala Ile
485 490 495

Met Gly Ile Phe Phe Cys Leu Ser Gly Val Gly Ser Leu Leu Gly
500 505 510

Ser Ser Leu Val Ala Leu Leu Ser Leu Pro Gly Gly Trp Leu His
515 520 525

Cys Pro Lys Asp Phe Gly Asn Ile Asn Asn Cys Arg Met Asp Leu
530 535 540

Tyr Phe Phe Leu Leu Ala Gly Ile Gln Ala Val Thr Ala Leu Leu
545 550 555

Phe Val Trp Ile Ala Gly Arg Tyr Glu Arg Ala Ser Gln Gly Pro
560 565 570

Ala Ser His Ser Arg Phe Ser Arg Asp Arg Gly
575 580

<210> 154
<211> 1266
<212> PRT
<213> Homo sapiens

<400> 154
Met Trp Val Asn Pro Glu Glu Val Leu Leu Ala Asn Ala Leu Trp
1 5 10 15

Ile Thr Glu Arg Ala Asn Pro Tyr Phe Ile Leu Gln Arg Arg Lys
20 25 30

Gly His Ala Gly Asp Gly Gly Gly Gly Gly Gly Leu Ala Gly Leu
35 40 45

Leu Val Gly Thr Leu Asp Val Val Leu Asp Ser Ser Ala Arg Val
50 55 60

Ala Pro Tyr Arg Ile Leu Tyr Gln Thr Pro Asp Ser Leu Val Tyr
65 70 75

Trp Thr Ile Ala Cys Gly Gly Ser Arg Lys Glu Ile Thr Glu His
80 85 90

Trp Glu Trp Leu Glu Gln Asn Leu Leu Gln Thr Leu Ser Ile Phe
95 100 105

Glu Asn Glu Asn Asp Ile Thr Thr Phe Val Arg Gly Lys Ile Gln
110 115 120

Gly Ile Ile Ala Glu Tyr Asn Lys Ile Asn Asp Val Lys Glu Asp
125 130 135

Asp Asp Thr Glu Lys Phe Lys Glu Ala Ile Val Lys Phe His Arg
140 145 150

1325

Leu Phe Gly Met Pro Glu Glu Glu Lys Leu Val Asn Tyr Tyr Ser
155 160 165

Cys Ser Tyr Trp Lys Gly Lys Val Pro Arg Gln Gly Trp Met Tyr
170 175 180

Leu Ser Ile Asn His Leu Cys Phe Tyr Ser Phe Leu Met Gly Arg
185 190 195

Glu Ala Lys Leu Val Ile Arg Trp Val Asp Ile Thr Gln Leu Glu
200 205 210

Lys Asn Ala Thr Leu Leu Leu Pro Asp Val Ile Lys Val Ser Thr
215 220 225

Arg Ser Ser Glu His Phe Phe Ser Val Phe Leu Asn Ile Asn Glu
230 235 240

Thr Phe Lys Leu Met Glu Gln Leu Ala Asn Ile Ala Met Arg Gln
245 250 255

Leu Leu Asp Asn Glu Gly Phe Glu Gln Asp Arg Ser Leu Pro Lys
260 265 270

Leu Lys Arg Lys Ser Pro Lys Lys Val Ser Ala Leu Lys Arg Asp
275 280 285

Leu Asp Ala Arg Ala Lys Ser Glu Arg Tyr Arg Ala Leu Phe Arg
290 295 300

Leu Pro Lys Asp Glu Lys Leu Asp Gly His Thr Asp Cys Thr Leu
305 310 315

Trp Thr Pro Phe Asn Lys Met His Ile Leu Gly Gln Met Phe Val
320 325 330

Ser Thr Asn Tyr Ile Cys Phe Thr Ser Lys Glu Glu Asn Leu Cys
335 340 345

Ser Leu Ile Ile Pro Leu Arg Glu Val Thr Ile Val Glu Lys Ala
350 355 360

Asp Ser Ser Ser Val Leu Pro Ser Pro Leu Ser Ile Ser Thr Arg
365 370 375

Asn Arg Met Thr Phe Leu Phe Ala Asn Leu Lys Asp Arg Asp Phe
380 385 390

Leu Val Gln Arg Ile Ser Asp Phe Leu Gln Gln Thr Thr Ser Lys
395 400 405

Ile Tyr Ser Asp Lys Glu Phe Ala Gly Ser Tyr Asn Ser Ser Asp
410 415 420

Asp Glu Val Tyr Ser Arg Pro Ser Ser Leu Val Ser Ser Ser Pro
425 430 435

Gln Arg Ser Thr Ser Ser Asp Ala Asp Gly Glu Arg Gln Phe Asn
440 445 450

Leu Asn Gly Asn Ser Val Pro Thr Ala Thr Gln Thr Leu Met Thr
455 460 465

Met Tyr Arg Arg Arg Ser Pro Glu Glu Phe Asn Pro Lys Leu Ala
470 475 480

Lys Glu Phe Leu Lys Glu Gln Ala Trp Lys Ile His Phe Ala Glu
485 490 495

Tyr Gly Gln Gly Ile Cys Met Tyr Arg Thr Glu Lys Thr Arg Glu
500 505 510

Leu Val Leu Lys Gly Ile Pro Glu Ser Met Arg Gly Glu Leu Trp
515 520 525

Leu Leu Leu Ser Gly Ala Ile Asn Glu Lys Ala Thr His Pro Gly

~~1320~~
1326

530 535 540

Tyr Tyr Glu Asp Leu Val Glu Lys Ser Met Gly Lys Tyr Asn Leu
545 550 555

Ala Thr Glu Glu Ile Glu Arg Asp Leu His Arg Ser Leu Pro Glu
560 565 570

His Pro Ala Phe Gln Asn Glu Met Gly Ile Ala Ala Leu Arg Arg
575 580 585

Val Leu Thr Ala Tyr Ala Phe Arg Asn Pro Asn Ile Gly Tyr Cys
590 595 600

Gln Ala Met Asn Ile Val Thr Ser Val Leu Leu Leu Tyr Ala Lys
605 610 615

Glu Glu Glu Ala Phe Trp Leu Leu Val Ala Leu Cys Glu Arg Met
620 625 630

Leu Pro Asp Tyr Tyr Asn Thr Arg Val Val Gly Ala Leu Val Asp
635 640 645

Gln Gly Val Phe Glu Glu Leu Ala Arg Asp Tyr Val Pro Gln Leu
650 655 660

Tyr Asp Cys Met Gln Asp Leu Gly Val Ile Ser Thr Ile Ser Leu
665 670 675

Ser Trp Phe Leu Thr Leu Phe Leu Ser Val Met Pro Phe Glu Ser
680 685 690

Ala Val Val Val Val Asp Cys Phe Phe Tyr Glu Gly Ile Lys Val
695 700 705

Ile Phe Gln Leu Ala Leu Ala Val Leu Asp Ala Asn Val Asp Lys
710 715 720

Leu Leu Asn Cys Lys Asp Asp Gly Glu Ala Met Thr Val Leu Gly
725 730 735

Arg Tyr Leu Asp Ser Val Thr Asn Lys Asp Ser Thr Leu Pro Pro
740 745 750

Ile Pro His Leu His Ser Leu Leu Ser Asp Asp Val Glu Pro Tyr
755 760 765

Pro Glu Val Asp Ile Phe Arg Leu Ile Arg Thr Ser Tyr Glu Lys
770 775 780

Phe Gly Thr Ile Arg Ala Asp Leu Ile Glu Gln Met Arg Phe Lys
785 790 795

Gln Arg Leu Lys Val Ile Gln Thr Leu Glu Asp Thr Thr Lys Arg
800 805 810

Asn Val Val Arg Thr Ile Val Thr Glu Thr Ser Phe Thr Ile Asp
815 820 825

Glu Leu Glu Glu Leu Tyr Ala Leu Phe Lys Ala Glu His Leu Thr
830 835 840

Ser Cys Tyr Trp Gly Gly Ser Ser Asn Ala Leu Asp Arg His Asp
845 850 855

Pro Ser Leu Pro Tyr Leu Glu Gln Tyr Arg Ile Asp Phe Glu Gln
860 865 870

Phe Lys Gly Met Phe Ala Leu Leu Phe Pro Trp Ala Cys Gly Thr
875 880 885

His Ser Asp Val Leu Ala Ser Arg Leu Phe Gln Leu Leu Asp Glu
890 895 900

Asn Gly Asp Ser Leu Ile Asn Phe Arg Glu Phe Val Ser Gly Leu
905 910 915

Ser Ala Ala Cys His Gly Asp Leu Thr Glu Lys Leu Lys Leu Leu
920 925 930

Tyr Lys Met His Val Leu Pro Glu Pro Ser Ser Asp Gln Asp Glu
935 940 945

Pro Asp Ser Ala Phe Glu Ala Thr Gln Tyr Phe Phe Glu Asp Ile
950 955 960

Thr Pro Glu Cys Thr His Val Val Gly Leu Asp Ser Arg Ser Lys
965 970 975

Gln Gly Ala Asp Asp Gly Phe Val Thr Val Ser Leu Lys Pro Asp
980 985 990

Lys Gly Lys Arg Ala Asn Ser Gln Glu Asn Arg Asn Tyr Leu Arg
995 1000 1005

Leu Trp Thr Pro Glu Asn Lys Ser Lys Ser Lys Asn Ala Lys Asp
1010 1015 1020

Leu Pro Lys Leu Asn Gln Gly Gln Phe Ile Glu Leu Cys Lys Thr
1025 1030 1035

Met Tyr Asn Met Phe Ser Glu Asp Pro Asn Glu Gln Glu Leu Tyr
1040 1045 1050

His Ala Thr Ala Ala Val Thr Ser Leu Leu Leu Glu Ile Gly Glu
1055 1060 1065

Val Gly Lys Leu Phe Val Ala Gln Pro Ala Lys Glu Gly Gly Ser
1070 1075 1080

Gly Gly Ser Gly Pro Ser Cys His Gln Gly Ile Pro Gly Val Leu
1085 1090 1095

Phe Pro Lys Lys Gly Pro Gly Gln Pro Tyr Val Val Glu Ser Val
1100 1105 1110

Glu Pro Leu Pro Ala Ser Leu Ala Pro Asp Ser Glu Glu His Ser
1115 1120 1125

Leu Gly Gly Gln Met Glu Asp Ile Lys Leu Glu Asp Ser Ser Pro
1130 1135 1140

Arg Asp Asn Gly Ala Cys Ser Ser Met Leu Ile Ser Asp Asp Asp
1145 1150 1155

Thr Lys Asp Asp Ser Ser Met Ser Ser Tyr Ser Val Leu Ser Ala
1160 1165 1170

Gly Ser His Glu Glu Asp Lys Leu His Cys Glu Asp Ile Gly Glu
1175 1180 1185

Asp Thr Val Leu Val Arg Ser Gly Gln Gly Thr Ala Ala Leu Pro
1190 1195 1200

Arg Ser Thr Ser Leu Asp Arg Asp Trp Ala Ile Thr Phe Glu Gln
1205 1210 1215

Phe Leu Ala Ser Leu Leu Thr Glu Pro Ala Leu Val Lys Tyr Phe
1220 1225 1230

Asp Lys Pro Val Cys Met Met Ala Arg Ile Thr Ser Ala Lys Asn
1235 1240 1245

Ile Arg Met Met Gly Lys Pro Leu Thr Ser Ala Ser Asp Tyr Glu
1250 1255 1260

Ile Ser Ala Met Ser Gly
1265

<210> 155
<211> 21
<212> DNA

<213> Homo sapiens

<400> 155
aaggaagagg cgcgggagaa g 21

<210> 156
<211> 21
<212> DNA
<213> Homo sapiens

<400> 156
ggaagaggcg cgggagaagt t 21

<210> 157
<211> 21
<212> DNA
<213> Homo sapiens

<400> 157
cuucucccgc gccucuucct t 21

<210> 158
<211> 21
<212> DNA
<213> Homo sapiens

<400> 158
cuggacuucc agaagaacat t 21

<210> 159
<211> 18
<212> DNA
<213> Homo sapiens

<400> 159
uguucuggaa guccagtt 18

Listado de Secuencias

<110> GENENTECH, INC.

<120> Composiciones y Métodos para el Diagnóstico y Tratamiento de Tumores

<130> P5001R1P1 PCT

<140> PCT/US2005/031798
<141> 2005-07-09

<150> US 10/936,626
<151> 2004-09-08

<150> US 10/177,488
<151> 2002-06-19

<150> US 60/299,500
<151> 2001-06-20

<150> US 60/301,880
<151> 2001-06-29

<150> US 60/323,268
<151> 2001-09-18

<160> 159

<210> 1
<211> 1039
<212> DNA
<213> Homo sapiens

<400> 1
ccaagaattc ggcacgaggg tgacccggac agctgtcctc tctgacacca 50

ccccggcctg cctctttgtt gccatgagag ctgcctacct cttcctgcta 100

ttcctgcctg caggcttgct ggctcagggc cagtatgacc tggacccgct 150

gccgccgttc cctgaccacg tccagtacac ccactatagc gaccagatcg 200

acaacccaga ctactatgat tatcaagagg tgactcctcg gccctccgag 250

gaacagttcc agttccagtc ccagcagcaa gtccaacagg aagtcatccc 300

agccccaacc ccagaaccag gaaatgcaga gctggagccc acagagcctg 350

ggcctcttga ctgccgtgag gaacagtacc cgtgcacccg cctctactcc 400

atacacaggc cttgcaaaca gtgtctcaac gaggtctgct tctacagcct 450

ccgccgtgtg tacgtcatta acaaggagat ctgtgttcgt acagtgtgtg 500

cccacgagga gctcctccga gctgacctct gtcgggacaa gttctccaaa 550

tgtggcgtga tggccagcag cggcctgtgc caatccgtgg cggcctcctg 600

tgccaggagc tgtgggagct gctagggtgg tgctggcatc ctgagtcctg 650

gccctcctgg gatctggggc cctcgggccc tgcctgacct ggtgcttttt 700

tccccatccc catgttcctt ttattctgta aaaagttagt ggactgcagc 750

cctgggggtt gcaggctgcg gtgcctcagg cccctccttc agcctgtggc 800

cacctctggg gcacgatggg ggctccccac tgcccagtct gcccctcggg 850

ttgggggagt atcccaggcc tctctgtggg acctgggccc ctgacgggcc 900

ttctcagccc gttttgagga cagacagtcc cccgaggtag gctacatccc 950

cccaccccag ctggtctgct tggatttcct acagcccccg tgggcatgga 1000

ccacctttat tttatacaaa attaaaaaca agtttttac 1039

1372
1330

Listado de Secuencias

<110> GENENTECH, INC.

<120> Composiciones y Métodos para el Diagnóstico y Tratamiento de Tumores

<130> P5001R1P1 PCT

<140> PCT/US2005/031798
<141> 2005-07-09

<150> US 10/936,626
<151> 2004-09-08

<150> US 10/177,488
<151> 2002-06-19

<150> US 60/299,500
<151> 2001-06-20

<150> US 60/301,880
<151> 2001-06-29

<150> US 60/323,268
<151> 2001-09-18

<160> 159

<210> 1
<211> 1039
<212> DNA
<213> Homo sapiens

<400> 1
ccaagaattc ggcacgaggg tgacccggac agctgtcctc tctgacacca 50

ccccggcctg cctctttgtt gccatgagag ctgcctacct cttcctgcta 100

ttcctgcctg caggcttgct ggctcagggc cagtatgacc tggacccgct 150

gccgccgttc cctgaccacg tccagtacac ccactatagc gaccagatcg 200

acaacccaga ctactatgat tatcaagagg tgactcctcg gccctccgag 250

gaacagttcc agttccagtc ccagcagcaa gtccaacagg aagtcatccc 300

agccccaacc ccagaaccag gaaatgcaga gctggagccc acagagcctg 350

ggcctcttga ctgccgtgag gaacagtacc cgtgcacccg cctctactcc 400

atacacaggc cttgcaaaca gtgtctcaac gaggtctgct tctacagcct 450

ccgccgtgtg tacgtcatta acaaggagat ctgtgttcgt acagtgtgtg 500

cccacgagga gctcctccga gctgacctct gtcgggacaa gttctccaaa 550

tgtggcgtga tggccagcag cggcctgtgc caatccgtgg cggcctcctg 600

tgccaggagc tgtgggagct gctagggtgg tgctggcatc ctgagtcctg 650

gccctcctgg gatctggggc cctcgggccc tgcctgacct ggtgcttttt 700

tccccatccc catgttcctt ttattctgta aaaagttagt ggactgcagc 750

cctgggggtt gcaggctgcg gtgcctcagg cccctccttc agcctgtggc 800

cacctctggg gcacgatggg ggctccccac tgcccagtct gcccctcggg 850

ttgggggagt atcccaggcc tctctgtggg acctgggccc ctgacgggcc 900

ttctcagccc gttttgagga cagacagtcc cccgaggtag gctacatccc 950

cccaccccag ctggtctgct tggatttcct acagcccccg tgggcatgga 1000

ccacctttat tttatacaaa attaaaaaca agtttttac 1039

Sequence Listing

<110> GENENTECH, INC.

<120> Compositions and Methods for the Diagnosis and Treatment of Tumor

<130> P5001R1P1 PCT

<140> PCT/US2005/031798
<141> 2005-07-09

<150> US 10/936,626
<151> 2004-09-08

<150> US 10/177,488
<151> 2002-06-19

<150> US 60/299,500
<151> 2001-06-20

<150> US 60/301,880
<151> 2001-06-29

<150> US 60/323,268
<151> 2001-09-18

<160> 159

<210> 1
<211> 1039
<212> DNA
<213> Homo sapiens

<400> 1
ccaagaattc ggcacgaggg tgacccggac agctgtcctc tctgacacca 50

ccccggcctg cctctttgtt gccatgagag ctgcctacct cttcctgcta 100

ttcctgcctg caggcttgct ggctcagggc cagtatgacc tggacccgct 150

gccgccgttc cctgaccacg tccagtacac ccactatagc gaccagatcg 200

acaacccaga ctactatgat tatcaagagg tgactcctcg gccctccgag 250

gaacagttcc agttccagtc ccagcagcaa gtccaacagg aagtcatccc 300

agccccaacc ccagaaccag gaaatgcaga gctggagccc acagagcctg 350

ggcctcttga ctgccgtgag gaacagtacc cgtgcacccg cctctactcc 400

atacacaggc cttgcaaaca gtgtctcaac gaggtctgct tctacagcct 450

ccgccgtgtg tacgtcatta acaaggagat ctgtgttcgt acagtgtgtg 500

cccacgagga gctcctccga gctgacctct gtcgggacaa gttctccaaa 550

tgtggcgtga tggccagcag cggcctgtgc caatccgtgg cggcctcctg 600

tgccaggagc tgtgggagct gctagggtgg tgctggcatc ctgagtcctg 650

gccctcctgg gatctggggc cctcgggccc tgcctgacct ggtgcttttt 700

tccccatccc catgttcctt ttattctgta aaaagttagt ggactgcagc 750

cctgggggtt gcaggctgcg gtgcctcagg cccctccttc agcctgtggc 800

cacctctggg gcacgatggg ggctccccac tgcccagtct gcccctcggg 850

ttgggggagt atcccaggcc tctctgtggg acctgggccc ctgacgggcc 900

ttctcagccc gttttgagga cagacagtcc cccgaggtag gctacatccc 950

cccaccccag ctggtctgct tggatttcct acagcccccg tgggcatgga 1000

ccacctttat tttatacaaa attaaaaaca agtttttac 1039

<210> 2
<211> 5062
<212> DNA
<213> Homo sapiens

<400> 2
cactaacgct cttcctagtc cccgggccaa ctcggacagt ttgctcattt 50

attgcaacgg tcaaggctgg cttgtgccag aacggcgcgc gcgcgacgca 100

cgcacacaca cgggggggaaa ctttttttaaa aatgaaaggc tagaagagct 150

cagcggcggc gcgggccgtg cgcgagggct ccggagctga ctcgccgagg 200

caggaaatcc ctccggtcgc gacgcccggc cccgctcggc gcccgcgtgg 250

gatggtgcag cgctcgccgc cgggcccgag agctgctgca ctgaaggccg 300

gcgacgatgg cagcgcgccc gctgcccgtg tcccccgccc gcgccctcct 350

gctcgccctg gccggtgctc tgctcgcgcc ctgcgaggcc cgaggggtga 400

gcttatggaa cgaaggaaga gctgatgaag ttgtcagtgc ctctgttcgg 450

agtggggacc tctggatccc agtgaagagc ttcgactcca agaatcatcc 500

agaagtgctg aatattcgac tacaacggga aagcaaagaa ctgatcataa 550

atctggaaag aaatgaaggt ctcattgcca gcagtttcac ggaaacccac 600

tatctgcaag acggtactga tgtctccctc gctcgaaatt acacggtaat 650

tctgggtcac tgttactacc atggacatgt acggggatat tctgattcag 700

cagtcagtct cagcacgtgt tctggtctca ggggacttat tgtgtttgaa 750

aatgaaagct atgtcttaga accaatgaaa agtgcaacca acagatacaa 800

actcttccca gcgaagaagc tgaaaagcgt ccggggatca tgtggatcac 850

atcacaacac accaaacctc gctgcaaaga atgtgtttcc accaccctct 900

cagacatggg caagaaggca taaaagagag accctcaagg caactaagta 950

tgtggagctg gtgatcgtgg cagacaaccg agagtttcag aggcaaggaa 1000

aagatctgga aaaagttaag cagcgattaa tagagattgc taatcacgtt 1050

gacaagtttt acagaccact gaacattcgg atcgtgttgg taggcgtgga 1100

agtgtggaat gacatggaca aatgctctgt aagtcaggac ccattcacca 1150

gcctccatga atttctggac tggaggaaga tgaagcttct acctcgcaaa 1200

tcccatgaca atgcgcagct tgtcagtggg gtttatttcc aagggaccac 1250

catcggcatg gccccaatca tgagcatgtg cacggcagac cagtctgggg 1300

gaattgtcat ggaccattca gacaatcccc ttggtgcagc cgtgaccctg 1350

gcacatgagc tgggccacaa tttcgggatg aatcatgaca cactggacag 1400

gggctgtagc tgtcaaatgg cggttgagaa aggaggctgc atcatgaacg 1450

cttccaccgg gtacccattt cccatggtgt tcagcagttg cagcaggaag 1500

gacttggaga ccagcctgga gaaaggaatg ggggtgtgcc tgtttaacct 1550

gccggaagtc agggagtctt tcgggggcca gaagtgtggg aacagatttg 1600

tggaagaagg agaggagtgt gactgtgggg agccagagga atgtatgaat 1650

cgctgctgca atgccaccac ctgtaccctg aagccggacg ctgtgtgcgc 1700

acatgggctg tgctgtgaag actgccagct gaagcctgca ggaacagcgt 1750

gcagggactc cagcaactcc tgtgacctcc cagagttctg cacaggggcc 1800

agccctcact gcccagccaa cgtgtacctg cacgatgggc actcatgtca 1850

ggatgtggac ggctactgct acaatggcat ctgccagact cacgagcagc 1900

agtgtgtcac actctgggga ccaggtgcta aacctgcccc tgggatctgc 1950

tttgagagag tcaattctgc aggtgatcct tatggcaact gtggcaaagt 2000

ctcgaagagt tcctttgcca aatgcgagat gagagatgct aaatgtggaa 2050

aaatccagtg tcaaggaggt gccagccggc cagtcattgg taccaatgcc 2100

gtttccatag aaacaaacat ccccctgcag caaggaggcc ggattctgtg 2150

ccgggggacc cacgtgtact tgggcgatga catgccggac ccagggcttg 2200

tgcttgcagg cacaaagtgt gcagatggaa aaatctgcct gaatcgtcaa 2250

tgtcaaaata ttagtgtctt tggggttcac gagtgtgcaa tgcagtgcca 2300

cggcagaggg gtgtgcaaca acaggaagaa ctgccactgc gaggcccact 2350

gggcacctcc cttctgtgac aagtttggct ttggaggaag cacagacagc 2400

ggccccatcc ggcaagcaga taaccaaggt ttaaccatag gaattctggt 2450

gaccatcctg tgtcttcttg ctgccggatt tgtggtttat ctcaaaagga 2500

agaccttgat acgactgctg tttacaaata agaagaccac cattgaaaaa 2550

ctaaggtgtg tgcgcccttc ccggccaccc cgtggcttcc aaccctgtca 2600

ggctcacctc ggccaccttg gaaaaggcct gatgaggaag ccgccagatt 2650

cctacccacc gaaggacaat cccaggagat tgctgcagtg tcagaatgtt 2700

gacatcagca gacccctcaa cggcctgaat gtccctcagc cccagtcaac 2750

tcagcgagtg cttcctcccc tccaccgggc cccacgtgca cctagcgtcc 2800

ctgccagacc cctgccagcc aagcctgcac ttaggcaggc ccaggggacc 2850

tgtaagccaa acccccctca gaagcctctg cctgcagatc ctctggccag 2900

aacaactcgg ctcactcatg ccttggccag gaccccagga caatgggaga 2950

ctgggctccg cctggcaccc ctcagacctg ctccacaata tccacaccaa 3000

gtgcccagat ccacccacac cgcctatatt aagtgagaag ccgacacctt 3050

ttttcaacag tgaagacaga agtttgcact atctttcagc tccagttgga 3100

gtttttgta ccaactttta ggatttttt taatgtttaa aacatcatta 3150

ctataagaac tttgagctac tgccgtcagt gctgtgctgt gctatggtgc 3200

tctgtctact tgcacaggta cttgtaaatt attaatttat gcagaatgtt 3250

gattacagtg cagtgcgctg tagtaggcat ttttaccatc actgagtttt 3300

ccatggcagg aaggcttgtt gtgcttttag tattttagtg aacttgaaat 3350

atcctgcttg atgggattct ggacaggatg tgtttgcttt ctgatcaagg 3400

ccttattgga aagcagtccc ccaactaccc ccagctgtgc ttatggtacc 3450

agatgcagct caagagatcc caagtagaat ctcagttgat tttctggatt 3500

ccccatctca ggccagagcc aaggggcttc aggtccaggc tgtgtttggc 3550

tttcagggag gccctgtgcc ccttgacaac tggcaggcag gctcccaggg 3600

acacctggga gaaatctggc ttctggccag gaagctttgg tgagaacctg 3650

ggttgcagac aggaatctta aggtgtagcc acaccaggat agagactgga 3700

acactagaca agccagaact tgaccctgag ctgaccagcc gtgagcatgt 3750

ttggaagggg tctgtagtgt cactcaaggc ggtgcttgat agaaatgcca 3800

agcacttctt tttctcgctg tcctttctag agcactgcca ccagtaggtt 3850

atttagcttg ggaaaggtgg tgtttctgta agaaacctac tgcccaggca 3900

ctgcaaaccg ccacctccct atactgcttg gagctgagca aatcaccaca 3950

aactgtaata caatgatcct gtattcagac agatgaggac tttccatggg 4000

accacaacta ttttcagatg tgaaccatta accagatcta gtcaatcaag 4050

tctgtttact gcaaggttca acttattaac aattaggcag actctttatg 4100

cttgcaaaaa ctacaaccaa tggaatgtga tgttcatggg tatagttcat 4150

gtctgctatc attattcgta gatattggac aaagaacctt ctctatgggg 4200

catcctcttt ttccaacttg gctgcaggaa tctttaaaag atgcttttaa 4250

cagagtctga acctatttct taaacacttg caacctacct gttgagcatc 4300

acagaatgtg ataaggaaat caacttgctt atcaacttcc taaatattat 4350

gagatgtggc ttgggcagca tccccttgaa ctcttcactc ttcaaatgcc 4400

tgactaggga gccatgtttc acaaggtctt taaagtgact aatggcatga 4450

gaaatacaaa aatactcaga taaggtaaaa tgccatgatg cctctgtctt 4500

ctggactggt tttcacatta gaagacaatt gacaacagtt acataattca 4550

ctctgagtgt tttatgagaa agccttcttt tggggtcaac agttttccta 4600

tgctttgaaa cagaaaaata tgtaccaaga atcttggttt gccttccaga 4650

aaacaaaact gcatttcact ttcccggtgt tccccactgt atctaggcaa 4700

catagtattc atgactatgg ataaactaaa cacgtgacac aaacacacac 4750

aaaagggaac ccagctctaa tacattccaa ctcgtatagc atgcatctgt 4800

ttattctata gttattaagt tctttaaaat gtaaagccat gctggaaaat 4850

aatactgctg agatacatac agaattactg taactgatta cacttggtaa 4900

ttgtactaaa gccaaacata tatatactat taaaaaggtt tacagaattt 4950

tatggtgcat tacgtgggca ttgtcttttt agatgcccaa atccttagat 5000

ctggcatgtt agcccttcct ccaattataa gaggatatga accaaaaaaa 5050

aaaaaaaaaa aa 5062

<210> 3
<211> 1150
<212> DNA
<213> Homo sapiens

<400> 3
ggagtttcaa agaagcgcta gtaaggtctc tgagatcctt gcactagcta 50

catcctcagg gtaggaggaa gatggcttcc agaagcatgc ggctgctcct 100

attgctgagc tgcctggcca aaacaggagt cctgggtgat atcatcatga 150

gacccagctg tgctcctgga tggttttacc acaagtccaa ttgctatggt 200

tacttcagga agctgaggaa ctggtctgat gccgagctcg agtgtcagtc 250

ttacggaaac ggagcccacc tggcatctat cctgagttta aaggaagcca 300

caccaggaaa accaggctac ggaagtcctg gactccaagg agagccaggg 300

ttgccaggac caccgggacc atcagctgta gggaaaccag gtgtgccagg 350

actcccagga aaaccaggag agagaggacc atatggacca aaaggagatg 400

ttggaccagc tggcctacca ggaccccggg gcccaccagg accacctgga 450

atccctggac cggctggaat ttctgtgcca ggaaaacctg gacaacaggg 500

acccacagga gccccaggac ccaggggctt tcctggagaa aagggtgcac 550

caggagtccc tggtatgaat ggacagaaag gggaaatggg atatggtgct 600

cctggtcgtc caggtgagag gggtcttcca ggccctcagg gtcccacagg 650

accatctggc cctcctggag tgggaaaaag aggtgaaaat ggggttccag 700

gacagccagg catcaaaggt gatagaggtt ttccgggaga aatgggacca 750

attggcccac caggtcccca aggccctcct ggggaacgag ggccagaagg 800

cattggaaag ccaggagctg ctggagcccc aggccagcca gggattccag 850

gaacaaaagg tctccctggg gctccaggaa tagctgggcc cccagggcct 900

cctggctttg ggaaaccagg cttgccaggc ctgaagggag aaagaggacc 950

tgctggcctt cctgggggtc caggtgccaa aggggaacaa gggccagcag 1000

gtcttcctgg gaagccaggt ctgactggac cccctgggaa tatgggaccc 1050

caaggaccaa aaggcatccc gggtagccat ggtctcccag gccctaaagg 1100

tgagacaggg ccagctgggc ctgcaggata ccctggggct aagggtgaaa 1150

ggggttcccc tgggtcagat ggaaaaccag ggtacccagg aaaaccaggt 1200

ctcgatggtc ctaagggtaa cccagggtta ccaggtccaa aaggtgatcc 1250

tggagttgga ggacctcctg gtctcccagg ccctgtgggc ccagcaggag 1300

caaagggaat gcccggacac aatggagagg ctggcccaag aggtgccctt 1350

ggaataccag gtactagagg ccctattggg ccaccaggca ttccaggatt 1400

ccctgggtct aaaggggatc caggaagtcc cggtcctcct ggcccagctg 1450

gcatagcaac taagggcctc aatggaccca ccgggccacc agggcctcca 1500

ggtccaagag gcccctctgg agagcctggt cttccagggc cccctgggcc 1550

tccaggccca ccaggtcaag cagtcatgcc tgagggtttt ataaaggcag 1600

gccaaaggcc cagtctttct gggacccctc ttgttagtgc caaccagggg 1650

gtaacaggaa tgcctgtgtc tgcttttact gttattctct ccaaagctta 1700

cccagcaata ggaactccca taccatttga taaaattttg tataacaggc 1750

aacagcatta tgacccaagg actggaatct ttacttgtca gataccagga 1800

atatactatt tttcatacca cgtgcatgtg aaagggactc atgtttgggt 1850

aggcctgtat aagaatggca cccctgtaat gtacacctat gatgaataca 1900

ccaaaggcta cctggatcag gcttcaggga gtgccatcat cgatctcaca 1950

gaaaatgacc aggtgtggct ccagcttccc aatgccgagt caaatggcct 2000

atactcctct gagtatgtcc actcctcttt ctcaggattc ctagtggctc 2050

caatgtgagt acaccccaca gagctaatct aaatcttgtg ctagaaaaag 2100

cattctctaa ctctacccca ccctacaaaa tgcatatgga ggtaggctga 2150

aaagaatgta attttattt tctgaaatac agatttgagc tatcagacca 2200

acaaaccttc cccctgaaaa gtgagcagca acgtaaaaac gtatgtgaag 2250

cctctcttga atttctagtt agcaatctta aggctcttta aggttttctc 2300

caatattaaa aaatatcacc aaagaagtcc tgctatgtta aaaacaaaca 2350

acaaaaaaca aagcaacaaa aaaaaaaatt aaaaaaaaaa acagaaatag 2400

agctctaagt tatgtgaaat ttgatttgag aaactcggca tttccttttt 2450

aaaaaagcct gtttctaact atgaatatga gaacttctag gaaacatcca 2500

ggaggtatca tataactttg tagaacttaa atacttgaat attcaaattt 2550

aaaagacact gtatccccta aaatatttct gatggtgcac tactctgagg 2600

cctgtatggc ccctttcatc aatatctatt caaatataca ggtgcatata 2650

tacttgttaa agctcttata taaaaaagcc ccaaaatatt gaagttcatc 2700

tgaaatgcaa ggtgctttca tcaatgaacc ttttcaaaac ttttctatga 2750

ttgcagagaa gctttttata tacccagcat aacttggaaa caggtatctg 2800

acctattctt atttagttaa cacaagtgtg attaatttga tttctttaat 2850

tccttattga atcttatgtg atatgatttt ctggatttac agaacattag 2900

cacatgtacc ttgtgcctcc cattcaagtg aagttataat ttacactgag 2950

ggtttcaaaa ttcgactaga agtggagata tattatttat ttatgcactg 3000

tactgtattt ttatattgct gtttaaaact tttaagctgt gcctcactta 3050

ttaaagcaca aaatgtttta cctactcctt atttacgaca caataaaata 3100

acatcaatag atttttaggc tgaattaatt tgaaagcagc aatttgctgt 3150

tctcaaccat tctttcaagg cttttcattc gacacaataa aataacatca 3200

atagattttt agggatgggt ggcttt 3226

<210> 6
<211> 1750
<212> DNA
<213> Homo sapiens

<400> 6
catgccgctg ccgccgctgc tgctgttgct cctggcggcg ccttggggac 50

gggcagttcc ctgtgtctct ggtggtttgc ctaaacctgc aaacatcacc 100

ttcttatcca tcaacatgaa gaatgtccta caatggactc caccagaggg 150

tcttcaagga gttaaagtta cttacactgt gcagtatttc atatatgggc 200

aaaagaaatg gctgaataaa tcagaatgca gaaatatcaa tagaacctac 250

tgtgatcttt ctgctgaaac ttctgactac gaacaccagt attatgccaa 300

agttaaggcc atttggggaa caaagtgttc caaatgggct gaaagtggac 350

ggttctatcc ttttttagaa acacaaattg gcccaccaga ggtggcactg 400

actacagatg agaagtccat ttctgttgtc ctgacagctc cagagaagtg 450

gaagagaaat ccagaagacc ttcctgtttc catgcaacaa atatactcca 500

atctgaagta taacgtgtct gtgttgaata ctaaatcaaa cagaacgtgg 550

tcccagtgtg tgaccaacca cacgctggtg ctcacctggc tggagccgaa 600

cactctttac tgcgtacacg tggagtcctt cgtcccaggg ccccctcgcc 650

gtgctcagcc ttctgagaag cagtgtgcca ggactttgaa agatcaatca 700

tcagagttca aggctaaaat catcttctgg tatgttttgc ccatatctat 750

taccgtgttt ctttttttctg tgatgggcta ttccatctac cgatatatcc 800

acgttggcaa agagaaaacac ccagcaaatt tgattttgat ttatggaaat 850

gaatttgaca aaagattctt tgtgcctgct gaaaaaatcg tgattaactt 900

tatcaccctc aatatctcgg atgattctaa aatttctcat caggatatga 950

gtttactggg aaaaagcagt gatgtatcca gccttaatga tcctcagccc 1000

agcgggaacc tgaggccccc tcaggaggaa gaggaggtga aacatttagg 1050

gtatgcttcg catttgatgg aaatttttttg tgactctgaa gaaaacacgg 1100

aaggtacttc tctcacccag caagagtccc tcagcagaac aataccccccg 1150

gataaaacag tcattgaata tgaatatgat gtcagaacca ctgacatttg 1200

tgcggggcct gaagagcagg agctcagttt gcaggaggag gtgtccacac 1250

aaggaacatt attggagtcg caggcagcgt tggcagtctt gggcccgcaa 1300

acgttacagt actcatacac ccctcagctc caagacttag acccccctggc 1350

gcaggagcac acagactcgg aggaggggcc ggaggaagag ccatcgacga 1400

ccctggtcga ctgggatccc caaactggca ggctgtgtat tccttcgctg 1450

tccagcttcg accaggattc agagggctgc gagccttctg agggggatgg 1500

gctcggagag gagggtcttc tatctagact ctatgaggag ccggctccag 1550

acaggccacc aggagaaaat gaaacctatc tcatgcaatt catggaggaa 1600

tgggggttat atgtgcagat ggaaaactga tgccaacact tccttttgcc 1650

ttttgtttcc tgtgcaaaca agtgagtcac ccctttgatc ccagccataa 1700

agtacctggg atgaaagaag ttttttccag tttgtcagtg tctgtgagaa 1750

<210> 7
<211> 1841
<212> DNA
<213> Homo sapiens

<400> 7
agcgggcgtg ggactgagca gtctgctgcc ccccgacatg tgacccagcc 50

ccgccgccca tgcgggctcc cggccgcccg gccctgcggc cgctgccgct 100

gccgccgctg ctgctgttgc tcctggcggc gccttgggga cgggcagttc 150

cctgtgtctc tggtggtttg cctaaacctg caaacatcac cttcttatcc 200

atcaacatga agaatgtcct acaatggact ccaccagagg gtcttcaagg 250

agttaaagtt acttacactg tgcagtattt catatatggg caaaagaaat 300

ggctgaataa atcagaatgc agaaatatca atagaaccta ctgtgatctt 350

tctgctgaaa cttctgacta cgaacaccag tattatgcca aagttaaggc 400

catttgggga acaaagtgtt ccaaatgggc tgaaagtgga cggttctatc 450

ctttttttaga aacacaaatt ggcccaccag aggtggcact gactacagat 500

gagaagtcca tttctgttgt cctgacagct ccagagaagt ggaagagaaa 550

tccagaagac cttcctgttt ccatgcaaca aatatactcc aatctgaagt 600

ataacgtgtc tgtgttgaat actaaatcaa acagaacgtg gtcccagtgt 650

gtgaccaacc acacgctggt gctcacctgg ctggagccga acactcttta 700

ctgcgtacac gtggagtcct tcgtcccagg gccccctcgc cgtgctcagc 750

cttctgagaa gcagtgtgcc aggactttga aagatcaatc atcagagttc 800

aaggctaaaa tcatcttctg gtatgttttg cccatatcta ttaccgtgtt 850

tctttttttct gtgatgggct attccatcta ccgatatatc cacgttggca 900

aagagaaaca cccagcaaat ttgattttga tttatggaaa tgaatttgac 950

aaaagattct ttgtgcctgc tgaaaaaatc gtgattaact ttatcaccct 1000

caatatctcg gatgattcta aaatttctca tcaggatatg agtttactgg 1050

gaaaaagcag tgatgtatcc agccttaatg atcctcagcc cagcgggaac 1100

ctgaggcccc ctcaggagga agaggaggtg aaacatttag ggtatgcttc 1150

gcatttgatg gaaatttttt gtgactctga agaaaacacg gaaggtactt 1200

ctctcaccca gcaagagtcc ctcagcagaa caatacccca ggataaaaca 1250

gtcattgaat atgaatatga tgtcagaacc actgacattt gtgcggggcc 1300

tgaagagcag gagctcagtt tgcaggagga ggtgtccaca caaggaacat 1350

tattggagtc gcaggcagcg ttggcagtct tgggcccgca aacgttacag 1400

tactcataca cccctcagct ccaagactta gacccctgg cgcaggagca 1450

cacagactcg gaggaggggc cggaggaaga gccatcgacg accctggtcg 1500

actgggatcc ccaaactggc aggctgtgta ttccttcgct gtccagcttc 1550

gaccaggatt cagagggctg cgagccttct gaggggatg ggctcggaga 1600

ggagggtctt ctatctagac tctatgagga gccggctcca gacaggccac 1650

caggagaaaa tgaaacctat ctcatgcaat tcatggagga atgggggtta 1700

tatgtgcaga tggaaaactg atgccaacac ttcctttgc cttttgtttc 1750

ctgtgcaaac aagtgagtca ccccttgat cccagccata aagtacctgg 1800

gatgaaagaa gttttttcca gtttgtcagt gtctgtgaga a 1841

<210> 8
<211> 536
<212> DNA
<213> Homo sapiens

<400> 8
gcttggtcta gtgcccattt actctggact ccggatggct gccgcacgcc 50

tctgcctctc cctgctgctc ctgtccacct gcgtggctct gttactacag 100

ccactgctgg gtgcccaggg agccccactg gagccagtgt acccagggga 150

caatgccaca ccagagcaga tggcccagta tgcagctgat ctccgtagat 200

acatcaacat gctgaccagg cctaggtatg ggaaaagaca caaagaggac 250

acgctggcct tctcggagtg ggggtccccg catgctgctg tccccaggga 300

gctcagcccg ctggacttat aatgccacct tctgtctcct acgactccat 350

gagcagcgcc agcccagctc tcccctctgc acccttggct ctggccaaag 400

cttgctccct gctcccacac aggctcaata aagcaagtca aagccaaaaa 450

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 500

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 536

<210> 9
<211> 2588
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 1940
<223> unknown base

<400> 9
gaaggacagc agggccaaca gtcacagcag ccctgaccag agcattcctg 50

gagctcaagc tcctctacaa agaggtggac agagaagaca gcagagacca 100

tgggaccccc ctcagcccct ccctgcagat tgcatgtccc ctggaaggag 150

gtcctgctca cagcctcact tctaaccttc tggaacccac ccaccactgc 200

caagctcact attgaatcca cgccgttcaa tgtcgcagag gggaaggagg 250

ttcttctact cgcccacaac ctgccccaga atcgtattgg ttacagctgg 300

tacaaaggcg aaagagtgga tggcaacagt ctaattgtag gatatgtaat 350

aggaactcaa caagctaccc cagggcccgc atacagtggt cgagagacaa 400

tataccccaa tgcatccctg ctgatccaga acgtcaccca gaatgacaca 450

ggattctata ccctacaagt cataaagtca gatcttgtga atgaagaagc 500

aaccggacag ttccatgtat acccggagct gcccaagccc tccatctcca 550

gcaacaactc caaccccgtg gaggacaagg atgctgtggc cttcacctgt 600

gaacctgagg ttcagaacac aacctacctg tggtgggtaa atggtcagag 650

cctcccggtc agtcccaggc tgcagctgtc caatggcaac atgaccctca 700

ctctactcag cgtcaaaagg aacgatgcag gatcctatga atgtgaaata 750

cagaacccag cgagtgccaa ccgcagtgac ccagtcaccc tgaatgtcct 800

ctatggccca gatggcccca ccatttcccc ctcaaaggcc aattaccgtc 850

caggggaaaa tctgaacctc tcctgccacg cagcctctaa cccacctgca 900

cagtactctt ggtttatcaa tgggacgttc cagcaatcca cacaagagct 950

ctttatcccc aacatcactg tgaataatag cggatcctat atgtgccaag 1000

cccataactc agccactggc ctcaatagga ccacagtcac gatgatcaca 1050

gtctctggaa gtgctcctgt cctctcagct gtggccaccg tcggcatcac 1100

gattggagtg ctggccaggg tggctctgat atagcagccc tggtgtattt 1150

tcgatatttc aggaagactg gcagattgga ccagaccctg aattcttcta 1200

gctcctccaa tcccatttta tcccatggaa ccactaaaaa caaggtctgc 1250

tctgctcctg aagccctata tgctggagat ggacaactca atgaaaattt 1300

aaagggaaaa ccctcaggcc tgaggtgtgt gccactcaga gacttcacct 1350

aactagagac aggcaaactg caaaccatgg tgagaaattg acgacttcac 1400

actatggaca gcttttccca agatgtcaaa acaagactcc tcatcatgat 1450

aaggctctta ccccctttta atttgtcctt gcttatgcct gcctctttcg 1500

cttggcagga tgatgctgtc attagtattt cacaagaagt agcttcagag 1550

ggtaacttaa cagagtatca gatctatctt gtcaatccca acgttttaca 1600

taaaataaga gatcctttag tgcacccagt gactgacatt agcagcatct 1650

ttaacacagc cgtgtgttca aatgtacagt ggtccttttc agagttggac 1700

ttctagactc acctgttctc actccctgtt ttaattcaac ccagccatgc 1750

aatgccaaat aatagaattg ctccctacca gctgaacagg gaggagtctg 1800

tgcagtttct gacacttgtt gttgaacatg gctaaataca atgggtatcg 1850

ctgagactaa gttgtagaaa ttaacaaatg tgctgcttgg ttaaaatggc 1900

tacactcatc tgactcattc tttattctat tttagttggn ttgtatcttg 1950

cctaaggtgc gtagtccaac tcttggtatt accctcctaa tagtcatact 2000

agtagtcata ctccctggtg tagtgtattc tctaaaagct ttaaatgtct 2050

gcatgcagcc agccatcaaa tagtgaatgg tctctctttg gctggaatta 2100

caaaactcag agaaatgtgt catcaggaga acatcataac ccatgaagga 2150

taaaagcccc aaatggtggt aactgataat agcactaatg ctttaagatt 2200

tggtcacact ctcacctagg tgagcgcatt gagccagtgg tgcataaatg 2250

ctacatactc caactgaaat gttaaggaag aagatagatc caattaaaaa 2300

aaattaaaac caatttaaaa aaaaaaagaa cacaggagat tccagtctac 2350

ttgagttagc ataatacaga agtcccctct actttaactt ttacaaaaaa 2400

gtaacctgaa ctaatctgat gttaaccaat gtatttattt ctgtggttct 2450

gtttccttgt tccaatttga caaaacccac tgttcttgta ttgtattgcc 2500

cagggggagc tatcactgta cttgtagagt ggtgctgctt taattcataa 2550

atcacaaata aaagccaatt agctctaaaa aaaaaaaa 2588

<210> 10
<211> 3963
<212> DNA
<213> Homo sapiens

<400> 10
gacctggggg ggcgagggcc ccccaaactc agttcggatc ctacccgagt 50

gaggcggcgc catggagctc cgggtgctgc tctgctgggc ttcgttggcc 100

gcagctttgg aagagaccct gctgaacaca aaattggaaa ctgctgatct 150

gaagtgggtg acattccctc aggtggacgg gcagtgggag gaactgagcg 200

gcctggatga ggaacagcac agcgtgcgca cctacgaagt gtgtgacgtg 250

cagcgtgccc cgggccaggc ccactggctt cgcacaggtt gggtcccacg 300

gcggggcgcc gtccacgtgt acgccacgct gcgcttcacc atgctcgagt 350

gcctgtccct gcctcgggct gggcgctcct gcaaggagac cttcaccgtc 400

ttctactatg agagcgatgc ggacacggcc acggccctca cgccagcctg 450

gatggagaac ccctacatca aggtggacac ggtggccgcg gagcatctca 500

cccggaagcg ccctggggcc gaggccaccg ggaaggtgaa tgtcaagacg 550

ctgcgtctgg gaccgctcag caaggctggc ttctacctgg ccttccagga 600

ccagggtgcc tgcatggccc tgctatccct gcacctcttc tacaaaaagt 650

gcgcccagct gactgtgaac ctgactcgat tcccggagac tgtgcctcgg 700

gagctggttg tgcccgtggc cggtagctgc gtggtggatg ccgtccccgc 750

ccctggcccc agccccagcc tctactgccg tgaggatggc cagtgggccg 800

aacagccggt cacgggctgc agctgtgctc cggggttcga ggcagctgag 850

gggaacacca agtgccgagc ctgtgcccag ggcaccttca agcccctgtc 900

aggagaaggg tcctgccagc catgcccagc caatagccac tctaacacca 950

ttggatcagc cgtctgccag tgccgcgtcg ggtacttccg ggcacgcaca 1000

gacccccggg gtgcaccctg caccaccccc ccttcggctc cgcggagcgt 1050

ggtttcccgc ctgaacggct cctccctgca cctggaatgg agtgcccccc 1100

tggagtctgg tggccgagag gacctcacct acgccctccg ctgccgggag 1150

tgccgacccg gaggctcctg tgcgccctgc gggggagacc tgacttttga 1200

ccccggcccc cgggacctgg tggagccctg ggtggtggtt cgagggctac 1250

gtcctgactt cacctatacc tttgaggtca ctgcattgaa cggggtatcc 1300

tccttagcca cggggcccgt cccatttgag cctgtcaatg tcaccactga 1350

ccgagaggta cctcctgcag tgtctgacat ccgggtgacg cggtcctcac 1400

ccagcagctt gagcctggcc tgggctgttc cccgggcacc cagtggggct 1450

gtgctggact acgaggtcaa ataccatgag aagggcgccg agggtcccag 1500

cagcgtgcgg ttcctgaaga cgtcagaaaa ccgggcagag ctgcgggggc 1550

tgaagcgggg agccagctac ctggtgcagg tacgggcgcg ctctgaggcc 1600

ggctacgggc ccttcggcca ggaacatcac agccagaccc aactggatga 1650

gagcgagggc tggcgggagc agctggccct gattgcgggc acggcagtcg 1700

tgggtgtggt cctggtcctg gtggtcattg tggtcgcagt tctctgcctc 1750

aggaagcaga gcaatgggag agaagcagaa tattcggaca aacacggaca 1800

gtatctcatc ggacatggta ctaaggtcta catcgacccc ttcacttatg 1850

aagaccctaa tgaggctgtg agggaatttg caaaagagat cgatgtctcc 1900

tacgtcaaga ttgaagaggt gattggtgca ggtgagtttg gcgaggtgtg 1950

tcgggggcgg ctcaaggccc cagggaagaa ggagagctgt gtggcaatca 2000

agaccctgaa gggtggctac acggagcggc agcggcgtga gtttctgagc 2050

gaggcctcca tcatgggcca gttcgagcac cccaatatca tccgcctgga 2100

gggcgtggtc accaacagca tgcccgtcat gattctcaca gagttcatgg 2150

agaacggcgc cctggactcc ttcctgcggc taaacgacgg acagttcaca 2200

gtcatccagc tcgtgggcat gctgcggggc atcgcctcgg gcatgcggta 2250

ccttgccgag atgagctacg tccaccgaga cctggctgct cgcaacatcc 2300

tagtcaacag caacctcgtc tgcaaagtgt ctgactttgg cctttcccga 2350

ttcctggagg agaactcttc cgatcccacc tacacgagct ccctgggagg 2400

aaagattccc atccgatgga ctgccccgga ggccattgcc ttccggaagt 2450

tcacttccgc cagtgatgcc tggagttacg ggattgtgat gtgggaggtg 2500

atgtcatttg gggagaggcc gtactgggac atgagcaatc aggacgtgat 2550

caatgccatt gaacaggact accggctgcc cccgccccca gactgtccca 2600

cctccctcca ccagctcatg ctggactgtt ggcagaaaga ccggaatgcc 2650

1342
1343

cggccccgct tcccccaggt ggtcagcgcc ctggacaaga tgatccggaa 2700

ccccgccagc ctcaaaatcg tggcccggga gaatggcggg gcctcacacc 2750

ctctcctgga ccagcggcag cctcactact cagcttttgg ctctgtgggc 2800

gagtggcttc gggccatcaa aatgggaaga tacgaagaaa gtttcgcagc 2850

cgctggcttt ggctccttcg agctggtcag ccagatctct gctgaggacc 2900

tgctccgaat cggagtcact ctggcgggac accagaagaa aatcttggcc 2950

agtgtccagc acatgaagtc ccaggccaag ccgggaaccc cgggtgggac 3000

aggaggaccg gccccgcagt actgacctgc aggaactccc caccccaggg 3050

acaccgcctc cccattttcc ggggcagagt ggggactcac agaggccccc 3100

agccctgtgc cccgctggat tgcactttga gcccgtgggg tgaggagttg 3150

gcaatttgga gagacaggat ttgggggttc tgccataata ggaggggaaa 3200

atcacccccc agccacctcg gggaactcca gaccaagggt gagggcgcct 3250

ttccctcagg actgggtgtg accagaggaa aaggaagtgc ccaacatctc 3300

ccagcctccc caggtgcccc cctcaccttg atgggtgcgt tcccgcagac 3350

caaagagagt gtgactccct tgccagctcc agagtggggg ggctgtccca 3400

gggggcaaga aggggtgtca gggcccagtg acaaaatcat tggggtttgt 3450

agtcccaact tgctgctgtc accaccaaac tcaatcattt ttttcccttg 3500

taaatgcccc tcccccagct gctgccttca tattgaaggt ttttgagttt 3550

tgtttttggt cttaattttt ctccccgttc ccttttttgtt tcttcgtttt 3600

gtttttctac cgtccttgtc ataactttgt gttggaggga acctgtttca 3650

ctatggcctc ctttgcccaa gttgaaacag gggcccatca tcatgtctgt 3700

ttccagaaca gtgccttggt catcccacat ccccggaccc cgcctgggac 3750

ccccaagctg tgtcctatga aggggtgtgg ggtgaggtag tgaaaagggc 3800

ggtagttggt ggtggaaccc agaaacggac gccggtgctt ggaggggttc 3850

ttaaattata tttaaaaaag taacttttg tataaataaa agaaaatggg 3900

acgtgtccca aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3950

aaaaaaaaaa aga 3963

<210> 11
<211> 3476
<212> DNA
<213> Homo sapiens

<400> 11
gctctatgcc gcctaccttg ctctcgccgc tgctgccgga gccgaagcag 50

agaaggcagc gggtcccgtg accgtcccga gagccccgcg ctcccgacca 100

gggggcgggg gcggccccgg ggagggcggg gcaggggcgg ggggaagaaa 150

gggggttttg tgctgcgccg ggagggccgg cgccctcttc cgaatgtcct 200

gcggccccag cctctcctca cgctcgcgca gtctccgccg cagtctcagc 250

tgcagctgca ggactgagcc gtgcacccgg aggagacccc cggaggaggc 300

gacaaacttc gcagtgccgc gacccaaccc cagccctggg tagcctgcag 350

catggcccag ctgttcctgc ccctgctggc agccctggtc ctggcccagg 400

ctcctgcagc tttagcagat gttctggaag gagacagctc agaggaccgc 450

gcttttcgcg tgcgcatcgc gggcgacgcg ccactgcagg gcgtgctcgg 500

cggcgccctc accatccctt gccacgtcca ctacctgcgg ccaccgccga 550

gccgccgggc tgtgctgggc tctccgcggg tcaagtggac tttcctgtcc 600

cggggccggg aggcagaggt gctggtggcg cggggagtgc gcgtcaaggt 650

gaacgaggcc taccggttcc gcgtggcact gcctgcgtac ccagcgtcgc 700

tcaccgacgt ctccctggcg ctgagcgagc tgcgccccaa cgactcaggt 750

atctatcgct gtgaggtcca gcacggcatc gatgacagca gcgacgctgt 800

ggaggtcaag gtcaaagggg tcgtctttct ctaccgagag ggctctgccc 850

gctatgcttt ctccttttct ggggcccagg aggcctgtgc ccgcattgga 900

gcccacatcg ccaccccgga gcagctctat gccgcctacc ttgggggcta 950

tgagcaatgt gatgctggct ggctgtcgga tcagaccgtg aggtatccca 1000

tccagacccc acgagaggcc tgttacggag acatggatgg cttccccggg 1050

gtccggaact atggtgtggt ggacccggat gacctctatg atgtgtactg 1100

ttatgctgaa gacctaaatg gagaactgtt cctgggtgac cctccagaga 1150

agctgacatt ggaggaagca cgggcgtact gccaggagcg gggtgcagag 1200

attgccacca cgggccaact gtatgcagcc tgggatggtg gcctggacca 1250

ctgcagccca gggtggctag ctgatggcag tgtgcgctac cccatcgtca 1300

cacccagcca gcgctgtggt gggggcttgc ctggtgtcaa gactctcttc 1350

ctcttcccca accagactgg cttccccaat aagcacagcc gcttcaacgt 1400

ctactgcttc cgagactcgg cccagccttc tgccatccct gaggcctcca 1450

acccagcctc caacccagcc tctgatggac tagaggctat cgtcacagtg 1500

acagagaccc tggaggaact gcagctgcct caggaagcca cagagagtga 1550

atcccgtggg gccatctact ccatccccat catggaggac ggaggaggtg 1600

gaagctccac tccagaagac ccagcagagg cccctaggac gctcctagaa 1650

tttgaaacac aatccatggt accgcccacg gggttctcag aagaggaagg 1700

taaggcattg gaggaagaag agaaatatga agatgaagaa gagaaagagg 1750

aggaagaaga agaggaggag gtggaggatg aggctctgtg ggcatggccc 1800

agcgagctca gcagcccggg ccctgaggcc tctctcccca ctgagccagc 1850

agcccaggag aagtcactct cccaggcgcc agcaagggca gtcctgcagc 1900

ctggtgcatc accacttcct gatggagagt cagaagcttc caggcctcca 1950

agggtccatg gaccacctac tgagactctg cccactccca gggagaggaa 2000

cctagcatcc ccatcacctt ccactctggt tgaggcaaga gaggtggggg 2050

aggcaactgg tggtcctgag ctatctgggg tccctcgagg agagagcgag 2100

gagacaggaa gctccgaggg tgccccttcc ctgcttccag ccacacgggc 2150

ccctgagggt accagggagc tggaggcccc ctctgaagat aattctggaa 2200

gaactgcccc agcagggacc tcagtgcagg cccagccagt gctgcccact 2250

gacagcgcca gccgaggtgg agtggccgtg gtccccgcat caggtgactg 2300

tgtccccagc ccctgccaca atggtgggac atgcttggag gaggaggaag 2350

gggtccgctg cctatgtctg cctggctatg gggggacct gtgcgatgtt 2400

ggcctccgct tctgcaaccc cggctgggac gccttccagg gcgcctgcta 2450

caagcacttt tccacacgaa ggagctggga ggaggcagag acccagtgcc 2500

ggatgtacgg cgcgcatctg gccagcatca gcacacccga ggaacaggac 2550

ttcatcaaca accggtaccg ggagtaccag tggatcggac tcaacgacag 2600

gaccatcgaa ggcgacttct tgtggtcgga tggcgtcccc ctgctctatg 2650

agaactggaa ccctgggcag cctgacagct acttcctgtc tggagagaac 2700

tgcgtggtca tggtgtggca tgatcaggga caatggagtg acgtgccctg 2750

caactaccac ctgtcctaca cctgcaagat ggggctggtg tcctgtgggc 2800

cgccaccgga gctgcccctg gctcaagtgt tcggccgccc acggctgcgc 2850

tatgaggtgg acactgtgct tcgctaccgg tgccgggaag gactggccca 2900

gcgcaatctg ccgctgatcc gatgccaaga gaacggtcgt tgggaggccc 2950

cccagatctc ctgtgtgccc agaagacctg cccgagctct gcacccagag 3000

gaggacccag aaggacgtca ggggaggcta ctgggacgct ggaaggcgct 3050

gttgatcccc ccttccagcc ccatgccagg tccctagggg gcaaggcctt 3100

gaacactgcc ggccacagca ctgccctgtc acccaaattt tccctcacac 3150

cttgcgctcc cgccaccaca ggaagtgaca acatgacgag gggtggtgct 3200

ggagtccagg tgacagttcc tgaaggggct tctgggaaat acctaggagg 3250

ctccagccca gcccaggccc tctcccccta ccctgggcac cagatcttcc 3300

atcagggccg gagtaaatcc ctaagtgcct caactgccct ctccctggca 3350

gccatcttgt cccctctatt cctctaggga gcactgtgcc cactctttct 3400

gggttttcca agggaatggg cttgcaggat ggagtgtctg taaaatcaac 3450

aggaaataaa actgtgtatg agccca 3476

<210> 12
<211> 2558
<212> DNA
<213> Homo sapiens

<400> 12
tgtggcactg cctgcgtacc caaccccagc cctgggtagc ctgcagcatg 50

gcccagctgt tcctgccccc gctggcagcc ctggtcctgg cccaggctcc 100

tgcagcttta gcagatgttc tggaaggaga cagctcagag gaccgcgctt 150

ttcgcgtgcg catcgcgggc gacgcgccac tgcagggcgt gctcggcggc 200

gccctcacca tcccttgcca cgtccactac ctgcggccac cgccgagccg 250

ccgggctgtg ctgggctctc cgcgggtcaa gtggactttc ctgtcccggg 300

gccgggaggc agaggtgctg gtggcgcggg gagtgcgcgt caaggtgaac 350

gaggcctacc ggttccgcgt ggcactgcct gcgtacccag cgtcgctcac 400

cgacgtctcc ctggcgctga gcgagctgcg ccccaacgac tcaggtatct 450

atcgctgtga ggtccagcac ggcatcgatg acagcagcga cgctgtggag 500

gtcaaggtca aaggggtcgt ctttctctac cgagagggct ctgcccgcta 550

tgctttctcc ttttctgggg cccaggaggc ctgtgcccgc attggagccc 600

acatcgccac cccggagcag ctctatgccg cctaccttgg gggctatgag 650

caatgtgatg ctggctggct gtcggatcag accgtgaggt atcccatcca 700

gaccccacga gaggcctgtt acggagacat ggatggcttc cccggggtcc 750

ggaactatgg tgtggtggac ccggatgacc tctatgatgt gtactgttat 800

gctgaagacc taaatggaga attgttcctg ggtgaccctc cagagaagct 850

gacattggag gaagcacggg cgtactgcca ggagcggggt gcagagattg 900

ccaccacggg ccaactgtat gcagcctggg atggtggcct ggaccactgc 950

agcccagggt ggctagctga tggcagtgtg cgctacccca tcgtcacacc 1000

cagccagcgc tgtggtgggg gcttgcctgg tgtcaagact ctcttcctct 1050

tccccaacca gactggcttc cccaataagc acagccgctt caacgtctac 1100

tgcttccgag actcggccca gccttctgcc atccctgagg cctccaaccc 1150

agcctccaac ccagcctctg atggactaga ggctatcgtc acagtgacag 1200

agaccctgga ggaactgcag ctgcctcagg aagccacaga gagtgaatcc 1250

cgtggggcca tctactccat ccccatcatg gaggacggag gaggtggaag 1300

ctccactcca gaagacccag cagaggcccc taggacgctc ctagaatttg 1350

aaacacaatc catggtaccg cccacggggt tctcagaaga ggaaggtaag 1400

gcattggagg aagaagagaa atatgaagat gaagaagaga aagaggagga 1450

agaagaagag gaggaggtgg aggatgaggc tctgtgggca tggcccagcg 1500

agctcagcag cccgggccct gaggcctctc tccccactga gccagcagcc 1550

caggaggagt cactctccca ggcgccagca agggcagtcc tgcagcctgg 1600

tgcatcacca cttcctgatg gagagtcaga agcttccagg cctccaaggg 1650

tccatggacc acctactgag actctgccca ctcccaggga gaggaaccta 1700

gcatccccat caccttccac tctggttgag gcaagagagg tgggggaggc 1750

aactggtggt cctgagctat ctggggtccc tcgaggagag agcgaggaga 1800

caggaagctc cgagggtgcc ccttccctgc ttccagccac acgggccccc 1850

gagggtacca gggagctgga ggccccctct gaagataatt ctggaagaac 1900

tgccccagca gggacctcag tgcaggccca gccagtgctg cccactgaca 1950

gcgccagccg aggtggagtg gccgtggtcc ccgcatcagg taattctgcc 2000

caaggctcaa ctgccctctc tatcctactc cttttcttcc ccctgcagct 2050

ctgggtcacc tgacctgtag tcctttaacc caccatcatc ccaaactctc 2100

ctgtcctttg ccttcattct cttacccacc tctacctatg ggtctccaat 2150

ctcggatatc caccttgtgg gtatctcagc tctccgcgtc tttaccctgt 2200

gatcccagcc ccgccactga ccatctgtga cccttccctg ccattgggcc 2250

ctccacctgt ggctcacatc tcgccagccc cacagagcat cctcaggcct 2300

ctccaagggt cctcatcacc tattgcagcc ttcagggctc ggcctatttt 2350

ccactactcc cttcatccgc ctgtgtgccg tccccctttag ctgcctccta 2400

ttgatctcag ggaagcctgg gagtcccttc tcacccctca acctccggag 2450

tccaggagaa cccgtacccc cacagagcct taagcaacta cttctgtgaa 2500

gtatttttg actgtttcat ggaaaacaag ccttggaaat aaatctctat 2550

taaaccgc 2558

<210> 13
<211> 401
<212> DNA
<213> Homo sapiens

<400> 13
agccagctca ggctacacta tcccaggatc agcatggccg tccgccagtg 50

ggtaatcgcc ctggccttgg ctgccctcct tgttgtggac agggaagtgc 100

cagtggcagc aggaaagctc cctttctcaa gaatgcccat ctgtgaacac 150

atggtagagt ctccaacctg ttcccagatg tccaacctgg tctgcggcac 200

tgatgggctc acatatacga atgaatgcca gctctgcttg gcccggataa 250

aaaccaaaca ggacatccag atcatgaaag atggcaaatg ctgatcccac 300

aggagcacct caagccatga agtgtcagct ggagaacagt ggtgggcatg 350

gagaggatat gacatgaaat aaaagatcca gcccaaaaaa aaaagaaaaa 400

a 401

<210> 14
<211> 1572
<212> DNA
<213> Homo sapiens

<400> 14
cgtttccaat gcacgtacag cccgtacaca ccgtgtgctg ggacacccca 50

cagtcagccg catggctccc ctgtgcccca gcccctggct ccctctgttg 100

atcccggccc ctgctccagg cctcactgtg caactgctgc tgtcactgct 150

gcttctggtg cctgtccatc cccagaggtt gccccggatg caggaggatt 200

ccccccctggg aggaggctct tctggggaag atgacccact gggcgaggag 250

gatctgccca gtgaagagga ttcacccaga gaggaggatc cacccggaga 300

ggaggatcta cctggagagg aggatctacc tggagaggag gatctacctg 350

aagttaagcc taaatcagaa gaagagggct ccctgaagtt agaggatcta 400

cctactgttg aggctcctgg agatcctcaa gaaccccaga ataatgccca 450

cagggacaaa gaaggggatg accagagtca ttggcgctat ggaggcgacc 500

cgccctggcc ccgggtgtcc ccagcctgcg cgggccgctt ccagtccccg 550

gtggatatcc gcccccagct cgccgccttc tgcccggccc tgcgcccccc 600

ggaactcctg ggcttccagc tcccgccgct cccagaactg cgcctgcgca 650

acaatggcca cagtgtgcaa ctgaccctgc ctcctgggct agagatggct 700

ctgggtcccg ggcgggagta ccgggctctg cagctgcatc tgcactgggg 750

ggctgcaggt cgtccgggct cggagcacac tgtggaaggc caccgtttcc 800

ctgccgagat ccacgtggtt cacctcagca ccgcctttgc cagagttgac 850

gaggccttgg ggcgcccggg aggcctggcc gtgttggccg cctttctgga 900

ggagggcccg gaagaaaaca gtgcctatga gcagttgctg tctcgcttgg 950

aagaaatcgc tgaggaaggc tcagagactc aggtcccagg actggacata 1000

tctgcactcc tgccctctga cttcagccgc tacttccgat atgaggggtc 1050

tctgactaca ccgccctgtg cccagggtgt catctggact gtgtttaacc 1100

agacagtgat gctgagtgct aagcagctcc acaccctctc tgacaccctg 1150

tggggacctg gtgactctcg gctacagctg aacttccgag cgacgcagcc 1200

tttgaatggg cgagtgattg aggcctcctt ccctgctgga gtggacagca 1250

gtcctcgggc tgctgagcca gtccagctga attcctgcct ggctgctggt 1300

gacatcctag ccctggtttt tggcctcctt tttgctgtca ccagcgtcgc 1350

gttccttgtg cagatgagaa ggcagcacag aaggggaacc aaaggggtg 1400

tgagctaccg cccagcagag gtagccgaga ctggagccta gaggctggat 1450

cttggagaat gtgagaagcc agccagaggc atctgagggg gagccggtaa 1500

ctgtcctgtc ctgctcatta tgccacttcc ttttaactgc caagaaattt 1550

tttaaaataa atatttataa ta 1572

<210> 15
<211> 2573
<212> DNA
<213> Homo sapiens

<400> 15
tccacatcca ggtctttgtg ctcctcgctt gcctgttcct tttccacgca 50

ttttccagga taactgtgac tccaggcccg caatggatgc cctgcaacta 100

gcaaattcgg cttttgccgt tgatctgttc aaacaactat gtgaaaagga 150

gccactgggc aatgtcctct tctctccaat ctgtctctcc acctctctgt 200

cacttgctca agtgggtgct aaaggtgaca ctgcaaatga aattggacag 250

gttcttcatt ttgaaaatgt caaagatgta ccctttggat ttcaaacagt 300

aacatcggat gtaaacaaac ttagttcctt ttactcactg aaactaatca 350

agcggctcta cgtagacaaa tctctgaatc tttctacaga gttcatcagc 400

tctacgaaga gaccgtatgc aaaggaattg gaaactgttg acttcaaaga 450

taaattggaa gaaacgaaag gtcagatcaa caactcaatt aaggatctca 500

cagatggcca ctttgagaac attttagctg acaacagtgt gaacgaccag 550

accaaaatcc ttgtggttaa tgctgcctac tttgttggca agtggatgaa 600

gaaattttct gaatcagaaa caaaagaatg tcctttcaga gtcaacaaga 650

cagacaccaa accagtgcag atgatgaaca tggaggccac gttctgtatg 700

ggaaacattg acagtatcaa ttgtaagatc atagagcttc cttttcaaaa 750

taagcatctc agcatgttca tcctactacc caaggatgtg gaggatgagt 800

ccacaggctt ggagaagatt gaaaaacaac tcaactcaga gtcactgtca 850

cagtggacta atcccagcac catggccaat gccaaggtca aactctccat 900

tccaaaattt aaggtggaaa agatgattga tcccaaggct tgtctggaaa 950

atctagggct gaaacatatc ttcagtgaag acacatctga tttctctgga 1000

atgtcagaga ccaagggagt ggccctatca aatgttatcc acaaagtgtg 1050

cttagaaata actgaagatg gtggggattc catagaggtg ccaggagcac 1100

ggatcctgca gcacaaggat gaattgaatg ctgaccatcc ctttatttac 1150

atcatcaggc acaacaaaac tcgaaacatc attttctttg gcaaattctg 1200

ttctccttaa gtggcatagc ccatgttaag tcctccctga cttttctgtg 1250

gatgccgatt tctgtaaact ctgcatccag agattcattt tctagataca 1300

ataaattgct aatgttgctg gatcaggaag ccgccagtac ttgtcatatg 1350

tagccttcac acagatagac ctttttttt tttccaattc tatcttttgt 1400

ttcctttttt cccataagac aatgacatac gcttttaatg aaaaggaatc 1450

acgttagagg aaaaatattt attcattatt tgtcaaattg tccgggttag 1500

ttggcagaaa tacagtcttc cacaaagaaa attcctataa ggaagatttg 1550

gaagctcttc ttcccagcac tatgctttcc ttctttggga tagagaatgt 1600

tccagacatt ctcgcttccc tgaaagactg aagaaagtgt agtgcatggg 1650

acccacgaaa ctgccctggc tccagtgaaa cttgggcaca tgctcaggct 1700

actataggtc cagaagtcct tatgttaagc cctggcaggc aggtgtttat 1750

taaaattctg aattttgggg attttcaaaa gataatattt tacatacact 1800

gtatgttata gaacttcatg gatcagatct ggggcagcac cctataaatc 1850

aacaccttaa tatgctgcaa caaaatgtag aatattcaga caaaatggat 1900

acataaagac taagtagccc ataaggggtc aaaatttgct gccaaatgcg 1950

tatgccacca acttacaaaa acacttcgtt cgcagagctt ttcagattgt 2000

ggaatgttgg ataaggaatt atagacctct agtagctgaa atgcaagacc 2050

ccaagaggaa gttcagatct taatataaat tcactttcat ttttgatagc 2100

tgtcccatct ggtcatttgg ttggcactag actggtggca ggggcttcta 2150

gctgacttgc acagggattc tcacaatagc cgatatcaga atttgtgttg 2200

aaggaacttg tctcttcatc taatatgata gcgggaaaag gagaggaaac 2250

tactgccttt agaaaatata agtaaagtga ttaaagtgct cacgttacct 2300

tgacacatag tttttcagtc tatgggttta gttactttag atggcaagca 2350

tgtaacttat attaatagta atttgtaaag ttggttggat aagctatccg 2400

tgttgcaggt tcatggatta cttctctata aaaaatatgt atttaccaaa 2450

aattttgtga cattccttct cccatctctt ccttgacctg cattgtaaat 2500

aggttcttct tgttctgaga ttcaatattg aatttttcct atgctattga 2550

caataaaata ttattgaact aca 2573

<210> 16
<211> 1145
<212> DNA
<213> Homo sapiens

<400> 16
gccgagccgg ggctccggaa gccggcgggg gcgccgcggc cgtgcggggc 50

gtcaatggat cgccactcca gctacatctt catctggctg cagctggagc 100

tctgcgccat ggccgtgctg ctcaccaaag gtgaaattcg atgctactgt 150

gatgctgccc actgtgtagc cactggttat atgtgtaaat ctgagctcag 200

cgcctgcttc tctagacttc ttgatcctca gaactcaaat tccccactca 250

cccatggctg cctggactct cttgcaagca cgacagacat ctgccaagcc 300

~~1350~~
1349

aaacaggccc gaaaccactc tggcaccacc atacccacat tggaatgctg 350

tcatgaagac atgtgcaatt acagagggct gcacgatgtt ctctctcctc 400

ccaggggtga ggcctcagga caaggaaaca ggtatcagca tgatggtagc 450

agaaacctta tcaccaaggt gcaggagctg acttcttcca aagagttgtg 500

gttccgggca gcggtcattg ccgtgcccat tgctggaggg ctgattttag 550

tgttgcttat tatgttggcc ctgaggatgc ttcgaagtga aaataagagg 600

ctgcaggatc agcggcaaca gatgctctcc cgtttgcact acagctttca 650

cggacaccat tccaaaaagg ggcaggttgc aaagttagac ttggaatgca 700

tggtgccggt cagtgggcac gagaactgct gtctgacctg tgataaaatg 750

agacaagcag acctcagcaa cgataagatc ctctcgcttg ttcactgggg 800

catgtacagt gggcacggga agctggaatt cgtatgacgg agtcttatct 850

gaactacact tactgaacag cttgaaggcc ttttgagttc tgctggacag 900

gagcacttta tctgaagaca aactcattta atcatctttg acagacaaaa 950

tgacctctgc aaacagaatc ttggatattt cttctgaagg attatttgca 1000

cagacttaaa tacagttaaa tgtgttattt gcttttaaaa ttataaaaag 1050

caaagagaag actttgtaca cactgtcacc agggttattt gcatccaagg 1100

gagctggaat tgagtaccta aataaacaaa aatgtgccct aaaaa 1145

<210> 17
<211> 5668
<212> DNA
<213> Homo sapiens

<400> 17
gtcacttagg aaaaggtgtc ctttcgggca gccgggctca gcatgaggaa 50

cagaaggaat gacactctgg acagcacccg gaccctgtac tccagcgcgt 100

ctcggagcac agacttgtct tacagtgaaa gcgacttggt gaattttatt 150

caagcaaatt ttaagaaacg agaatgtgtc ttctttacca aagattccaa 200

ggccacggag aatgtgtgca agtgtggcta tgcccagagc cagcacatgg 250

aaggcaccca gatcaaccaa agtgagaaat ggaactacaa gaaacacacc 300

aaggaatttc ctaccgacgc ctttggggat attcagtttg agacactggg 350

gaagaaaggg aagtatatac gtctgtcctg cgacacggac gcggaaatcc 400

tttacgagct gctgacccag cactggcacc tgaaaacacc caacctggtc 450

atttctgtga ccgggggcgc caagaacttc gccctgaagc cgcgcatgcg 500

caagatcttc agccggctca tctacatcgc gcagtccaaa ggtgcttgga 550

ttctcacggg aggcacccat tatggcctga cgaagtacat cggggaggtg 600

gtgagagata acaccatcag caggagttca gaggagaata ttgtggccat 650

tggcatagca gcttggggca tggtctccaa ccgggacacc ctcatcagga 700

attgcgatgc tgagggctat tttttagccc agtaccttat ggatgacttc 750

acaagggatc cactgtatat cctggacaac aaccacacac atttgctgct 800

cgtggacaat ggctgtcatg gacatcccac tgtcgaagca aagctccgga 850

atcagctaga gaagcatatc tctgagcgca ctattcaaga ttccaactat 900

ggtggcaaga tccccattgt gtgttttgcc caaggaggtg gaaaagagac 950

tttgaaagcc atcaatacct ccatcaaaaa taaaattcct tgtgtggtgg 1000

tggaaggctc gggccggatc gctgatgtga tcgctagcct ggtggaggtg 1050

gaggatgccc cgacatcttc tgccgtcaag gagaagctgg tgcgcttttt 1100

accccgcacg gtgtcccggc tgtctgagga ggagactgag agttggatca 1150

aatggctcaa agaaattctc gaatgttctc acctattaac agttattaaa 1200

atggaagaag ctggggatga aattgtgagc aatgccatct cctacgctct 1250

atacaaagcc ttcagcacca gtgagcaaga caaggataac tggaatgggc 1300

agctgaagct tctgctggag tggaaccagc tggacttagc caatgatgag 1350

attttcacca atgaccgccg atgggagtct gctgaccttc aagaagtcat 1400

gtttacggct ctcataaagg acagacccaa gtttgtccgc ctctttctgg 1450

agaatggctt gaacctacgg aagtttctca cccatgatgt cctcactgaa 1500

ctcttctcca accacttcag cacgcttgtg taccggaatc tgcagatcgc 1550

caagaattcc tataatgatg ccctcctcac gtttgtctgg aaactggttg 1600

cgaacttccg aagaggcttc cggaaggaag acagaaatgg ccgggacgag 1650

atggacatag aactccacga cgtgtctcct attactcggc acccccctgca 1700

agctctcttc atctgggcca ttcttcagaa taagaaggaa ctctccaaag 1750

tcatttggga gcagaccagg ggctgcactc tggcagccct gggagccagc 1800

aagcttctga agactctggc caaagtgaag aacgacatca atgctgctgg 1850

ggagtccgag gagctggcta atgagtacga gacccgggct gttgagctgt 1900

tcactgagtg ttacagcagc gatgaagact tggcagaaca gctgctggtc 1950

tattcctgtg aagcttgggg tggaagcaac tgtctggagc tggcggtgga 2000

ggccacagac cagcatttca ccgcccagcc tggggtccag aattttcttt 2050

ctaagcaatg gtatggagag atttcccgag acaccaagaa ctggaagatt 2100

atcctgtgtc tgtttattat acccttggtg ggctgtggct ttgtatcatt 2150

taggaagaaa cctgtcgaca agcacaagaa gctgctttgg tactatgtgg 2200

cgttcttcac ctccccccttc gtggtcttct cctggaatgt ggtcttctac 2250

atcgccttcc tcctgctgtt tgcctacgtg ctgctcatgg atttccattc 2300

ggtgccacac ccccccgagc tggtcctgta ctcgctggtc tttgtcctct 2350

tctgtgatga agtgagacag tggtacgtaa atggggtgaa ttattttact 2400

gacctgtgga atgtgatgga cacgctgggg ctttttttact tcatagcagg 2450

aattgtattt cggctccact cttctaataa aagctctttg tattctggac 2500

gagtcatttt ctgtctggac tacattattt tcactctaag attgatccac 2550

atttttactg taagcagaaa cttaggaccc aagattataa tgctgcagag 2600

gatgctgatc gatgtgttct tcttcctgtt cctctttgcg gtgtggatgg 2650

tggcctttgg cgtggccagg caagggatcc ttaggcagaa tgagcagcgc 2700

tggaggtgga tattccgttc ggtcatctac gagccctacc tggccatgtt 2750

cggccaggtg cccagtgacg tggatggtac cacgtatgac tttgcccact 2800

gcaccttcac tgggaatgag tccaagccac tgtgtgtgga gctggatgag 2850

cacaacctgc cccggttccc cgagtggatc accatccccc tggtgtgcat 2900

ctacatgtta tccaccaaca tcctgctggt caacctgctg gtcgccatgt 2950

ttggctacac ggtgggcacc gtccaggaga acaatgacca ggtctggaag 3000

ttccagaggt acttcctggt gcaggagtac tgcagccgcc tcaatatccc 3050

cttccccttc atcgtcttcg cttacttcta catggtggtg aagaagtgct 3100

tcaagtgttg ctgcaaggag aaaaacatgg agtcttctgt ctgctgtttc 3150

aaaaatgaag acaatgagac tctggcatgg gagggtgtca tgaaggaaaa 3200

ctaccttgtc aagatcaaca caaaagccaa cgacacctca gaggaaatga 3250

ggcatcgatt tagacaactg gatacaaagc ttaatgatct caagggtctt 3300

ctgaaagaga ttgctaataa aatcaaataa aactgtatga aactctaatg 3350

gagaaaaatc taattatagc aagatcatat taaggaatgc tgatgaacaa 3400

ttttgctatc gactactaaa tgagagattt tcagacccct gggtacatgg 3450

tggatgattt taaatcaccc tagtgtgctg agaccttgag aataaagtgt 3500

gtgattggtt tcatacttga agacggatat aaaggaagaa tatttccttt 3550

atgtgtttct ccagaatggt gcctgtttct ctctgtgtct caatgcctgg 3600

gactggaggt tgatagttta agtgtgttct taccgcctcc tttttccttt 3650

aatcttattt ttgatgaaca catatatagg agaacatcta tcctatgaat 3700

aagaacctgg tcatgcttta ctcctgtatt gttattttgt tcatttccaa 3750

ttgattctct acttttccct tttttgtatt atgtgactaa ttagttggca 3800

tattgttaaa agtctctcaa attaggccag attctaaaac atgctgcagc 3850

aagaggaccc cgctctcttc aggaaaagtg ttttcatttc tcaggatgct 3900

tcttacctgt cagaggaggt gacaaggcag tctcttgctc tcttggactc 3950

accaggctcc tattgaagga accaccccca ttcctaaata tgtgaaaagt 4000

cgcccaaaat gcaaccttga aaggcactac tgactttgtt cttattggat 4050

actcctctta tttattattt ttccattaaa aataatagct ggctattata 4100

gaaaatttag accatacaga gatgtagaaa gaacataaat tgtccccatt 4150

accttaaggt aatcactgct aacaatttct ggatggtttt tcaagtctat 4200

ttttttttcta tgtatgtctc aattctcttt caaaatttta cagaatgtta 4250

tcatactaca tatatacttt ttatgtaagc tttttcactt agtattttat 4300

caaatatgtt tttattatat tcatagcctt cttaaacatt atatcaataa 4350

ttgcataata ggcaacctct agcgattacc ataatttttgc tcattgaagg 4400

ctatctccag ttgatcattg ggatgagcat ctttgtgcat gaatcctatt 4450

gctgtatttg ggaaaatttt ccaaggttag attccaataa atatctattt 4500

attattaaat attaaaatat cgatttatta ttaaaaccat ttataaggct 4550

ttttcataaa tgtatagcaa ataggaatta ttaacttgag cataagatat 4600

gagatacatg aacctgaact attaaaataa aatattatat ttaaccctag 4650

tttaagaaga agtcaatatg cttatttaaa tattatggat ggtgggcaga 4700

1353
1352

tcacttgagg tcaggagttc gagaccagcc tggccaacat ggcaaaacca 4750

catctctact aaaaataaaa aaattagctg ggtgtggtgg tgcactcctg 4800

taatcccagc tactcagaag gctgaggtac aagaattgct ggaacctggg 4850

aggcggaggt tgcagtgaac caagattgca ccactgcact ccagccgggg 4900

tgacagagtg agactccgac tgaaaataaa taaataaata aataaataaa 4950

taaataaata aatattatgg atggtgaagg gaatggtata gaattggaga 5000

gattatctta ctgaacacct gtagtcccag ctttctctgg aagtggtggt 5050

atttgagcag gatgtgcaca aggcaattga aatgcccata attagtttct 5100

cagctttgaa tacactataa actcagtggc tgaaggagga aattttagaa 5150

ggaagctact aaaagatcta atttgaaaaa ctacaaaagc attaactaaa 5200

aaagtttatt ttccttttgt ctgggcagta gtgaaaataa ctactcacaa 5250

cattcactat gtttgcaagg aattaacaca aataaaagat gccttttttac 5300

ttaaacgcca agacagaaaa cttgcccaat actgagaagc aacttgcatt 5350

agagagggaa ctgttaaatg ttttcaaccc agttcatctg gtggatgttt 5400

ttgcaggtta ctctgagaat tttgcttatg aaaaatcatt atttttagtg 5450

tagttcacaa taatgtattg aacatacttc taatcaaagg tgctatgtcc 5500

ttgtgtatgg tactaaatgt gtcctgtgta cttttgcaca actgagaatc 5550

ctgcggcttg gtttaatgag tgtgttcatg aaataaataa tggaggaatt 5600

gtcaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 5650

aaaaaaaaaa aaaaaaaa 5668

<210> 18
<211> 3740
<212> DNA
<213> Homo sapiens

<400> 18
gaattcgccc ttggctcagc atgaggaaca gaaggaatga cactctggac 50

agcacccgga ccctgtactc cagcgcgtct cggagcacag acttgtctta 100

cagtgaaagc gacttggtga attttattca agcaaatttt aagaaacgag 150

aatgtgtctt ctttaccaaa gattccaagg ccacggagaa tgtgtgcaag 200

tgtggctatg cccagagcca gcacatggaa ggcacccaga tcaaccaaag 250

tgagaaatgg aactacaaga aacacaccaa ggaatttcct accgacgcct 300

ttggggatat tcagtttgag acactgggga agaaagggaa gtatatacgt 350

ctgtcctgcg acacggacgc ggaaatcctt tacgagctgc tgacccagca 400

ctggcacctg aaaacaccca acctggtcat ttctgtgacc gggggcgcca 450

agaacttcgc cctgaagccg cgcatgcgca agatcttcag ccggctcatc 500

tacatcgcgc agtccaaagg tgcttggatt ctcacgggag gcacccatta 550

tggcctgatg aagtacatcg gggaggtggt gagagataac accatcagca 600

ggagttcaga ggagaatatt gtggccattg gcatagcagc ttggggcatg 650

gtctccaacc gggacaccct catcaggaat tgcgatgctg agggctattt 700

tttagcccag taccttatgg atgacttcac aagagatcca ctgtatatcc 750

1353

tggacaacaa ccacacacat ttgctgctcg tggacaatgg ctgtcatgga 800

catcccactg tcgaagcaaa gctccggaat cagctagaga agtatatctc 850

tgagcgcact attcaagatt ccaactatgg tggcaagatc cccattgtgt 900

gttttgccca aggaggtgga aaagagactt tgaaagccat caatacctcc 950

atcaaaaata aaattccttg tgtggtggtg gaaggctcgg gccagatcgc 1000

tgatgtgatc gctagcctgg tggaggtgga ggatgccctg acatcttctg 1050

ccgtcaagga gaagctggtg cgctttttac cccgcacggt gtcccggctg 1100

cctgaggagg agactgagag ttggatcaaa tggctcaaag aaattctcga 1150

atgttctcac ctattaacag ttattaaaat ggaagaagct ggggatgaaa 1200

ttgtgagcaa tgccatctcc tacgctctat acaaagcctt cagcaccagt 1250

gagcaagaca aggataactg gaatgggcag ctgaagcttc tgctggagtg 1300

gaaccagctg gacttagcca atgatgagat tttcaccaat gaccgccgat 1350

gggagtctgc tgaccttcaa gaagtcatgt ttacggctct cataaaggac 1400

agacccaagt ttgtccgcct ctttctggag aatggcttga acctacggaa 1450

gtttctcacc catgatgtcc tcactgaact cttctccaac cacttcagca 1500

cgcttgtgta ccggaatctg cagatcgcca agaattccta taatgatgcc 1550

ctcctcacgt ttgtctggaa actggttgcg aacttccgaa gaggcttccg 1600

gaaggaagac agaaatggcc gggacgagat ggacatagaa ctccacgacg 1650

tgtctcctat tactcggcac cccctgcaag ctctcttcat ctgggccatt 1700

cttcagaata agaaggaact ctccaaagtc atttgggagc agaccagggg 1750

ctgcactctg gcagccctgg gagccagcaa gcttctgaag actctggcca 1800

aagtgaagaa cgacatcaat gctgctgggg agtccgagga gctggctaat 1850

gagtacgaga cccgggctgt tgagctgttc actgagtgtt acagcagcga 1900

tgaagacttg gcagaacagc tgctggtcta ttcctgtgaa gcttggggtg 1950

gaagcaactg tctggagctg gcggtggagg ccacagacca gcatttcatc 2000

gcccagcctg gggtccagaa ttttctttct aagcaatggt atggagagat 2050

ttcccgagac accaagaact ggaagattat cctgtgtctg tttattatac 2100

ccttggtggg ctgtggcttt gtatcattta ggaagaaacc tgtcgacaag 2150

cacaagaagc tgctttggta ctatgtggcg ttcttcacct ccccccttcgt 2200

ggtcttctcc tggaatgtgg tcttctacat cgccttcctc ctgctgtttg 2250

cctacgtgct gctcatggat ttccattcgg tgccacaccc ccccgagctg 2300

gtcctgtact cgctggtctt tgtcctcttc tgtgatgaag tgagacagtg 2350

gtacgtaaat ggggtgaatt attttactga cctgtggaat gtgatggaca 2400

cgctggggct tttttacttc atagcaggaa ttgtatttcg gctccactct 2450

tctaataaaa gctctttgta ttctggacga gtcattttct gtctggacta 2500

cattattttc actctaagat tgatccacat ttttactgta agcagaaact 2550

taggacccaa gattataatg ctgcagagga tgctgatcga tgtgttcttc 2600

ttcctgttcc tctttgcggt gtggatggtg gcctttggcg tggccaggca 2650

agggatcctt aggcagaatg agcagcgctg gaggtggata ttccgttcgg 2700

tcatctacga gccctacctg gccatgttcg gccaggtgcc cagtgacgtg 2750

gatggtacca cgtatgactt tgcccactgc accttcactg ggaatgagtc 2800

caagccactg tgtgtggagc tggatgagca caacctgccc cggttccccg 2850

agtggatcac catccccctg gtgtgcatct acatgttatc caccaacatc 2900

ctgctggtca acctgctggt cgccatgttt ggctacacgg tgggcaccgt 2950

ccaggagaac aatgaccagg tctggaagtt ccagaggtac ttcctggtgc 3000

aggagtactg cagccgcctc aatatccccc tccccttcat tgtcttcgct 3050

tacttctaca tggtggtgaa gaagtgcttc aagtgttgct gcaaggagaa 3100

aaacatggag tcttctgtct gctgtttcaa aaatgaagac aatgagactc 3150

tggcatggga gggtgtcatg aaggaaaact accttgtcaa gatcaacaca 3200

aaagccaacg acacctcaga ggaaatgagg catcgattta gacaactgga 3250

tacaaagctt aatgatctca agggtcttct gaaagagatt gctaataaaa 3300

tcaaataaaa ctgtatgaac tctaatggag aaaaatctaa ttatagcaag 3350

atcatattaa ggaatgctga tgaacaattt tgctatcgac tactaaatga 3400

gagattttca gacccctggg tacatggtgg atgattttaa atcaccctag 3450

tgtgctgaga ccttgagaat aaagtgtgtg attggtttca tacttgaaga 3500

cggatataaa ggaagaatat ttcctttatg tgtttctcca gaatggtgcc 3550

tgtttctctc tgtgtctcaa tgcctgggac tggaggttga tagtttaagt 3600

gtgttcttac cgcctccttt ttcctttaat cttatttttg atgaacacat 3650

atataggaga acatctatcc tatgaataag aacctggtca tgctttaaaa 3700

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3740

<210> 19
<211> 3809
<212> DNA
<213> Homo sapiens

<400> 19
gcggcggggg cgaggtgagg tgttggcagt ggaaaggggt tcgggctcgg 50

ggggcggggg gacgcggagc ggatggcccg cgccggccgc aggggcggat 100

aaaaagccgt cgcgctgcgg gagtgggcgg gagggagagg gggtgtccga 150

gggccacaag agtatgacgg ggctgtacga gctggtgtgg cgggtgctgc 200

acgcgctgct ctgtctgcac cgcacgctca cctcctggct ccgcgttcgg 250

ttcggcacct ggaactggat ctggcggcgc tgctgccgcg ccgcctctgc 300

cgcggtccta gcgccgctcg gcttcacgct ccgcaagccc ccggcagtcg 350

gcaggaaccg ccgtcaccac cggcacccgc gcgggggggtc gtgcctggca 400

gccgcacacc accggatgcg ctggcgcgcg gacggtcgtt ccttggagaa 450

gctgcctgtg catatgggcc tggtgatcac cgaggtggag caggaaccca 500

gcttctcgga catcgcgagc ctcgtggtgt ggtgtatggc cgtgggcatc 550

tcctacatta gcgtctacga ccaccaaggt attttcaaaa gaaataattc 600

cagattgatg gatgaaattt taaaacaaca gcaagaactt ctgggcctag 650

attgttcaaa atactcacca gaatttgcaa atagtaatga caaagatgat 700

caagttttaa attgccattt ggcagtgaag gtgctgtctc cggaagatgg 750

aaaagcagat attgtaagag ctgctcagga cttttgccag ttagtagccc 800

agaagcaaag gagacccaca gatttggatg tagatacgtt agccagttta 850

cttagttcaa atggttgtcc tgatcctgat ttagtattga agttcggtcc 900

tgtggacagc acattaggct ttcttccctg gcacatcaga ttgactgaga 950

ttgtctcttt gccttcccac ctaaacatca gttatgagga ctttttctct 1000

gcccttcgtc aatatgcagc ctgtgaacag cgtctgggaa agtagtggtc 1050

attggttgca taatttgatt tgaggcttgt ggaggaaagg aaccaagtga 1100

ctctgatgtt tacaaagcac ctatgaaacc ctgtacacac ctatgaaacc 1150

ctgtacacac ctagttcata atcctcataa tttatcaaca aacacaaaga 1200

agtgtcttac ttgagagtga gtgtgtgcgt gtgtgcgtgc acacatgtgc 1250

acgtttgtat gtgtggaaat aaacataaat ggggacgtgt tggagaagga 1300

aatacataga cctacaactt tgagcatata gcagtgatgt tttaggaact 1350

gaaatgtcac acttaataaa gtcttcagcc cagctacttc cctgttttcg 1400

tggggagaag ggggcctgat tagaactgtt ctggttgtgt ttggcgggag 1450

gggaataaat tttgttcagt ccttcttagt gaccaaactt taatttttag 1500

gaataacata ttgacttact gaactgaagc attctgagtt gaaaggagcc 1550

ccagaggaaa ggagttctgt gttgctcaca tgttaaaagc ttgctcacct 1600

tcagagcaga gggaatacct atcttcagat atccgcccat tttcatctct 1650

tcattatagt caaacagtgt gacttgagag tgttgctctg gtgtctgtat 1700

tctggcttat gaagattatt tgaaaaagaa ctcttactac attgaaatgc 1750

agacttttaa aaatttaaat attggattag gcagtcaaaa aaccaaacaa 1800

gcataaaagg tcagtaagtt gtaatcttaa aagtaaaggt ggaaaactca 1850

ttataaatgg aagaaaagtt ttgatttcct tttttgtttg atgggcagta 1900

tgccatatta tatccaaagt tggtttaaaa aatacttcca tcaactattt 1950

ttatttaaaa taaacatttg agggaagtta ccaaggcagc ttttttcctc 2000

aaaagtaacc tgttcctctt tggaacagca cattttaggg gcatggttaa 2050

tacctgagat ttttactcag taaacctgat ggttaccgtg tgtaaaatat 2100

ctttaagtag gattgaaggc ctctgtgggg aatataatgt taccaaagcc 2150

tataaaaata aattttacat gttctttttt atgacagaga gcagcactgg 2200

ttctgttatt ttttaaaatg aataactgat ttcttgacag gtgtttagta 2250

tttcttccct cactgctgat tcttggatag aaaccattct ttatatttga 2300

tagactgttt tcagaaaact cttatcaaca agtgtacaat agttatctaa 2350

aactgtacat ttagaatgga gcagtttaat actagatctc agaaatttga 2400

aaaatagcaa agaagactgg atttggaaag catggtctac aattgcttgt 2450

caattctgaa gctatgaaga ataaatgttt cagctttgga ttatgaaacc 2500

ccgtttatga tttttaaaac acacttgaaa taaaaatgat tgaactaaat 2550

1356

tttggtccgg tgacatcatt ttgcactgca tagcccatta tacattgtaa 2600

gtctttttgt tttaatttat tactgagagt tttctgtgaa gctacagcat 2650

atctaatcag agaatttctg atttgtttcc ttatactgtg attatgttat 2700

attaaggcat ttgtagtgca gcttaagact gaatttatgc tttttgtaaa 2750

catgatcggt gtaaatgtct tataaacatt ctggagtatg tataacttta 2800

atgaatgaaa tttaatggac ctgattaaaa taaagggatt taatctttgt 2850

taaagtgaag ttagtggaat aaattaccta ctggaatata gtccaagtca 2900

ctaaaggttt aatatttgca tttttgtgct tttattttct ccttccattc 2950

ataattatat acttgaaagt acatctgtag cctatgactt gagtctcttg 3000

aacttctagg aagaggcaaa ctacaaacta ctaggattct gatttcagat 3050

ataggcattc cagaatcttc tctttacgag ttcacctgct agtataatct 3100

ccacaacttg aatggccttg gttgttctgt aattgctgcc aaaatcatca 3150

caagctgtac gtcatcaagg ctccctttgc actcccaaga agaactgttc 3200

attttaaaca aaagtgtatg tctttatttg tattggaaaa tattgtcttt 3250

aaattgtttc ttgttgacac ttcccacaat ggaaaaatta ccaaattaaa 3300

cctgttttat ggatggcagc ttggagcata gcaagaattt ggaggatttg 3350

aattccattc ccagttctca ttgcgttttg tttcttaaaa ctataataat 3400

tggttactgt tataaagttt aaaaggtggt tttaatatga atagcaaatt 3450

ctggtatatt gtgactaacg cttaagaatg cctgtctttg agaggaaggt 3500

gttataatat taatcaacag tgccaaatac actgtgcata tctacaattt 3550

aatctttgaa tgtttgttac tggattagct ccctcctcct tctgtgtgat 3600

ggtaccatgc atagagtcaa tcaaatcctt gtgatgtttt gtatggactt 3650

tgacaacatg taactaatgt gtaaagcaag tttttatgat taaggaatca 3700

aatttattga attttattat tgaaagttga aacttaacat gtataaacaa 3750

aaaacaataa aataataaac tattttcatt gactataaaa aaaagcccac 3800

gcgtccgcc 3809

<210> 20
<211> 1048
<212> DNA
<213> Homo sapiens

<400> 20
gaggtgaggt gttggcagtg gaaaggggtt cgggctcggg gggcgggggg 50

acgcggagcg atggcccgcg ccggccgcag gggcggataa aaagccgtcg 100

cgctgcggga gtgggcggga gggagagggg gtgtccgagg gccacaagag 150

tatgacgggg ctgtacgagc tggtgtggcg ggtgctgcac gcgctgctct 200

gtctgcaccg cacgctcacc tcctggctcc gcgttcggtt cggcacctgg 250

aactggatct ggcggcgctg ctgccgcgcc gcctctgccg cggtcctagc 300

gccgctcggc ttcacgctcc gcaagccccc tggcagtcggc aggaaccgcc 350

gtcaccaccg gcacccgcgc ggggggtcgt gcctggcagc cgcacaccac 400

cggatgcgct ggcgcgcgga cggtcgttcc ttggagaagc tgcctgtgca 450

tatgggcctg gtgatcaccg aggtggagca ggaacccagc ttctcggaca 500

tcgcgagcct cgtggtgtgg tgtatggccg tgggcatctc ctacattagc 550

gtctacgacc accaaggtat tttcaaaaga aataattcca gattgatgga 600

tgaaatttta aaacaacagc aagaacttct gggcctagat tgttcaaaat 650

actcaccaga atttgcaaat agtaatgaca aagatgatca agttttaaat 700

tgccatttgg cagtgaaggt gctgtctccg gaagatggaa aagcagatat 750

tgtaagagct gctcaggact tttgccagtt agtagcccag aagcaaaaga 800

gacccacaga tttggatgta gatacgttag ccagtttact tagttcaaat 850

ggttgtcctg atcctgattt agtattgaag ttcggtcctg tggacagcac 900

attaggcttt cttccctggc acatcagatt gactgagatt gtctctttgc 950

cttcccacct aaacatcagt tatgaggact ttttctctgc ccttcgtcaa 1000

tatgcagcct gtgaacagcg tctgggaaag tagtggtcat tggttgca 1048

<210> 21
<211> 3332
<212> DNA
<213> Homo sapiens

<400> 21
gggggggcgga accagcctgc acgcgctggc tccgggtgac agccgcgcgc 50

ctcggccagg atctgagtga tgagacgtgt ccccactgag gtgccccaca 100

gcagcaggtg ttgagcatgg gctgagaagc tggaccggca ccaaagggct 150

ggcagaaatg ggcgcctggc tgattcctag gcagttggcg gcagcaagga 200

ggagaggccg cagcttctgg agcagagccg agacgaagca gttctggagt 250

gcctgaacgg ccccctgagc cctacccgcc tggcccacta tggtccagag 300

gctgtgggtg agccgcctgc tgcggcaccg gaaagcccag ctcttgctgg 350

tcaacctgct aacctttggc ctggaggtgt gtttggccgc aggcatcacc 400

tatgtgccgc ctctgctgct ggaagtgggg gtagaggaga agttcatgac 450

catggtgctg ggcattggtc cagtgctggg cctggtctgt gtcccgctcc 500

taggctcagc cagtgaccac tggcgtggac gctatggccg ccgccggccc 550

ttcatctggg cactgtcctt gggcatcctg ctgagcctct ttctcatccc 600

aagggccggc tggctagcag ggctgctgtg cccggatccc aggcccctgg 650

agctggcact gctcatcctg ggcgtggggc tgctggactt ctgtggccag 700

gtgtgcttca ctccactgga ggccctgctc tctgacctct tccgggaccc 750

ggaccactgt cgccaggcct actctgtcta tgccttcatg atcagtcttg 800

ggggctgcct gggctacctc ctgcctgcca ttgactggga caccagtgcc 850

ctggcccccc acctgggcac ccaggaggag tgcctctttg gcctgctcac 900

cctcatcttc ctcacctgcg tagcagccac actgctggtg gctgaggagg 950

cagcgctggg ccccaccgag ccagcagaag ggctgtcggc cccctccttg 1000

tcgccccact gctgtccatg ccgggcccgc ttggctttcc ggaacctggg 1050

cgccctgctt ccccggctgc accagctgtg ctgccgcatg ccccgcaccc 1100

tgcgccggct cttcgtggct gagctgtgca gctggatggc actcatgacc 1150

ttcacgctgt tttacacgga tttcgtgggc gaggggctgt accagggcgt 1200

gcccagagct gagccgggca ccgaggcccg gagacactat gatgaaggcg 1250

ttcggatggg cagcctgggg ctgttcctgc agtgcgccat ctccctggtc 1300

ttctctctgg tcatggaccg gctggtgcag cgattcggca ctcgagcagt 1350

ctatttggcc agtgtggcag ctttccctgt ggctgccggt gccacatgcc 1400

tgtcccacag tgtggccgtg gtgacagctt cagccgccct caccgggttc 1450

accttctcag ccctgcagat cctgccctac acactggcct ccctctacca 1500

ccgggagaag caggtgttcc tgcccaaata ccgaggggac actggaggtg 1550

ctagcagtga ggacagcctg atgaccagct tcctgccagg ccctaagcct 1600

ggagctccct tccctaatgg acacgtgggt gctggaggca gtggcctgct 1650

cccacctcca cccgcgctct gcggggcctc tgcctgtgat gtctccgtac 1700

gtgtggtggt gggtgagccc accgaggcca gggtggttcc gggccggggc 1750

atctgcctgg acctcgccat cctggatagt gccttcctgc tgtcccaggt 1800

ggccccatcc ctgtttatgg gctccattgt ccagctcagc cagtctgtca 1850

ctgcctatat ggtgtctgcc gcaggcctgg gtctggtcgc catttacttt 1900

gctacacagg tagtatttga caagagcgac ttggccaaat actcagcgta 1950

gaaaacttcc agcacattgg ggtggagggc ctgcctcact gggtcccagc 2000

tccccgctcc tgttagcccc atggggctgc cgggctggcc gccagtttct 2050

gttgctgcca aagtaatgtg gctctctgct gccaccctgt gctgctgagg 2100

tgcgtagctg cacagctggg ggctggggcg tccctctcct ctctccccag 2150

tctctagggc tgcctgactg gaggccttcc aaggggttt cagtctggac 2200

ttatacaggg aggccagaag ggctccatgc actggaatgt ggggactctg 2250

caggtggatt acccaggctc agggttaaca gctagcctcc tagttgagac 2300

acacctagag aagggttttt gggagctgaa taaactcagt cacctggttt 2350

cccatctcta agccccttaa cctgcagctt cgtttaatgt agctcttgca 2400

tgggagtttc taggatgaaa cactcctcca tgggatttga acatatgaaa 2450

gttatttgta ggggaagagt cctgaggggc aacacacaag aaccaggtcc 2500

cctcagccca cagcactgtc tttttgctga tccacccccc tcttaccttt 2550

tatcaggatg tggcctgttg gtccttctgt tgccatcaca gagacacagg 2600

catttaaata tttaacttat ttatttaaca aagtagaagg gaatccattg 2650

ctagcttttc tgtgttggtg tctaatattt gggtagggtg ggggatcccc 2700

aacaatcagg tcccctgaga tagctggtca ttgggctgat cattgccaga 2750

atcttcttct cctggggtct ggcccccccaa aatgcctaac ccaggacctt 2800

ggaaattcta ctcatcccaa atgataattc caaatgctgt tacccaaggt 2850

tagggtgttg aaggaaggta gagggtgggg cttcaggtct caacggcttc 2900

cctaaccacc cctcttctct tggcccagcc tggttccccc cacttccact 2950

cccctctact ctctctagga ctgggctgat gaaggcactg cccaaaattt 3000

cccctacccc caactttccc ctacccccaa ctttccccac cagctccaca 3050

1359

accctgtttg gagctactgc aggaccagaa gcacaaagtg cggtttccca 3100

agcctttgtc catctcagcc cccagagtat atctgtgctt ggggaatctc 3150

acacagaaac tcaggagcac cccctgcctg agctaaggga ggtcttatct 3200

ctcaggggg tttaagtgcc gtttgcaata atgtcgtctt atttatttag 3250

cggggtgaat attttatact gtaagtgagc aatcagagta taatgtttat 3300

ggtgacaaaa ttaaagcctt tcttatatgt tt 3332

<210> 22
<211> 3768
<212> DNA
<213> Homo sapiens

<400> 22
aacaaaagct ggagctccac cgcggtggcg gccgctctag cccctggttc 50

ggcccacctc tgaaggttcc agaatcgata gtgaattcgt ggggaagcgc 100

agccatggct ctgcggaggc tgggggccgc gctgctgctg ctgccgctgc 150

tcgccgccgt ggaagaaacg ctaatggact ccactacagc gactgctgag 200

ctgggctgga tggtgcatcc tccatcaggg tgggaagagg tgagtggcta 250

cgatgagaac atgaacacga tccgcacgta ccaggtgtgc aacgtgtttg 300

agtcaagcca gaacaactgg ctacggacca agtttatccg gcgccgtggc 350

gcccaccgca tccacgtgga gatgaagttt tcggtgcgtg actgcagcag 400

catccccagc gtgcctggct cctgcaagga gaccttcaac ctctattact 450

atgaggctga ctttgactcg gccaccaaga ccttccccaa ctggatggag 500

aatccatggg tgaaggtgga taccattgca gccgacgaga gcttctccca 550

ggtggacctg ggtggccgcg tcatgaaaat caacaccgag gtgcggagct 600

tcggacctgt gtcccgcagc ggcttctacc tggccttcca ggactatggc 650

ggctgcatgt ccctcatcgc cgtgcgtgtc ttctaccgca agtgcccccg 700

catcatccag aatggcgcca tcttccagga aaccctgtcg ggggctgaga 750

gcacatcgct ggtggctgcc cggggcagct gcatcgccaa tgcggaagag 800

gtggatgtac ccatcaagct ctactgtaac ggggacggcg agtggctggt 850

gcccatcggg cgctgcatgt gcaaagcagg cttcgaggcc gttgagaatg 900

gcaccgtctg ccgaggttgt ccatctggga ctttcaaggc caaccaaggg 950

gatgaggcct gtacccactg tcccatcaac agccggacca cttctgaagg 1000

ggccaccaac tgtgtctgcc gcaatggcta ctacagagca gacctggacc 1050

ccctggacat gccctgcaca accatccccc ccgcgcccca ggctgtgatt 1100

tccagtgtca atgagacctc cctcatgctg gagtggaccc ctccccgcga 1150

ctccggaggc cgagaggacc tcgtctacaa catcatctgc aagagctgtg 1200

gctcgggccg gggtgcctgc acccgctgcg gggacaatgt acagtacgca 1250

ccacgccagc taggcctgac cgagccacgc atttacatca gtgacctgct 1300

ggcccacacc cagtacacct tcgagatcca ggctgtgaac ggcgttactg 1350

accagagccc cttctcgcct cagttcgcct ctgtgaacat caccaccaac 1400

caggcagctc catcggcagt gtccatcatg catcaggtga gccgcaccgt 1450

ggacagcatt accctgtcgt ggtcccagcc agaccagccc aatggcgtga 1500

tcctggacta tgagctgcag tactatgaga aggagctcag tgagtacaac 1550

gccacagcca taaaaagccc caccaacacg gtcaccgtgc agggcctcaa 1600

agccggcgcc atctatgtct tccaggtgcg ggcacgcacc gtggcaggct 1650

acgggcgcta cagcggcaag atgtacttcc agaccatgac agaagccgag 1700

taccagacaa gcatccagga gaagttgcca ctcatcatcg gctcctcggc 1750

cgctggcctg gtcttcctca ttgctgtggt tgtcatcgcc atcgtgtgta 1800

acagaagacg ggggtttgag cgtgctgact cggagtacac ggacaagctg 1850

caacactaca ccagtggcca catgacccca ggcatgaaga tctacatcga 1900

tcctttcacc tacgaggacc ccaacgaggc agtgcgggag tttgccaagg 1950

aaattgacat ctcctgtgtc aaaattgagc aggtgatcgg agcaggggag 2000

tttggcgagg tctgcagtgg ccacctgaag ctgccaggca agagagagat 2050

ctttgtggcc atcaagacgc tcaagtcggg ctacacggag aagcagcgcc 2100

gggacttcct gagcgaagcc tccatcatgg gccagttcga ccatcccaac 2150

gtcatccacc tggagggtgt cgtgaccaag agcacacctg tgatgatcat 2200

caccgagttc atggagaatg gctccctgga ctcctttctc cggcaaaacg 2250

atgggcagtt cacagtcatc cagctggtgg gcatgcttcg gggcatcgca 2300

gctggcatga agtacctggc agacatgaac tatgttcacc gtgacctggc 2350

tgcccgcaac atcctcgtca acagcaacct ggtctgcaag gtgtcggact 2400

ttgggctctc acgctttcta gaggacgata cctcagaccc cacctacacc 2450

agtgccctgg gcggaaagat ccccatccgc tggacagccc cggaagccat 2500

ccagtaccgg aagttcacct cggccagtga tgtgtggagc tacggcattg 2550

tcatgtggga ggtgatgtcc tatggggagc ggccctactg ggacatgacc 2600

aaccaggatg taatcaatgc cattgagcag gactatcggc tgccaccgcc 2650

catggactgc ccgagcgccc tgcaccaact catgctggac tgttggcaga 2700

aggaccgcaa ccaccggccc aagttcggcc aaattgtcaa cacgctagac 2750

aagatgatcc gcaatcccaa cagcctcaaa gccatggcgc ccctctcctc 2800

tggcatcaac ctgccgctgc tggaccgcac gatccccgac tacaccagct 2850

ttaacacggt ggacgagtgg ctggaggcca tcaagatggg gcagtacaag 2900

gagagcttcg ccaatgccgg cttcacctcc tttgacgtcg tgtctcagat 2950

gatgatggag gacattctcc gggttgggct cactttggct ggccaccaga 3000

aaaaaatcct gaacagtatc caggtgatgc gggcgcagat gaaccagatt 3050

cagtctgtgg aggtttgaca ttcacctgcc tcggctcacc tcttcctcca 3100

agccccgccc cctctgcccc acgtgccggc cctcctggtg ctctatccac 3150

tgcagggcca gccactcgcc aggaggccac gggccacggg aagaaccaag 3200

cggtgccagc cacgagacgt caccaagaaa acatgcaact caaacgacgg 3250

aaaaaaaaag ggaatgggaa aaaagaaaac agatcctggg agggggcggg 3300

aaatacaagg aatatttttt aaagaggatt ctcataagga aagcaatgac 3350

tgttcttgcg ggggataaaa aagggcttgg gagattcatg cgatgtgtcc 3400

aatcggagac aaaagcagtt tctctccaac tccctctggg aaggtgacct 3450

ggccagagcc aagaaacact ttcagaaaaa caaatgtgaa ggggagagac 3500

aggggccgcc cttggctcct gtccctgctg ctcctctagg cctcactcaa 3550

caaccaagcg cctggaggac gggacagatg gacagacagc caccctgaga 3600

acccctctgg gaaaatctat tcctgccacc actgggcaaa cagaagaatt 3650

tttctgtctt tggagagtat tttagaaact ccaatgaaag acactgtttc 3700

tcctgttggc tcacagggct gaaaggggct tttgtcctcc tgggtcaggg 3750

agaacgcggg gacccccag 3768

<210> 23
<211> 3800
<212> DNA
<213> Homo sapiens

<400> 23
cgggaagcgc agccatggct ctgcggaggc tgggggccgc gctgctgctg 50

ctgccgctgc tcgccgccgt ggaagaaacg ctaatggact ccactacagc 100

gactgctgag ctgggctgga tggtgcatcc tccatcaggg tgggaagagg 150

tgagtggcta cgatgagaac atgaacacga tccgcacgta ccaggtgtgc 200

aacgtgtttg agtcaagcca gaacaactgg ctacggacca agtttatccg 250

gcgccgtggc gcccaccgca tccacgtgga gatgaagttt tcggtgcgtg 300

actgcagcag catccccagc gtgcctggct cctgcaagga gaccttcaac 350

ctctattact atgaggctga ctttgactcg gccaccaaga ccttccccaa 400

ctggatggag aatccatggg tgaaggtgga taccattgca gccgacgaga 450

gcttctccca ggtggacctg ggtggccgcg tcatgaaaat caacaccgag 500

gtgcggagct tcggacctgt gtcccgcagc ggcttctacc tggccttcca 550

ggactatggc ggctgcatgt ccctcatcgc cgtgcgtgtc ttctaccgca 600

agtgcccccg catcatccag aatggcgcca tcttccagga aaccctgtcg 650

ggggctgaga gcacatcgct ggtggctgcc cggggcagct gcatcgccaa 700

tgcggaagag gtggatgtac ccatcaagct ctactgtaac ggggacggcg 750

agtggctggt gcccatcggg cgctgcatgt gcaaagcagg cttcgaggcc 800

gttgagaatg gcaccgtctg ccgaggttgt ccatctggga ctttcaaggc 850

caaccaaggg gatgaggcct gtacccactg tcccatcaac agccggacca 900

cttctgaagg ggccaccaac tgtgtctgcc gcaatggcta ctacagagca 950

gacctggacc ccctggacat gccctgcaca accatccct ccgcgcccca 1000

ggctgtgatt tccagtgtca atgagacctc cctcatgctg gagtggaccc 1050

ctccccgcga ctccggaggc cgagaggacc tcgtctacaa catcatctgc 1100

aagagctgtg gctcgggccg gggtgcctgc acccgctgcg gggacaatgt 1150

acagtacgca ccacgccagc taggcctgac cgagccacgc atttacatca 1200

gtgacctgct ggcccacacc cagtacacct tcgagatcca ggctgtgaac 1250

ggcgttactg accagagccc cttctcgcct cagttcgcct ctgtgaacat 1300

caccaccaac caggcagctc catcggcagt gtccatcatg catcaggtga 1350

gccgcaccgt ggacagcatt accctgtcgt ggtcccagcc ggaccagccc 1400

aatggcgtga tcctggacta tgagctgcag tactatgaga aggagctcag 1450

tgagtacaac gccacagcca taaaaagccc caccaacacg gtcaccgtgc 1500

agggcctcaa agccggcgcc atctatgtct tccaggtgcg ggcacgcacc 1550

gtggcaggct acgggcgcta cagcggcaag atgtacttcc agaccatgac 1600

agaagccgag taccagacaa gcatccagga gaagttgcca ctcatcatcg 1650

gctcctcggc cgctggcctg gtcttcctca ttgctgtggt tgtcatcgcc 1700

atcgtgtgta acagacgggg gtttgagcgt gctgactcgg agtacacgga 1750

caagctgcaa cactacacca gtggccacat gaccccaggc atgaagatct 1800

acatcgatcc tttcacctac gaggacccca acgaggcagt gcgggagttt 1850

gccaaggaaa ttgacatctc ctgtgtcaaa attgagcagg tgatcggagc 1900

aggggagttt ggcgaggtct gcagtggcca cctgaagctg ccaggcaaga 1950

gagagatctt tgtggccatc aagacgctca agtcgggcta cacggagaag 2000

cagcgccggg acttcctgag cgaagcctcc atcatgggcc agttcgacca 2050

tcccaacgtc atccacctgg agggtgtcgt gaccaagagc acacctgtga 2100

tgatcatcac cgagttcatg gagaatggct ccctggactc ctttctccgg 2150

caaaacgatg ggcagttcac agtcatccag ctggtgggca tgcttcgggg 2200

catcgcagct ggcatgaagt acctggcaga catgaactat gttcaccgtg 2250

acctggctgc ccgcaacatc ctcgtcaaca gcaacctggt ctgcaaggtg 2300

tcggactttg ggctctcacg ctttctagag gacgatacct cagaccccac 2350

ctacaccagt gccctgggcg gaaagatccc catccgctgg acagccccgg 2400

aagccatcca gtaccggaag ttcacctcgg ccagtgatgt gtggagctac 2450

ggcattgtca tgtgggaggt gatgtcctat ggggagcggc cctactggga 2500

catgaccaac caggatgtaa tcaatgccat tgagcaggac tatcggctgc 2550

caccgcccat ggactgcccg agcgccctgc accaactcat gctggactgt 2600

tggcagaagg accgcaacca ccggcccaag ttcggccaaa ttgtcaacac 2650

gctagacaag atgatccgca atcccaacag cctcaaagcc atggcgcccc 2700

tctcctctgg catcaacctg ccgctgctgg accgcacgat ccccgactac 2750

accagcttta acacggtgga cgagtggctg gaggccatca agatggggca 2800

gtacaaggag agcttcgcca atgccggctt cacctccttt gacgtcgtgt 2850

ctcagatgat gatggaggac attctccggg ttggggtcac tttggctggc 2900

caccagaaaa aaatcctgaa cagtatccag gtgatgcggg cgcagatgaa 2950

ccagattcag tctgtggagg tttgacattc acctgcctcg gctcacctct 3000

tcctccaagc cccgcccccc ctgccccacg tgccggccct cctggtgctc 3050

tatccactgc agggccagcc actcgccagg aggccacggg ccacgggaag 3100

aaccaagcgg tgccagccac gagacgtcac caagaaaaca tgcaactcaa 3150

acgacggaaa aaaaaaggga atgggaaaaa agaaaacaga tcctgggagg 3200

gggcgggaaa tacaaggaat atttttaaa gaggattctc ataaggaaag 3250

caatgactgt tcttgcgggg gataaaaaag ggcttgggag attcatgcga 3300

tgtgtccaat cggagacaaa agcagtttct ctccaactcc ctctgggaag 3350

gtgacctggc cagagccaag aaacactttc agaaaaacaa atgtgaaggg 3400

gagagacagg ggccgccctt ggctcctgtc cctgctgctc ctctaggcct 3450

cactcaacaa ccaagcgcct ggaggacggg acagatggac agacagccac 3500

cctgagaacc cctctgggaa aatctattcc tgccaccact gggcaaacag 3550

aagaattttt ctgtctttgg agagtatttt agaaactcca atgaaagaca 3600

ctgtttctcc tgttggctca cagggctgaa aggggctttt gtcctcctgg 3650

gtcagggaga acgcggggac cccagaaagg tcagccttcc tgaggatggg 3700

caacccccag gtctgcagct ccaggtacat atcacgcgca cagcctggca 3750

gcctggccct cctggtgccc actcccgcca gcccctgcct cgaggtcgac 3800

<210> 24
<211> 853
<212> DNA
<213> Homo sapiens

<400> 24
ccgccagctc cttcggcaat gaacttctcc accagcagca gcagcttcgc 50

ctacgaccgg gagttcctcc gcaccctgcc cggcttcctc atcgtggccg 100

agatcgttct ggggctgctg gtatggacgc ttattgctgg aactgagtac 150

ttccgggtcc ccgcatttgg ctgggtcatg tttgtagctg tattttactg 200

ggtcctcacc gtcttcttcc tcattatcta cataacaatg acctacacca 250

ggattcccca ggtgccctgg acaacagtgg gcctgtgctt taacggcagt 300

gccttcgtct tgtacctctc tgccgctgtt gtagatgcat cttccgtctc 350

ccctgagagg gacagtcaca acttcaacag ctgggcggcc tcatcgttct 400

ttgccttcct ggtcaacatc tgctacgctg gaaatacata tttcagtttt 450

atagcatgga gatccaggac catacagtga tttaccattt tgataattaa 500

aaggaaaaaa aaaggaagac tctcactgta aaaacagctg taggtataat 550

gtatattccc agagaattgt atttaactaa ttaatgtttt ttatattctt 600

aaatttgctc acaaattgtg gtttgttaca attaaactgg atacttattt 650

gcaaagtgtt gtagcttata atgaactctt aagtatctta ttaatgtatt 700

aatgtcttca tagatcatat tttcttagac aatgtttaaa tagataaatt 750

gctaatattg agaatgtgtc aagtttgtaa acctaacttt taagatgcca 800

gattcttttt tgattaaatg ttgcaaaatc ccaaaaaaaa aaaaaaaaaa 850

aaa 853

<210> 25
<211> 485
<212> DNA
<213> Homo sapiens

<400> 25
gcccttggca atgaacttct ccaccagcag cagcagcttc gcctacgacc 50

gggagttcct ccgcaccctg cccggcttcc tcatcgtggc cgagatcgtt 100

ctggggctgc tggtatggac gcttattgct ggaactgagt acttccgggt 150

ccccgcattt ggctgggtca tgtttgtagc tgtattttac tgggtcctca 200

ccgtcttctt cctcattatc tacataacaa tgacctacac caggattccc 250

caggtgccct ggacaacagt gggcctgtgc tttaacggca gtgccttcgt 300

cttgtacctc tctgccgctg ttgtagatgc atcttccgtc tcccctgaga 350

gggacagtca caacttcaac agctgggcgg cctcatcgtt ctttgccttc 400

ctggtcacca tctgctacgc tggaaataca tatttcagtt ttatagcatg 450

gagatccagg accatacagt gatttaccaa agggc 485

<210> 26
<211> 2257
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 2159, 2231
<223> unknown base

<400> 26
ggccctgccc tgaaggctgg tcacttgcag aggtaaactc ccctctttga 50

cttctggcca gggtttgtgc tgagctggct gcagccgctc tcagcctcgc 100

tccgggcacg tcgggcagcc tcgggccctc ctgcctgcag gatcatgccc 150

accaccgtgg acgatgtcct ggagcatgga ggggagtttc actttttcca 200

gaagcaaatg tttttcctct tggctctgct ctcggctacc ttcgcgccca 250

tctacgtggg catcgtcttc ctgggcttca cccctgacca ccgctgccgg 300

agccccggag tggccgagct gagtctgcgc tgcggctgga gtcctgcaga 350

ggaactgaac tacacggtgc cgggcccagg acctgcgggc gaagcctccc 400

caagacagtg taggcgctac gaggtggact ggaaccagag caccttttgac 450

tgcgtggacc ccctggccag cctggacacc aacaggagcc gcctgccact 500

gggcccctgc cgggacggct gggtgtacga gacgcctggc tcgtccatcg 550

tcaccgagtt taacctggta tgtgccaact cctggatgtt ggacctattc 600

cagtcatcag tgaatgtagg attctttatt ggctctatga gtatcggcta 650

catagcagac aggtttggcc gtaagctctg cctcctaact acagtcctca 700

taaatgctgc agctggagtt ctcatggcca tttccccaac ctatacgtgg 750

atgttaattt ttcgcttaat ccaaggactg gtcagcaaag caggctggtt 800

aataggctac atcctgatta cagaatttgt tgggcggaga tatcggagaa 850

cagtggggat tttttaccaa gttgcctata cagttgggct cctggtgcta 900

gctggggtgg cttacgcact tcctcactgg aggtggttgc agttcacagt 950

tgctctgccc aacttcttct tcttgctcta ttactggtgc atacctgagt 1000

ctcccaggtg gctgatctcc cagaataaga atgctgaagc catgagaatc 1050

attaagcaca tcgcaaagaa aaatggaaaa tctctacccg cctcccttca 1100

gcgcctgaga cttgaagagg aaactggcaa gaaattgaac ccttcatttc 1150

ttgacttggt cagaactcct cagataagga aacatactat gatattgatg 1200

tacaactggt tcacgagctc tgtgctctac cagggcctca tcatgcacat 1250

gggccttgca ggtgacaata tctacctgga tttcttctac tctgccctgg 1300

ttgaattccc agctgccttc atgatcatcc tcaccatcga ccgcatcgga 1350

cgccgttacc cttgggctgc atcaaatatg gttgcagggg cagcctgtct 1400

ggcctcagtt tttatacctg gtgatctaca atggctaaaa attattatct 1450

catgcttggg aagaatgggg atcacaatgg cctatgagat agtctgcctg 1500

gtcaatgctg agctgtaccc cacattcatt aggaatcttg gcgtccacat 1550

ctgttcctca atgtgtgaca ttggtggcat catcacgcca ttcctggtct 1600

accggctcac taacatctgg cttgagctcc cgctgatggt tttcggcgta 1650

cttggcttgg ttgctggagg tctggtgctg ttgcttccag aaactaaagg 1700

gaaagctttg cctgagacca tcgaggaagc cgaaaatatg caaagaccaa 1750

gaaaaaataa agaaaagatg atttacctcc aagttcagaa actagacatt 1800

ccattgaact aagaagagag accgttgctg ctgtcatgac ctagctttga 1850

tggcagcaag accaaaagta gaaatccctg cactcatcac aaagcccata 1900

caactcaacc aaacttaccc ctgagcccta tcaacctagg tctacagcca 1950

gtggagtcta ttgtacactg tggaaaaata cccatgggac cagatcctgc 2000

caaattcttc cagctcactt tattctcagc attcctagga cattggacat 2050

tggttttctg gagggttttt tttccgatct ttgtattttt ttaaatttga 2100

ttcttttctt tgcaatgcta gcaaccagaa tacatagggg aactgtgggc 2150

taggcaaana aaatagaaaa agtgtgaaaa acagtaaagt tgggagagga 2200

gcatctattt tcttaaagaa ataaaacacc naaaacaaaa aaaaaaaaaa 2250

aaaaaaa 2257

<210> 27
<211> 1725
<212> DNA
<213> Homo sapiens

<400> 27
gcccttctcc tgcctgcagg atcatgccca ccaccgtgga cgatgtcctg 50

gagcatggag gggagtttca ctttttccag aagcaaatgt ttttcctctt 100

ggctctgctc tcggctacct tcgcgcccat ctacgtgggc atcgtcttcc 150

tgggcttcac ccctgaccac cgctgccgga gccccggagt ggccgagctg 200

agtctgcgct gcggctggag tcctgcagag gaactgaact acacggtgcc 250

gggcccagga cctgcgggcg aagcctcccc aagacagtgt aggcgctacg 300

aggtggactg gaaccagagc accttcgact gcgtggaccc cctggccagc 350

ctggacacca acaggagccg cctgccactg ggcccctgcc gggacggctg 400

ggtgtacgag acgcctggct cgtccatcgt caccgagttt aacctggtat 450

gtgccaactc ctggatgttg gacctattcc agtcatcagt gaatgtagga 500

ttctttattg gctctatgag tatcggctac atagcagaca ggtttggccg 550

taagctctgc ctcctaacta cagtcctcat aaatgctgca gctggagttc 600

tcatggccat ttccccaacc tatacgtgga tgttaatttt tcgcttaatc 650

caaggactgg tcagcaaagc aggctggtta ataggctaca tcctgattac 700

agaatttgtt gggcggagat atcggagaac agtggggatt ttttaccaag 750

ttgcctatac agttgggctc ctggtgctag ctggggtggc ttacgcactt 800

cctcactgga ggtggttgca gttcacagtt tctctgccca acttcttctt 850

cttgctctat tactggtgca tacctgagtc tcccaggtgg ctgatctccc 900

agaataagaa tgctgaagcc atgagaatca ttaagcacat cgcaaagaaa 950

aatggaaaat ctctacccgc ctcccttcag cgcctgagac ttgaagagga 1000

aactggcaag aaattgaacc cttcatttct tgacttggtc agaactcctc 1050

agataaggaa acatactatg atattgatgt acaactggtt cacgagctct 1100

gtgctctacc agggcctcat catgcacatg ggccttgcag gtgacaatat 1150

ctacctggat ttcttctact ctgccctggt tgaattccca gctgccttca 1200

tgatcatcct caccatcgac cgcatcggac gccgttaccc ttgggctgca 1250

tcaaatatgg ttgcaggggc agcctgtctg gcctcagttt ttatacctgg 1300

tgatctacaa tggctaaaaa ttattatctc atgcttggga agaatgggga 1350

tcacaatggc ctatgagata gtctgcctgg tcaatgctga gctgtacccc 1400

acattcatta ggaatcttgg cgtccacatc tgttcctcaa tgtgtgacat 1450

tggtggcatc atcacgccat tcctggtcta ccggctcact aacatctggc 1500

ttgagctccc gctgatggtt ttcggcgtgc ttggcttggt tgctggaggt 1550

ctggtgctgt tgcttccaga aactaaaggg aaagctttgc ctgagaccat 1600

cgaggaagcc gaaaatatgc aaagaccaag aaaaaataaa gaaaagatga 1650

tttacctcca agttcagaaa ctagacattc cattgaacta agaagagaga 1700

ccgttgctgc tgtcatgaca agggc 1725

<210> 28
<211> 2280
<212> DNA
<213> Homo sapiens

<400> 28
cgggccaggt ttccaggctc ggccgccgcc tccatcccag cacctgcgga 50

gggagcgctg accatggctc cctggcctga attgggagat gcccagccca 100

accccgataa gtacctcgaa ggggccgcag gtcagcagcc cactgcccct 150

gataaaagca aagagaccaa caaaacagat aacactgagg cacctgtaac 200

caagattgaa cttctgccgt cctactccac ggctacactg atagatgagc 250

ccactgaggt ggatgacccc tggaacctac ccactcttca ggactcgggg 300

atcaagtggt cagagagaga caccaaaggg aagattctct gtttcttcca 350

agggattggg agattgattt tacttctcgg atttctctac tttttcgtgt 400

gctccctgga tattcttagt agcgccttcc agctggttgg aggaaaaatg 450

gcaggacagt tcttcagcaa cagctctatt atgtccaacc ctttgttggg 500

gctggtgatc ggggtgctgg tgaccgtctt ggtgcagagc tccagcacct 550

caacgtccat cgttgtcagc atggtgtcct cttcattgct cactgttcgg 600

gctgccatcc ccattatcat gggggccaac attggaacgt caatcaccaa 650

cactattgtt gcgctcatgc aggtgggaga tcggagtgag ttcagaagag 700

cttttgcagg agccactgtc catgacttct tcaactggct gtccgtgttg 750

gtgctcttgc ccgtggaggt ggccacccat tacctcgaga tcataaccca 800

gcttatagtg gagagcttcc acttcaagaa tggagaagat gccccagatc 850

ttctgaaagt catcactaag cccttcacaa agctcattgt ccagctggat 900

aaaaaagtta tcagccaaat tgcaatgaac gatgaaaaag cgaaaaacaa 950

gagtcttgtc aagatttggt gcaaaacttt taccaacaag acccagatta 1000

acgtcactgt tccctcgact gctaactgca cctccccttc cctctgttgg 1050

acggatggca tccaaaactg gaccatgaag aatgtgacct acaaggagaa 1100

catcgccaaa tgccagcata tctttgtgaa tttccacctc ccggatcttg 1150

ctgtgggcac catcttgctc atactctccc tgctggtcct ctgtggttgc 1200

ctgatcatga ttgtcaagat cctgggctct gtgctcaagg ggcaggtcgc 1250

cactgtcatc aagaagacca tcaacactga tttccccttt ccctttgcat 1300

ggttgactgg ctacctggcc atcctcgtcg gggcaggcat gaccttcatc 1350

gtacagagca gctctgtgtt cacgtcggcc ttgacccccc tgattggaat 1400

cggcgtgata accattgaga gggcttatcc actcacgctg ggctccaaca 1450

tcggcaccac caccaccgcc atcctggccg ccttagccag ccctggcaat 1500

gcattgagga gttcactcca gatcgccctg tgccactttt tcttcaacat 1550

ctccggcatc ttgctgtggt acccgatccc gttcactcgc ctgcccatcc 1600

gcatggccaa ggggctgggc aacatctctg ccaagtatcg ctggttcgcc 1650

gtcttctacc tgatcatctt cttcttcctg atcccgctga cggtgtttgg 1700

cctctcgctg gccggctggc gggtgctggt tggtgtcggg gttcccgtcg 1750

tcttcatcat catcctggta ctgtgcctcc gactcctgca gtctcgctgc 1800

ccacgcgtcc tgccgaagaa actccagaac tggaacttcc tgccgctgtg 1850

gatgcgctcg ctgaagccct gggatgccgt cgtctccaag ttcaccggct 1900

gcttccagat gcgctgctgc tactgctgcc gcgtgtgctg ccgcgcgtgc 1950

tgcttgctgt gtggctgccc caagtgctgc cgctgcagca agtgctgcga 2000

ggacttggag gaggcgcagg aggggcagga tgtccctgtc aaggctcctg 2050

agacctttga taacataacc attagcagag aggctcaggg tgaggtccct 2100

gcctcggact caaagaccga atgcacggcc ttgtagggga cgccccagat 2150

tgtcagggat ggggggatgg tccttgagtt ttgcatgctc tcctccctcc 2200

cacttctgca ccctttcacc acctcgagga gatttgctcc ccattagcga 2250

atgaaattga tgcagtccta aaaaaaaaaa 2280

<210> 29
<211> 2366
<212> DNA
<213> Homo sapiens

<400> 29
cagcccagca cctgcggagg gagcgctgac catggctccc tggcctgaat 50

tgggagatgc ccagccccaac cccgataagt acctcgaagg ggccgcaggt 100

cagcagccca ctgcccctga taaaagcaaa gagaccaaca aaacagataa 150

cactgaggca cctgtaacca agattgaact tctgccgtcc tactccacgg 200

ctacactgat agatgagccc actgaggtgg atgacccctg gaacctaccc 250

actcttcagg actcggggat caagtggtca gagagagaca ccaaagggaa 300

gattctctgt ttcttccaag ggattgggag attgatttta cttctcggat 350

ttctctactt tttcgtgtgc tccctggata ttcttagtag cgccttccag 400

ctggttggag gaaaaatggc aggacagttc ttcagcaaca gctctattat 450

gtccaaccct ttgttggggc tggtgatcgg ggtgctggtg accgtcttgg 500

tgcagagctc cagcacctca acgtccatcg ttgtcagcat ggtgtcctct 550

tcattgctca ctgttcgggc tgccatcccc attatcatgg gggccaacat 600

tggaacgtca atcaccaaca ctattgttgc gctcatgcag gtgggagatc 650

ggagtgagtt cagaagagct tttgcaggag ccactgtcca tgacttcttc 700

aactggctgt ccgtgttggt gctcttgccc gtggaggtgg ccacccatta 750

cctcgagatc ataacccagc ttatagtgga gagcttccac ttcaagaatg 800

gagaagatgc cccagatctt ctgaaagtca tcactaagcc cttcacaaag 850

ctcattgtcc agctggataa aaaagttatc agccaaattg caatgaacga 900

tgaaaaagcg aaaaacaaga gtcttgtcaa gatttggtgc aaaactttta 950

ccaacaagac ccagattaac gtcactgttc cctcgactgc taactgcacc 1000

tcccccttccc tctgttggac ggatggcatc caaaactgga ccatgaagaa 1050

tgtgacctac aaggagaaca tcgccaaatg ccagcatatc tttgtgaatt 1100

tccacctccc ggatcttgct gtgggcacca tcttgctcat actctccctg 1150

ctggtcctct gtggttgcct gatcatgatt gtcaagatcc tgggctctgt 1200

gctcaagggg caggtcgcca ctgtcatcaa gaagaccatc aacactgatt 1250

tcccctttcc ctttgcatgg ttgactggct acctggccat cctcgtcggg 1300

gcaggcatga ccttcatcgt acagagcagc tctgtgttca cgtcggcctt 1350

gaccccccctg attggaatcg gcgtgataac cattgagagg gcttatccac 1400

tcacgctggg ctccaacatc ggcaccacca ccaccgccat cctggccgcc 1450

ttagccagcc ctggcaatgc attgaggagt tcactccaga tcgccctgtg 1500

ccacttttte ttcaacatct ccggcatctt gctgtggtac ccgatcccgt 1550

tcactcgcct gcccatccgc atggccaagg ggctgggcaa catctctgcc 1600

aagtatcgct ggttcgccgt cttctacctg atcatcttct tcttcctgat 1650

cccgctgacg gtgtttggcc tctcgctggc cggctggcgg gtgctggttg 1700

gtgtcggggt tcccgtcgtc ttcatcatca tcctggtact gtgcctccga 1750

ctcctgcagt ctcgctgccc acgcgtcctg ccgaagaaac tccagaactg 1800

gaacttcctg ccgctgtgga tgcgctcgct gaagccctgg gatgccgtcg 1850

tctccaagtt caccggctgc ttccagatgc gctgctgctg ctgctgccgc 1900

gtgtgctgcc gcgcgtgctg cttgctgtgt ggctgcccca agtgctgccg 1950

ctgcagcaag tgctgcgagg acttggagga ggcgcaggag gggcaggatg 2000

tccctgtcaa ggctcctgag acctttgata acataaccat tagcagagag 2050

1369

gctcagggtg aggtccctgc ctcggactca aagaccgaat gcacggcctt 2100

gtaggggacg ccccagattg tcagggatgg ggggatggtc cttgagtttt 2150

gcatgctctc ctccctccca cttctgcacc ctttcaccac ctcgaggaga 2200

tttgctcccc attagcgaat gaaattgatg cagtcctacc taactcgatt 2250

ccctttggct tggtgggtag gcctgcaggg cacttttatt ccaacccatg 2300

gcctccatga ctttttcaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2350

aaaaaaaaaa aaaaaa 2366

<210> 30
<211> 1633
<212> DNA
<213> Homo sapiens

<400> 30
ggcgtgggag ctgcctaggc cgggccctgc cagggagcaa gtctgcatcg 50

cagaggccag ggcagagtgg ctcctcacag cctgaagctc atccttctgc 100

acgggccagc caggccagca cagaggcacc agggcagcag tgcacacagg 150

tccccgggga ccccaccatg tggagcggat ggtggctgtg gccccttgtg 200

gccgtctgca ctgcagactt ctttcgggac gaggcagaga ggatcatgag 250

ggactcccct gtcattgatg ggcacaatga cctcccctgg cagctgctgg 300

atatgttcaa caaccggctg caggacgaga gggccaacct gaccaccttg 350

gccggcacac acaccaacat ccccaagctg agggccggct ttgtgggagg 400

ccagttctgg tccgtgtaca cgccctgcga cacccagaac aaagacgccg 450

tgcggaggac gctggagcag atggacgtgg tccaccgcat gtgccggatg 500

tacccggaga ccttcctgta tgtcaccagc agtgcaggca ttcggcaggc 550

cttccgggaa gggaaggtgg ccagcctgat cggcgtggag ggcggccact 600

ccattgacag cagtttgggc gtcctgcggg cactctatca gctgggcatg 650

cggtacctga ccctcaccca cagctgcaac acgccctggg ctgacaactg 700

gctggtggac acgggagaca gcgagcccca gagccaaggc ttgtcaccct 750

ttgggcagcg tgtggtgaag gagctgaacc gtctgggggt cctcatcgac 800

ttggctcacg tgtctgtggc caccatgaag gccaccctgc agctgtccag 850

agccccggtc atcttcagcc actcctcggc ctacagcgtg tgcgcaagcc 900

ggcgcaacgt gcctgacgac gtcctgaggc tggtgaaaca gacagacagc 950

ctggtgatgg tgaacttcta caacaattac atttcctgca ccaacaaggc 1000

caacctgtcc caagtggccg accatctgga tcacatcaag gaggtggcag 1050

gagccagagc cgtgggtttt ggtggggact ttgatggtgt tccaagggtc 1100

cctgaggggc tggaggacgt ctccaagtat ccagacctga tcgctgagct 1150

gctcaggagg aactggacgg aggcggaggt caagggcgca ctggctgaca 1200

acctgctgag ggtcttccag gctgtggaac aggccagcaa cctcacacag 1250

gctcccgagg aggagcccat cccgctggac cagctgggtg gctcctgcag 1300

gacccattac ggctactcct ctggggcttc cagcctccat cgccactggg 1350

ggctcctgct ggcctccctc gctcccctgg tcctctgtct gtctctcctg 1400

1370

tgaaacctgg gagaccagag tccccttag ggttcccgga gctccgggaa 1450

gacccgccca tcccaggact ccagatgcca ggagccctgc tgcccacatg 1500

caaggaccag catctcctga gaggacgcct gggcttacct ggggggcagg 1550

atgcctgggg acagttcagg acacacacac agtaggcccg caataaaagc 1600

aacacccctt caaaaaaaaa aaaaaaaaaa aaa 1633

<210> 31
<211> 1589
<212> DNA
<213> Homo sapiens

<400> 31
catctaaagc ctcctcagcc ttctgagtca gcctgaaagg aacaggccga 50

actgctgtat gggctctact gccagtgtga cctcaccctc tccagtcacc 100

cctcctcagt tccagctatg agttcctgca acttcacaca tgccaccttt 150

gtgcttattg gtatcccagg attagagaaa gcccatttct gggttggctt 200

ccccctcctt tccatgtatg tagtggcaat gtttggaaac tgcatcgtgg 250

tcttcatcgt aaggacggaa cgcagcctgc acgctccgat gtacctcttt 300

ctctgcatgc ttgcagccat tgacctggcc ttatccacat ccaccatgcc 350

taagatcctt gccctttttct ggtttgattc ccgagagatt agctttgagg 400

cctgtcttac ccagatgttc tttattcatg ccctctcagc cattgaatcc 450

accatcctgc tggccatggc ctttgaccgt tatgtggcca tctgccaccc 500

actgcgccat gctgcagtgc tcaacaatac agtaacagcc cagattggca 550

tcgtggctgt ggtccgcgga tccctctttt ttttcccact gcctctgctg 600

atcaagcggc tggccttctg ccactccaat gtcctctcgc actcctattg 650

tgtccaccag gatgtaatga agttggccta tgcagacact ttgcccaatg 700

tggtatatgg tcttactgcc attctgctgg tcatgggcgt ggacgtaatg 750

ttcatctcct tgtcctattt tctgataata cgaacggttc tgcaactgcc 800

ttccaagtca gagcgggcca aggcctttgg aacctgtgtg tcacacattg 850

gtgtggtact cgccttctat gtgccactta ttggcctctc agttgtacac 900

cgctttggaa acagccttca tcccattgtg cgtgttgtca tgggtgacat 950

ctacctgctg ctgcctcctg tcatcaatcc catcatctat ggtgccaaaa 1000

ccaaacagat cagaacacgg gtgctggcta tgttcaagat cagctgtgac 1050

aaggacttgc aggctgtggg aggcaagtga cccttaacac tacacttctc 1100

cttatcttta ttggcttgat aaacataatt atttctaaca ctagcttatt 1150

tccagttgcc cataagcaca tcagtacttt tctctggctg gaatagtaaa 1200

ctaaagtatg gtacatctac ctaaaggact attatgtgga ataatacata 1250

ctaatgaagt attacatgat ttaaagacta caataaaacc aaacatgctt 1300

ataacattaa gaaaaacaat aaagatacat gattgaaacc aagttgaaaa 1350

atagcatatg ccttggagga aatgtgctca aattactaat gatttagtgt 1400

tgtccctact ttctctctct ttttttcttc ttttttttt attatggtta 1450

gctgtcacat acaacttttt ttttttttga gatggggtct ccagcctggg 1500

(377
1371

caacagagca agaccctgtc tcaaagcata aaatggaata acatatcaaa 1550

tgaaacaggg aaaatgaagc tgacaattta tgggagcca 1589

<210> 32
<211> 7941
<212> DNA
<213> Homo sapiens

<400> 32
cacacatacg cacgcacgat ctcacttcga tctatacact ggaggattaa 50

aacaaacaaa caaaaaaaac atttccttcg ctccccctcc ctctccactc 100

tgagaagcag aggagccgca cggcgagggg ccgcagaccg tctggaaatg 150

cgaatcctaa agcgtttcct cgcttgcatt cagctcctct gtgtttgccg 200

cctggattgg gctaatggat actacagaca acagagaaaa cttgttgaag 250

agattggctg gtcctataca ggagcactga atcaaaaaaa ttggggaaag 300

aaatatccaa catgtaatag cccaaaacaa tctcctatca atattgatga 350

agatcttaca caagtaaatg tgaatcttaa gaaacttaaa tttcagggtt 400

gggataaaac atcattggaa aacacattca ttcataacac tgggaaaaca 450

gtggaaatta atctcactaa tgactaccgt gtcagcggag gagtttcaga 500

aatggtgttt aaagcaagca agataacttt tcactgggga aaatgcaata 550

tgtcatctga tggatcagag catagtttag aaggacaaaa atttccactt 600

gagatgcaaa tctactgctt tgatgcggac cgattttcaa gttttgagga 650

agcagtcaaa ggaaaaggga agttaagagc tttatccatt ttgtttgagg 700

ttgggacaga agaaaatttg gatttcaaag cgattattga tggagtcgaa 750

agtgttagtc gttttgggaa gcaggctgct ttagatccat tcatactgtt 800

gaaccttctg ccaaactcaa ctgacaagta ttacatttac aatggctcat 850

tgacatctcc tccctgcaca gacacagttg actggattgt ttttaaagat 900

acagttagca tctctgaaag ccagttggct gtttttttgtg aagttcttac 950

aatgcaacaa tctggttatg tcatgctgat ggactactta caaaacaatt 1000

ttcgagagca acagtacaag ttctctagac aggtgttttc ctcatacact 1050

ggaaaggaag agattcatga agcagtttgt agttcagaac cagaaaatgt 1100

tcaggctgac ccagagaatt ataccagcct tcttgttaca tgggaaagac 1150

ctcgagtcgt ttatgatacc atgattgaga agtttgcagt tttgtaccag 1200

cagttggatg gagaggacca aaccaagcat gaatttttga cagatggcta 1250

tcaagacttg ggtgctattc tcaataattt gctacccaat atgagttatg 1300

ttcttcagat agtagccata tgcactaatg gcttatatgg aaaatacagc 1350

gaccaactga ttgtcgacat gcctactgat aatcctgaac ttgatctttt 1400

ccctgaatta attggaactg aagaaataat caaggaggag gaagagggaa 1450

aagacattga agaaggcgct attgtgaatc ctggtagaga cagtgctaca 1500

aaccaaatca ggaaaaagga accccagatt tctaccacaa cacactacaa 1550

tcgcataggg acgaaataca atgaagccaa gactaaccga tccccaacaa 1600

gaggaagtga attctctgga aagggtgatg ttcccaatac atctttaaat 1650

tccacttccc aaccagtcac taaattagcc acagaaaaag atatttcctt 1700

gacttctcag actgtgactg aactgccacc tcacactgtg gaaggtactt 1750

cagcctcttt aaatgatggc tctaaaactg ttcttagatc tccacatatg 1800

aacttgtcgg ggactgcaga atccttaaat acagtttcta taacagaata 1850

tgaggaggag agtttattga ccagtttcaa gcttgatact ggagctgaag 1900

attcttcagg ctccagtccc gcaacttctg ctatcccatt catctctgag 1950

aacatatccc aagggtatat attttcctcc gaaaacccag agacaataac 2000

atatgatgtc cttataccag aatctgctag aaatgcttcc gaagattcaa 2050

cttcatcagg ttcagaagaa tcactaaagg atccttctat ggagggaaat 2100

gtgtggtttc ctagctctac agacataaca gcacagcccg atgttggatc 2150

aggcagagag agctttctcc agactaatta cactgagata cgtgttgatg 2200

aatctgagaa gacaaccaag tccttttctg caggcccagt gatgtcacag 2250

ggtccctcag ttacagatct ggaaatgcca cattattcta cctttgccta 2300

cttcccaact gaggtaacac ctcatgcttt taccccatcc tccagacaac 2350

aggatttggt ctccacggtc aacgtggtat actcgcagac aacccaaccg 2400

gtatacaatg gtgagacacc tcttcaacct tcctacagta gtgaagtctt 2450

tcctctagtc acccctttgt tgcttgacaa tcagatcctc aacactaccc 2500

ctgctgcttc aagtagtgat tcggccttgc atgctacgcc tgtatttccc 2550

agtgtcgatg tgtcatttga atccatcctg tcttcctatg atggtgcacc 2600

tttgcttcca ttttcctctg cttccttcag tagtgaattg tttcgccatc 2650

tgcatacagt ttctcaaatc cttccacaag ttacttcagc taccgagagt 2700

gataaggtgc ccttgcatgc ttctctgcca gtggctgggg gtgatttgct 2750

attagagccc agccttgctc agtattctga tgtgctgtcc actactcatg 2800

ctgcttcaga gacgctggaa tttggtagtg aatctggtgt tctttataaa 2850

acgcttatgt tttctcaagt tgaaccaccc agcagtgatg ccatgatgca 2900

tgcacgttct tcagggcctg aaccttctta tgccttgtct gataatgagg 2950

gctcccaaca catcttcact gtttcttaca gttctgcaat acctgtgcat 3000

gattctgtgg gtgtaactta tcagggttcc ttatttagcg gccctagcca 3050

tataccaata cctaagtctt cgttaataac cccaactgca tcattactgc 3100

agcctactca tgccctctct ggtgatgggg aatggtctgg agcctcttct 3150

gatagtgaat ttcttttacc tgacacagat gggctgacag cccttaacat 3200

ttcttcacct gtttctgtag ctgaatttac atatacaaca tctgtgtttg 3250

gtgatgataa taaggcgctt tctaaaagtg aaataatata tggaaatgag 3300

actgaactgc aaattccttc tttcaatgag atggtttacc cttctgaaag 3350

cacagtcatg cccaacatgt atgataatgt aaataagttg aatgcgtctt 3400

tacaagaaac ctctgtttcc atttctagca ccaagggcat gtttccaggg 3450

tcccttgctc ataccaccac taaggttttt gatcatgaga ttagtcaagt 3500

tccagaaaat aacttttcag ttcaacctac acatactgtc tctcaagcat 3550

ctggtgacac ttcgcttaaa cctgtgctta gtgcaaactc agagccagca 3600

tcctctgacc ctgcttctag tgaaatgtta tctccttcaa ctcagctctt 3650

attttatgag acctcagctt cttttagtac tgaagtattg ctacaacctt 3700

cctttcaggc ttctgatgtt gacaccttgc ttaaaactgt tcttccagct 3750

gtgcccagtg atccaatatt ggttgaaacc cccaaagttg ataaaattag 3800

ttctacaatg ttgcatctca ttgtatcaaa ttctgcttca agtgaaaaca 3850

tgctgcactc tacatctgta ccagtttttg atgtgtcgcc tacttctcat 3900

atgcactctg cttcacttca aggtttgacc atttcctatg caagtgagaa 3950

atatgaacca gttttgttaa aaagtgaaag ttcccaccaa gtggtacctt 4000

ctttgtacag taatgatgag ttgttccaaa cggccaattt ggagattaac 4050

caggcccatc ccccaaaagg aaggcatgta tttgctacac ctgttttatc 4100

aattgatgaa ccattaaata cactaataaa taagcttata cattccgatg 4150

aaattttaac ctccaccaaa agttctgtta ctggtaaggt atttgctggt 4200

attccaacag ttgcttctga tacatttgta tctactgatc attctgttcc 4250

tataggaaat gggcatgttg ccattacagc tgtttctccc cacagagatg 4300

gttctgtaac ctcaacaaag ttgctgtttc cttctaaggc aacttctgag 4350

ctgagtcata gtgccaaatc tgatgccggt ttagtgggtg gtggtgaaga 4400

tggtgacact gatgatgatg gtgatgatga tgatgacaga gatagtgatg 4450

gcttatccat tcataagtgt atgtcatgct catcctatag agaatcacag 4500

gaaaaggtaa tgaatgattc agacacccac gaaaacagtc ttatggatca 4550

gaataatcca atctcatact cactatctga gaattctgaa gaagataata 4600

gagtcacaag tgtatcctca gacagtcaaa ctggtatgga cagaagtcct 4650

ggtaaatcac catcagcaaa tgggctatcc caaaagcaca atgatggaaa 4700

agaggaaaat gacattcaga ctggtagtgc tctgcttcct ctcagccctg 4750

aatctaaagc atgggcagtt ctgacaagtg atgaagaaag tggatcaggg 4800

caaggtacct cagatagcct taatgagaat gagacttcca cagatttcag 4850

ttttgcagac actaatgaaa aagatgctga tgggatcctg gcagcaggtg 4900

actcagaaat aactcctgga ttcccacagt ccccaacatc atctgttact 4950

agcgagaact cagaagtgtt ccacgtttca gaggcagagg ccagtaatag 5000

tagccatgag tctcgtattg gtctagctga ggggttggaa tccgagaaga 5050

aggcagttat acccccttgtg atcgtgtcag ccctgacttt tatctgtcta 5100

gtggttcttg tgggtattct catctactgg aggaaatgct tccagactgc 5150

acacttttac ttagaggaca gtacatcccc tagagttata tccacacctc 5200

caacacctat ctttccaatt tcagatgatg tcggagcaat tccaataaag 5250

cactttccaa agcatgttgc agatttacat gcaagtagtg ggtttactga 5300

agaatttgag acactgaaag agttttacca ggaagtgcag agctgtactg 5350

ttgacttagg tattacagca gacagctcca accacccaga caacaagcac 5400

aagaatcgat acataaatat cgttgcctat gatcatagca gggttaagct 5450

agcacagctt gctgaaaagg atggcaaact gactgattat atcaatgcca 5500

attatgttga tggctacaac agaccaaaag cttatattgc tgcccaaggc 5550

ccactgaaat ccacagctga agatttctgg agaatgatat gggaacataa 5600

tgtggaagtt attgtcatga taacaaacct cgtggagaaa ggaaggagaa 5650

aatgtgatca gtactggcct gccgatggga gtgaggagta cgggaacttt 5700

ctggtcactc agaagagtgt gcaagtgctt gcctattata ctgtgaggaa 5750

ttttactcta agaaacacaa aaataaaaaa gggctcccag aaaggaagac 5800

ccagtggacg tgtggtcaca cagtatcact acacgcagtg gcctgacatg 5850

ggagtaccag agtactccct gccagtgctg acctttgtga gaaaggcagc 5900

ctatgccaag cgccatgcag tggggcctgt tgtcgtccac tgcagtgctg 5950

gagttggaag aacaggcaca tatattgtgc tagacagtat gttgcagcag 6000

attcaacacg aaggaactgt caacatattt ggcttcttaa aacacatccg 6050

ttcacaaaga aattatttgg tacaaactga ggagcaatat gtcttcattc 6100

atgatacact ggttgaggcc atacttagta aagaaactga ggtgctggac 6150

agtcatattc atgcctatgt taatgcactc ctcattcctg gaccagcagg 6200

caaaacaaag ctagagaaac aattccagct cctgagccag tcaaatatac 6250

agcagagtga ctattctgca gccctaaagc aatgcaacag ggaaaagaat 6300

cgaacttctt ctatcatccc tgtggaaaga tcaagggttg gcatttcatc 6350

cctgagtgga gaaggcacag actacatcaa tgcctcctat atcatgggct 6400

attaccagag caatgaattc atcattaccc agcaccctct ccttcatacc 6450

atcaaggatt tctggaggat gatatgggac cataatgccc aactggtggt 6500

tatgattcct gatggccaaa acatggcaga agatgaattt gtttactggc 6550

caaataaaga tgagcctata aattgtgaga gctttaaggt cactcttatg 6600

gctgaagaac acaaatgtct atctaatgag gaaaaactta taattcagga 6650

ctttatctta gaagctacac aggatgatta tgtacttgaa gtgaggcact 6700

ttcagtgtcc taaatggcca aatccagata gccccattag taaaactttt 6750

gaacttataa gtgttataaa agaagaagct gccaataggg atgggcctat 6800

gattgttcat gatgagcatg gaggagtgac ggcaggaact ttctgtgctc 6850

tgacaaccct tatgcaccaa ctagaaaaag aaaattccgt ggatgtttac 6900

caggtagcca agatgatcaa tctgatgagg ccaggagtct ttgctgacat 6950

tgagcagtat cagtttctct acaaagtgat cctcagcctt gtgagcacaa 7000

ggcaggaaga gaatccatcc acctctctgg acagtaatgg tgcagcattg 7050

cctgatggaa atatagctga gagcttagag tctttagttt aacacagaaa 7100

ggggtggggg gactcacatc tgagcattgt tttcctcttc ctaaaattag 7150

gcaggaaaat cagtctagtt ctgttatctg ttgatttccc atcacctgac 7200

agtaactttc atgacatagg attctgccgc caaatttata tcattaacaa 7250

tgtgtgcctt tttgcaagac ttgtaattta cttattatgt ttgaactaaa 7300

atgattgaat tttacagtat ttctaagaat ggaattgtgg tatttttttc 7350

tgtattgatt ttaacagaaa atttcaattt atagaggtta ggaattccaa 7400

actacagaaa atgtttgttt ttagtgtcaa atttttagct gtatttgtag 7450

caattatcag gtttgctaga aatataactt ttaatacagt agcctgtaaa 7500

taaaacactc ttccatatga tattcaacat tttacaactg cagtattcac 7550

ctaaagtaga aataatctgt tacttattgt aaatactgcc ctagtgtctc 7600

catggaccaa atttatattt ataattgtag atttttatat tttactactg 7650

agtcaagttt tctagttctg tgtaattgtt tagtttaatg acgtagttca 7700

ttagctggtc ttactctacc agttttctga cattgtattg tgttacctaa 7750

gtcattaact ttgtttcagc atgtaatttt aacttttgtg gaaaatagaa 7800

ataccttcat tttgaaagaa gtttttatga gaataacacc ttaccaaaca 7850

ttgttcaaat ggtttttatc caaggaattg caaaaataaa tataaatatt 7900

gccattaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa a 7941

<210> 33
<211> 2755
<212> DNA
<213> Homo sapiens

<400> 33
attgtctggg aattgcagcc gcggggcggg cggcggcggc ggcggcggcg 50

gccgggaccc agcgggccag gtggggacgg cgcggagcgg gtgcgggaga 100

tgccgtgcgg gactggggcc acctgagccg cccgcctcgt ccccgccttc 150

tgtgggaagg atgtgcgcgc ggatggccgg tcgcacaaga gcggcccctc 200

gggggcccta cggcccctgg ctctgcctcc tggtggccct cgccctggac 250

gtcgtgagag tggactgtgg ccaggctccc ctggaccctg tctacctgca 300

tgtgacagcc gcccgcccag cccagcccac actctggact gccaagctag 350

accgcttcaa gggctccagg caccacacca ccctcatcac ctgccaccgt 400

gctgggctca cagagccaga ttccagcagt ccccttgaac tgtctgagtt 450

cctatgggtg gactttgtgg tggagaatag cactggtggg ggcgtagcgg 500

tcactcgccc cgtcacgtgg cagctggagt acccaggcca ggcccctgaa 550

gcagagaagg acaaaatggt gtgggaaatc ctggtgtctg agcgggacat 600

cagagcccct atcccactgg ccaaggctga ggagctggtg aatacagcac 650

cactgactgg agtgccccag catgtccccg tgcgccttgt cactgtggac 700

ggcgggggcg ccttggtgga ggtgacagag catgtcggct gcgagtctgc 750

caacacacag gtcctgcagg tgtctgaggc ctgtgatgcc gtgttcgtgg 800

ctggcaagga gagccgggcc gcccgggggg tgcgagtgga cttctggtgg 850

cgccggctcc gcgcctcgct gcggctgacc gtgtgggccc cgctgctacc 900

gctgcgtatc gagctcaccg acaccaccct cgagcaggtc cgcggctgga 950

gggtacctgg ccctgctgaa gggcctgcgg aacccgctgc agaggcgtca 1000

gatgaggccg agcggcgcgc ccgtggctgc cacctgcagt accagcgggc 1050

cggtgtgcgc ttcctcgccc ccttcgcggc ccacccgctg gacggcggcc 1100

gccgcctcac gcacctgctt ggccccgact ggctgctaga cgtgtcccac 1150

ctcgtggcgc cacacgcccg cgtgctggac tcgcgtgtag cctctctgga 1200

gggtggccgt gtcgtggtgg gccgggagcc cggtgtcacc tccattgagg 1250

tgcgttcccc actgtctgac tccatcctgg gggagcaggc gctggctgtg 1300

acggacgaca aggtctcagt gctggagctg agggtgcagc cagtgatggg 1350

catctcgctg accttgagcc ggggcactgc ccaccccggg gaggtcacag 1400

ctacgtgctg ggcacagtca gcccttcccg ccccaaagca ggaggtggcc 1450

ctctccctat ggctgtcctt ctctgatcac actgtggccc cagctgagct 1500

ctacgaccgc cgtgacctgg gactgtccgt ctcagccgag gagcctggtg 1550

ccatcctgcc agctgaggag cagggtgccc agctcggggt ggtggtgagt 1600

ggggcaggcg ccgaggggct gccgctgcat gtggctctgc acccgcccga 1650

gccctgccgc cggggccgcc accgtgtgcc tctggcctct ggcaccgcct 1700

ggctggggct gcccccctgcc tccactccag cccctgctct cccatccagc 1750

cctgcttgga gcccaccagc cacagaagcc accatgggtg gtaaacggca 1800

ggtggcaggc agtgtcgggg gcaacacagg tgtgaggggc aagtttgagc 1850

gggcagagga ggaggccagg aaggaggaga ccgaagccag ggaggaggag 1900

gaggaagagg aggaggagat ggtccctgcc cctcagcatg tcactgagct 1950

agagctgggc atgtacgccc tgctgggagt cttctgcgtg gccatcttca 2000

tcttcttggt caatggtgtg gtcttcgtcc tgcgctatca gcgcaaagaa 2050

cctcccgaca gtgccactga ccccacctcc ccccagcccc acaactgggt 2100

ctggctgggc actgaccagg aggaactgag ccgccagctg gaccggcagt 2150

cccctggccc gcccaagggg gaggggagct gcccctgtga gagtggggga 2200

ggaggggagg cccctaccct ggcccctggc cctcctgggg gcaccaccag 2250

ctcctcaagc accctggccc gaaaggaggc tgggggggcgg cggaagcgag 2300

tagagtttgt gacatttgtg ccagcccctc cagcccagtc acctgaggag 2350

cctgtagggg cccctgctgt gcagtccatc cttgtggcag gcgaggagga 2400

catccgctgg gtgtgtgagg acatggggct gaaggaccct gaggagcttc 2450

gcaactacat ggagaggatc cggggcagct cctgaccctc cacagccacc 2500

tggtcagcca ccagctgggg caacgagggt ggaggtccca ctgagcctct 2550

cgcctgcccc cgccactcgt ctggtgcttg ttgatccaag tcccctgcct 2600

ggtcccccac aaggactccc atccaggccc cctctgccct gccccttgtc 2650

atggaccatg gtcgtgagga agggctcatg ccccttattt atgggaacca 2700

tctcattcta acagaataaa ccgagaagga aaccagaaaa aaaaaaaaaa 2750

aaaaa 2755

<210> 34
<211> 1795
<212> DNA
<213> Homo sapiens

<400> 34
acgcgtccgc ccacgcgtcc gcccacgcgt ccggtcgggg ccagagcgca 50

ggtgtacctg gcggccgtgc tggagcacct gaccgccgag atcctggagc 100

tggctggcaa cccggcccgc gacaagaaga cccgcatcat cctgcgccac 150

ctgtagctgg ccattcgcaa cggcgaggag cttaacaagc tgctgggcga 200

agtcaccatc gcgcagggcg gtgtcctgcc caacattcag ggcgtgcttc 250

tgccccagaa gaccaagagc caccacaagg ccaagggtga aaaccattca 300

ctaggagagg agaaacacaa tggccaccaa gacagagttg agtcccacag 350

caagggagag caagaacgca caagatatgc aagtggatga gacactgatc 400

cccaggaaag gtccaagttt atgttctgct cgctatggaa tagccctcgt 450

cttacatttc tgcaatttca caacgatagc acaaaatgtc atcatgaaca 500

tcaccatggt agccatggtc aacagcacaa gccctcaatc ccagctcaat 550

gattcctctg aggtgctgcc tgttgactca tttggtggcc taagtaaagc 600

cccaaagagt cttcctgcaa agtcctcaat acttgggggt cagtttgcaa 650

tttgggaaaa gtggggccct ccacaagaac gaagcagact ctgcagcatt 700

gctttatcag gaatgttact gggatgcttt actgccatcc tcataggtgg 750

cttcattagt gaaacccttg ggtggccctt tgtcttctat atctttggag 800

gtgttggctg tgtctgctgc cttctctggt ttgttgtgat ttatgatgac 850

ccctttтcct atccatggat aagcacctca gaaaaagaat acatcatatc 900

ctccttgaaa caacaggtcg ggtcttctaa gcagcctctt cccatcaaag 950

ctatgctcag atctctaccc atttggtcca tatgtttagg ctgtttcagc 1000

catcaatggt tagttagcac aatggttgta tacataccaa cttacatcag 1050

ctctgtgtac catgttaaca tcagagacaa tggacttcta tctgcccttc 1100

cttttattgt tgcctgggtc ataggcatgg tgggaggcta tctggcagat 1150

ttccttctaa ccaaaaagtt tagactcatc actgtgagga aaattgccac 1200

aattttagga agtctcccct cttcagcact cattgtgtct ctgccttacc 1250

tcaattccgg ctatatcaca gcaactgcct tgctgacgct ctcttgcgga 1300

ttaagcacat tgtgtcagtc agggatttat atcaatgtct tagatattgc 1350

tccaaggtat tccagttttc tcatgggagc atcaagagga ttttcgagca 1400

tagcacctgt cattgtaccc actgtcagcg gatttcttct tagtcaggac 1450

cctgagtttg ggtggaggaa tgtcttcttc ttgctgtttg ccgttaacct 1500

gttaggacta ctcttctacc tcatatttgg agaagcagat gtccaagaat 1550

gggctaaaga gagaaaactc actcgtttat gaagttatcc caccttggat 1600

ggaaaagtca ttaggcaccg tattgcataa aatagaaggc ttccgtgatg 1650

aaaataccag tgaaaagatt ttttttтcct gtggctcttt tcaattatga 1700

gatcagttca ttattttatt cagacttttt tttgagagaa atgtaagatg 1750

aataaaaatt caaataaaat gataactaag aaaaaaaaaa aaaaa 1795

<210> 35
<211> 1795
<212> DNA
<213> Homo sapiens

<400> 35
acgcgtccgc ccacgcgtcc gcccacgcgt ccggtcgggg ccagagcgca 50

1378

ggtgtacctg gcggccgtgc tggagcacct gaccgccgag atcctggagc 100

tggctggcaa cccggcccgc gacaagaaga cccgcatcat cctgcgccac 150

ctgtagctgg ccattcgcaa cggcgaggag cttaacaagc tgctgggcga 200

agtcaccatc gcgcagggcg gtgtcctgcc caacattcag ggcgtgcttc 250

tgccccagaa gaccaagagc caccacaagg ccaagggtga aaaccattca 300

ctaggagagg agaaacacaa tggccaccaa gacagagttg agtcccacag 350

caagggagag caagaacgca caagatatgc aagtggatga gacactgatc 400

cccaggaaag gtccaagttt atgttctgct cgctatggaa tagccctcgt 450

cttacatttc tgcaatttca caacgatagc acaaaatgtc atcatgaaca 500

tcaccatggt agccatggtc aacagcacaa gccctcaatc ccagctcaat 550

gattcctctg aggtgctgcc tgttgactca tttggtggcc taagtaaagc 600

cccaaagagt cttcctgcaa agtcctcaat acttgggggt cagtttgcaa 650

tttgggaaaa gtggggccct ccacaagaac gaagcagact ctgcagcatt 700

gctttatcag gaatgttact gggatgcttt actgccatcc tcataggtgg 750

cttcattagt gaaacccttg ggtggccctt tgtcttctat atctttggag 800

gtgttggctg tgtctgctgc cttctctggt ttgttgtgat ttatgatgac 850

ccctttтcct atccatggat aagcacctca gaaaaagaat acatcatatc 900

ctccttgaaa caacaggtcg ggtcttctaa gcagcctctt cccatcaaag 950

ctatgctcag atctctaccc atttggtcca tatgtttagg ctgtttcagc 1000

catcaatggt tagttagcac aatggttgta tacataccaa cttacatcag 1050

ctctgtgtac catgttaaca tcagagacaa tggacttcta tctgcccttc 1100

cttttattgt tgcctgggtc ataggcatgg tgggaggcta tctggcagat 1150

ttccttctaa ccaaaaagtt tagactcatc actgtgagga aaattgccac 1200

aattttagga agtctccccт cttcagcact cattgtgtct ctgccttacc 1250

tcaattccgg ctatatcaca gcaactgcct tgctgacgct ctcttgcgga 1300

ttaagcacat tgtgtcagtc agggatttat atcaatgtct tagatattgc 1350

tccaaggtat tccagttttc tcatgggagc atcaagagga ttttcgagca 1400

tagcacctgt cattgtaccc actgtcagcg gatttcttct tagtcaggac 1450

cctgagtttg ggtggaggaa tgtcttcttc ttgctgtttg ccgttaacct 1500

gttaggacta ctcttctacc tcatatttgg agaagcagat gtccaagaat 1550

gggctaaaga gagaaaactc actcgtttat gaagttatcc caccttggat 1600

ggaaaagtca ttaggcaccg tattgcataa aatagaaggc ttccgtgatg 1650

aaaataccag tgaaaagatt ttttttтcct gtggctcttt tcaattatga 1700

gatcagttca ttatttтatt cagacttttt tttgagagaa atgtaagatg 1750

aataaaaatt caaataaaat gataactaag aaaaaaaaaa aaaaa 1795

<210> 36
<211> 4524
<212> DNA
<213> Homo sapiens

<400> 36

gctcgctggg ccgctgctcc cgggtgtccc aggcccggcc ggtgcgcaga 50

gcatggcggg tgcgggcccg aagcggcgcg cgctagcggc gccggcggcc 100

gaggagaagg aagaggcgcg ggagaagatg ctggccgcca agagcgcgga 150

cggctcggcg ccggcaggcg agggcgaggg cgtgaccctg cagcggaaca 200

tcacgctgct caacggcgtg gccatcatcg tggggaccat tatcggctcg 250

ggcatcttcg tgacgcccac gggcgtgctc aaggaggcag gctcgccggg 300

gctggcgctg gtggtgtggg ccgcgtgcgg cgtcttctcc atcgtgggcg 350

cgctctgcta cgcggagctc ggcaccacca tctccaaatc gggcggcgac 400

tacgcctaca tgctggaggt ctacggctcg ctgcccgcct tcctcaagct 450

ctggatcgag ctgctcatca tccggccttc atcgcagtac atcgtggccc 500

tggtcttcgc cacctacctg ctcaagccgc tcttccccac ctgcccggtg 550

cccgaggagg cagccaagct cgtggcctgc ctgtgcgtgc tgctgctcac 600

ggccgtgaac tgctacagcg tgaaggccgc cacccgggtc caggatgcct 650

ttgccgccgc caagctcctg gccctggccc tgatcatcct gctgggcttc 700

gtccagatcg gaaagggtga tgtgtccaat ctagatccca agttctcatt 750

tgaaggcacc aaactggatg tggggaacat tgtgctggca ttatacagcg 800

gcctctttgc ctatggagga tggaattact tgaatttcgt cacagaggaa 850

atgatcaacc cctacagaaa cctgcccctg gccatcatca tctccctgcc 900

catcgtgacg ctggtgtacg tgctgaccaa cctggcctac ttcaccaccc 950

tgtccaccga gcagatgctg tcgtccgagg ccgtggccgt ggacttcggg 1000

aactatcacc tgggcgtcat gtcctggatc atccccgtct tcgtgggcct 1050

gtcctgcttt ggctccgtca atgggtccct gttcacatcc tccaggctct 1100

tcttcgtggg gtcccgggaa ggccacctgc cctccatcct ctccatgatc 1150

cacccacagc tcctcacccc cgtgccgtcc ctcgtgttca cgtgtgtgat 1200

gacgctgctc tacgccttct ccaaggacat cttctccgtc atcaacttct 1250

tcagcttctt caactggctc tgcgtggccc tggccatcat cggcatgatc 1300

tggctgcgcc acagaaagcc tgagcttgag cggcccatca aggtgaacct 1350

ggccctgcct gtgttcttca tcctggcctg cctcttcctg atcgccgtct 1400

ccttctggaa gacacccgtg gagtgtggca tcggcttcac catcatcctc 1450

agcgggctgc ccgtctactt cttcggggtc tggtggaaaa acaagcccaa 1500

gtggctcctc cagggcatct tctccacgac cgtcctgtgt cagaagctca 1550

tgcaggtggt cccccaggag acatagccag gaggccgagt ggctgccgga 1600

ggagcatgcg cagaggccag ttaaagtaga tcacctcctc gaacccactc 1650

cggttccccg caacccacag ctcagctgcc catcccagtc ctcgccgtcc 1700

ctcccaggtc gggcagtgga ggctgctgtg aaaactctgg tacgaatctc 1750

atccctcaac tgagggccag ggacccaggt gtgcctgtgc tcctgcccag 1800

gagcagcttt tggtctcctt gggccctttt tcccttccct cctttgttta 1850

cttatatata tatttttttt aaacttaaat tttgggtcaa cttgacacca 1900

ctaagatgat tttttaagga gctgggggaa ggcaggagcc ttcctttctc 1950

ctgccccaag ggcccagacc ctgggcaaac agagctactg agacttggaa 2000

cctcattgct accacagact tgcactgaag ccagacagct gcccagacac 2050

atgggcttgt gacattcgtg aaaaccaacc ctgtgggctt atgtctctgc 2100

cttagggttt gcagagtgga aactcagccg tagggtggca ctgggagggg 2150

gtgggggatc tgggcaaggt gggtgattcc tcccaggagg tgcttgaggc 2200

cccgatggac tcctgaccat aatcctagcc ccgagacacc atcctgagcc 2250

agggaacagc cccagggttg gggggtgccg gcatctcccc tagctcacca 2300

ggcctggcct ctgggcagtg tggcctcttg gctatttctg ttccagtttt 2350

ggaggctgag ttctggttca tgcagacaaa gccctgtcct tcagtcttct 2400

agaaacagag acaagaaagg cagacacacc gcggccaggc acccatgtgg 2450

gcgcccaccc tgggctccac acagcagtgt cccctgcccc agaggtcgca 2500

gctaccctca gcctccaatg cattggcctc tgtaccgccc ggcagcccct 2550

tctggccggt gctgggttcc cactcccggc ctaggcacct ccccgctctc 2600

cctgtcacgc tcatgtcctg tcctggtcct gatgcccgtt gtctaggaga 2650

cagagccaag cactgctcac gtctctgccg cctgcgtttg gaggcccctg 2700

ggctctcacc cagtccccac ccgcctgcag agagggaact agggcacccc 2750

ttgtttctgt tgttcccgtg aatttttttc gctatgggag gcagccgagg 2800

cctggccaat gcggcccact ttcctgagct gtcgctgcct ccatggcagc 2850

agccaaggac ccccagaaca agaagacccc cccgcaggat ccctcctgag 2900

ctcggggggc tctgccttct caggccccgg gcttcccttc tccccagcca 2950

gaggtggagc caagtggtcc agcgtcactc cagtgctcag ctgtggctgg 3000

aggagctggc ctgtggcaca gccctgagtg tcccaagccg ggagccaacg 3050

aagccggaca cggcttcact gaccagcggc tgctcaagcc gcaagctctc 3100

agcaagtgcc cagtggagcc tgccgccccc acctgggcac cgggaccccc 3150

tcaccatcca gtgggcccgg agaaacctga tgaacagttt ggggactcag 3200

gaccagatgt ccgtctctct tgcttgagga atgaagacct ttattcaccc 3250

ctgccccgtt gcttcccgct gcacatggac agacttcaca gcgtctgctc 3300

ataggacctg catccttcct ggggacgaat tccactcgtc caagggacag 3350

cccacggtct ggaggccgag gaccaccagc aggcaggtgg actgactgtg 3400

ttgggcaaga cctcttccct ctgggcctgt tctcttggct gcaaataagg 3450

acagcagctg gtgccccacc tgcctggtgc attgctgtgt gaatccagga 3500

ggcagtggac atcgtaggca gccacggccc caggtccagg agaagtgctc 3550

cctggaggca cggaccactg cttcccactg gggccggcgg ggcccacgca 3600

cgacgtcagc ctcttacctt cccgcctcgg ctaggggtcc tcgggatgcc 3650

gttctgttcc aacctcctgt tctgggaggt ggacatgcct caaggataca 3700

gggagccggc ggcctctcga cggcacgcac ttcctgttgg ctgctgcggc 3750

tgtgggcgag catgggggct gccagcgtct gttgtggaaa gtagctgcta 3800

gtgaaatggc tggggccgct ggggtccgtc ttcacactgc gcaggtctct 3850

tctgggcgtc tgagctgggg tgggagctcc tccgcagaag gttggtgggg 3900

ggtccagtct gtgatccttg gtgctgtgtg ccccactcca gcctggggac 3950

cccacttcag aaggtagggg ccgtgtcccg cggtgctgac tgaggcctgc 4000

ttccccctcc ccctcctgct gtgctggaat tccacaggga ccagggccac 4050

cgcaggggac tgtctcagaa gacttgattt ttccgtccct ttttctccac 4100

actccactga caaacgtccc cagcggtttc cacttgtggg cttcaggtgt 4150

tttcaagcac aacccaccac aacaagcaag tgcattttca gtcgttgtgc 4200

ttttttgttt tgtgctaacg tcttactaat ttaaagatgc tgtcggcacc 4250

atgtttattt atttccagtg gtcatgctca gccttgctgc tctgcgtggc 4300

gcaggtgcca tgcctgctcc ctgtctgtgt cccagccacg cagggccatc 4350

cactgtgacg tcggccgacc aggctggaca ccctctgccg agtaatgacg 4400

tgtgtggctg ggaccttctt tattctgtgt taatggctaa cctgttacac 4450

tgggctgggt tgggtagggt gttctggctt ttttgtgggg tttttatttt 4500

taaagaaaca ctcaatcatc ctag 4524

<210> 37
<211> 1590
<212> DNA
<213> Homo sapiens

<400> 37
attcgccctt tgctcccggg tgtcccaggc ccggccggtg cgcagagcat 50

ggcgggtgcg ggcccgaagc ggcgcgcgct agcggcgccg gcggccgagg 100

agaaggaaga ggcgcgggag aagatgctgg ccgccaagag cgcggacggc 150

tcggcgccgg caggcgaggg cgagggcgtg accctgcagc ggaacatcac 200

gctgctcaac ggcgtggcca tcatcgtggg gaccattatc ggctcgggca 250

tcttcgtgac gcccacgggc gtgctcaagg aggcaggctc gccggggctg 300

gcgctggtgg tgtgggccgc gtgcggcgtc ttctccatcg tgggcgcgct 350

ctgctacgcg gagctcggca ccaccatctc caaatcgggc ggcgactacg 400

cctacatgct ggaggtctac ggctcgctgc ccgccttcct caagctctgg 450

atcgagctgc tcatcatccg gccttcatcg cagtacatcg tggccctggt 500

cttcgccacc tacctgctca agccgctctt ccccacctgc ccggtgcccg 550

aggaggcagc caagctcgtg gcctgcctct gcgtgctgct gctcacggcc 600

gtgaactgct acagcgtgaa ggccgccacc cgggtccagg atgcctttgc 650

cgccgccaag ctcctggccc tggccctgat catcctgctg ggcttcgtcc 700

agatcgggaa gggtgatgtg tccaatctag atcccaactt ctcatttgaa 750

ggcaccaaac tggatgtggg gaacattgtg ctggcattat acagcggcct 800

ctttgcctat ggaggatgga attacttgaa tttcgtcaca gaggaaatga 850

tcaaccccta cagaaacctg cccctggcca tcatcatctc cctgcccatc 900

gtgacgctgg tgtacgtgct gaccaacctg gcctacttca ccaccctgtc 950

caccgagcag atgctgtcgt ccgaggccgt ggccgtggac ttcgggaact 1000

1382

atcacctggg cgtcatgtcc tggatcatcc ccgtcttcgt gggcctgtcc 1050

tgcttcggct ccgtcaatgg gtccctgttc acatcctcca ggctcttctt 1100

cgtggggtcc cgggaaggcc acctgccctc catcctctcc atgatccacc 1150

cacagctcct cacccccgtg ccgtccctcg tgttcacgtg tgtgatgacg 1200

ctgctctacg ccttctccaa ggacatcttc tccgtcatca acttcttcag 1250

cttcttcaac tggctctgcg tggccctggc catcatcggc atgatctggc 1300

tgcgccacag aaagcctgag cttgagcggc ccatcaaggt gaacctggcc 1350

ctgcctgtgt tcttcatcct ggcctgcctc ttcctgatcg ccgtctcctt 1400

ctggaagaca cccgtggagt gtggcatcgg cttcaccatc atcctcagcg 1450

ggctgcccgt ctacttcttc ggggtctggt ggaaaaacaa gcccaagtgg 1500

ctcctccagg gcatcttctc cacgaccgtc ctgtgtcaga agctcatgca 1550

ggtggtcccc caggagacat agccaggagg ccgaagggcg 1590

<210> 38
<211> 2436
<212> DNA
<213> Homo sapiens

<400> 38
ccaggctctc cacccccact tcccaattga ggaaaccgag gcagaggagg 50

ctcaggtgtg gccaatcacc ctgcacatca gagttaccct gggcagggcc 100

cactgagacc tgggaggggc cactcgggac ctggagggct gggggctgcc 150

cgggcgttag gggtaaagct ccctacccaa ctgcgcagaa ggcctcagag 200

gcctgggggc tgggcttccc ctttcacatc gccctttaga ggcccacgtg 250

tgggcattgg cccgcgatct gaaaggggct gtcctgttcc tcatgggcgc 300

tgccagcgcc acgcactcct ctttctgcct ggccggccac tcccgtctgc 350

tgtgacgcgc ggacagagag ctaccggtgg acccacggtg cctccctccc 400

tgggatctac acagaccatg gccttgccaa cggctcgacc cctgttgggg 450

tcctgtggga ccccgccct cggcagcctc ctgttcctgc tcttcagcct 500

cggatgggtg cagccctcga ggaccctggc tggagagaca gggcaggagg 550

ctgcacccct ggacggagtc ctggccaacc cacctaacat ttccagcctc 600

tcccctcgcc aactccttgg cttcccgtgt gcggaggtgt ccggcctgag 650

cacggagcgt gtccgggagc tggctgtggc cttggcacag aagaatgtca 700

agctctcaac agagcagctg cgctgtctgg ctcaccggct ctctgagccc 750

cccgaggacc tggacgccct cccattggac ctgctgctat tcctcaaccc 800

agatgcgttc tcggggcccc aggcctgcac ccgtttcttc tcccgcatca 850

cgaaggccaa tgtggacctg ctcccgaggg gggctcccga gcgacagcgg 900

ctgctgcctg cggctctggc ctgctggggt gtgcgggggt ctctgctgag 950

cgaggctgat gtgcgggctc tgggaggcct ggcttgcgac ctgcctgggc 1000

gctttgtggc cgagtcggcc gaagtgctgc taccccggct ggtgagctgc 1050

ccgggacccc tggaccagga ccagcaggag gcagccaggg cggctctgca 1100

gggcgggggа ccccctacg gccccccgtc gacatggtct gtctccacga 1150

tggacgctct gcggggcctg ctgcccgtgc tgggccagcc catcatccgc 1200

agcatcccgc agggcatcgt ggccgcgtgg cggcaacgct cctctcggga 1250

cccatcctgg cggcagcctg aacggaccat cctccggccg cggttccggc 1300

gggaagtgga gaagacagcc tgtccttcag gcaagaaggc ccgcgagata 1350

gacgagagcc tcatcttcta caagaagtgg gagctggaag cctgcgtgga 1400

tgcggccctg ctggccaccc agatggaccg cgtgaacgcc atcccccttca 1450

cctacgagca gctggacgtc ctaaagcata aactggatga gctctaccca 1500

caaggttacc ccgagtctgt gatccagcac ctgggctacc tcttcctcaa 1550

gatgagccct gaggacattc gcaagtggaa tgtgacgtcc ctggagaccc 1600

tgaaggcttt gcttgaagtc aacaaagggc acgaaatgag tcctcaggtg 1650

gccaccctga tcgaccgctt tgtgaaggga aggggccagc tagacaaaga 1700

caccctagac accctgaccg ccttctaccc tgggtacctg tgctccctca 1750

gccccgagga gctgagctcc gtgccccccca gcagcatctg ggcggtcagg 1800

ccccaggacc tggacacgtg tgacccaagg cagctggacg tcctctatcc 1850

caaggcccgc cttgctttcc agaacatgaa cgggtccgaa tacttcgtga 1900

agatccagtc cttcctgggt ggggcccccca cggaggattt gaaggcgctc 1950

agtcagcaga atgtgagcat ggacttggcc acgttcatga agctgcggac 2000

ggatgcggtg ctgccgttga ctgtggctga ggtgcagaaa cttctgggac 2050

cccacgtgga gggcctgaag gcggaggagc ggcaccgccc ggtgcgggac 2100

tggatcctac ggcagcggca ggacgacctg gacacgctgg ggctggggct 2150

acagggcggc atccccaacg gctacctggt cctagacctc agcgtgcaag 2200

aggccctctc ggggacgccc tgcctcctag gacctggacc tgttctcacc 2250

gtcctggcac tgctcctagc ctccaccctg gcctgagggc cccactccct 2300

tgctggcccc agccctgctg gggatccccg cctggccagg agcaggcacg 2350

ggtgatcccc gttccacccc aagagaactc gcgctcagta aacgggaaca 2400

tgccccctgc agacacgtaa aaaaaaaaaa aaaaaa 2436

<210> 39
<211> 1979
<212> DNA
<213> Homo sapiens

<400> 39
tgggatctac acagaccatg gccttgccaa cggctcgacc cctgttgggg 50

tcctgtggga cccccgccct cggcagcctc ctgttcctgc tcttcagcct 100

cggatgggtg cagccctcga ggaccctggc tggagagaca gggcaggagg 150

ctgcgccccct ggacggagtc ctggccaacc cacctaacat ttccagcctc 200

tcccctcgcc aactccttgg cttcccgtgt gcggaggtgt ccggcctgag 250

cacggagcgt gtccgggagc tggctgtggc cttggcacag aagaatgtca 300

agctctcaac agagcagctg cgctgtctgg ctcaccggct ctctgagccc 350

cccgaggacc tggacgccct cccattggac ctgctgctat tcctcaaccc 400

agatgcgttc tcggggcccc aggcctgcac ccgtttcttc tcccgcatca 450

1383
1384

cgaaggccaa tgtggacctg ctcccgaggg gggctcccga gcgacagcgg 500

ctgctgcctg cggctctggc ctgctggggt gtgcgggggt ctctgctgag 550

cgaggctgat gtgcgggctc tgggaggcct ggcttgcgac ctgcctgggc 600

gctttgtggc cgagtcggcc gaagtgctgc taccccggct ggtgagctgc 650

ccgggacccc tggaccagga ccagcaggag gcagccaggg cggctctgca 700

gggcgggggа ccccctacg gccccccgtc gacatggtct gtctccacga 750

tggacgctct gcggggcctg ctgcccgtgc tgggccagcc catcatccgc 800

agcatcccgc agggcatcgt ggccgcgtgg cggcaacgct cctctcggga 850

cccatcctgg cggcagcctg aacggaccat cctccggccg cggttccggc 900

gggaagtgga gaagacagcc tgtccttcag gcaagaaggc ccgcgagata 950

gacgagagcc tcatcttcta caagaagtgg gagctggaag cctgcgtgga 1000

tgcggccctg ctggccaccc agatggaccg cgtgaacgcc atccccttca 1050

cctacgagca gctggacgtc ctaaagcata aactggatga gctctaccca 1100

caaggttacc ccgagtctgt gatccagcac ctgggctacc tcttcctcaa 1150

gatgagccct gaggacattc gcaagtggaa tgtgacgtcc ctggagaccc 1200

tgaaggcttt gcttgaagtc aacaaagggc acgaaatgag tcctcaggtg 1250

gccaccctga tcgaccgctt tgtgaaggga aggggccagc tagacaaaga 1300

caccctagac accctgaccg ccttctaccc tgggtacctg tgctccctca 1350

gccccgagga gctgagctcc gtgcccccca gcagcatctg ggcggtcagg 1400

ccccaggacc tggacacgtg tgacccaagg cagctggacg tcctctatcc 1450

caaggcccgc cttgctttcc agaacatgaa cgggtccgaa tacttcgtga 1500

agatccagtc cttcctgggt ggggcccccа cggaggattt gaaggcgctc 1550

agtcagcaga atgtgagcat ggacttggcc acgttcatga agctgcggac 1600

ggatgcggtg ctgccgttga ctgtggctga ggtgcagaaa cttctgggac 1650

cccacgtgga gggcctgaag gcggaggagc ggcaccgccc ggtgcgggac 1700

tggatcctac ggcagcggca ggacgacctg gacacgctgg ggctggggct 1750

acagggcggc atccccaacg gctacctggt cctagacctc agcgtgcaag 1800

aggccctctc ggggacgccc tgcctcctag gacctggacc tgttctcacc 1850

gtcctggcac tgctcctagc ctccaccctg gcctgagggc cccactccct 1900

tgctggcccc agccctgctg gggatccccg cctggccagg agcaggcacg 1950

ggtgatcccc gttccacccc aagagaact 1979

<210> 40
<211> 2456
<212> DNA
<213> Homo sapiens

<400> 40
ataacagcat gaagtgccgt ggaactggaa taggcgtgtc ctctccctcg 50

accctccccc tccttgtccc tctgctcacc cctcgctcgt tccctccctc 100

cggcgagggc cgcctttata acaactgctc agagtgcgag ggcgggatag 150

ctgtccaagg tctcccccag cactgaggag ctcgcctgct gccctcttgc 200

~~1384~~
1385

gcgcgggaag cagcaccaag ttcacggcca acgccttggc actagggtcc 250

agaatggcta caacagtccc tgatggttgc cgcaatggcc tgaaatccaa 300

gtactacaga ctttgtgata aggctgaagc ttggggcatc gtcctagaaa 350

cggtggccac agccggggtt gtgacctcgg tggccttcat gctcactctc 400

ccgatcctcg tctgcaaggt gcaggactcc aacaggcgaa aaatgctgcc 450

tactcagttt ctcttcctcc tgggtgtgtt gggcatcttt ggcctcacct 500

tcgccttcat catcggactg gacgggagca cagggcccac acgcttcttc 550

ctctttggga tcctcttttc catctgcttc tcctgcctgc tggctcatgc 600

tgtcagtctg accaagctcg tccgggggag gaagcccctt tccctgttgg 650

tgattctggg tctggccgtg ggcttcagcc tagtccagga tgttatcgct 700

attgaatata ttgtcctgac catgaatagg accaacgtca atgtcttttc 750

tgagctttcc gctcctcgtc gcaatgaaga ctttgtcctc ctgctcacct 800

acgtcctctt cttgatggcg ctgaccttcc tcatgtcctc cttcaccttc 850

tgtggttcct tcacgggctg gaagagacat ggggcccaca tctacctcac 900

gatgctcctc tccattgcca tctgggtggc ctggatcacc ctgctcatgc 950

ttcctgactt tgaccgcagg tgggatgaca ccatcctcag ctccgccttg 1000

gctgccaatg gctgggtgtt cctgttggct tatgttagtc ccgagttttg 1050

gctgctcaca aagcaacgaa accccatgga ttatcctgtt gaggatgctt 1100

tctgtaaacc tcaactcgtg aagaagagct atggtgtgga gaacagagcc 1150

tactctcaag aggaaatcac tcaaggtttt gaagagacag gggacacgct 1200

ctatgccccc tattccacac attttcagct gcagaaccag cctccccaaa 1250

aggaattctc catcccacgg gcccacgctt ggccgagccc ttacaaagac 1300

tatgaagtaa agaaagaggg cagctaactc tgtcctgaag agtgggacaa 1350

atgcagccgg gcggcagatc tagcgggagc tcaaagggat gtgggcgaaa 1400

tcttgagtct tctgagaaaa ctgtacaaga cactacggga acagtttgcc 1450

tccctcccag cctcaaccac aattcttcca tgctggggct gatgtgggct 1500

agtaagactc cagttcttag aggcgctgta gtatttttt ttttttgtct 1550

catcctttgg atacttcttt taagtgggag tctcaggcaa ctcaagttta 1600

gacccttact ctttttgttt gtttttgaa acaggatctt gctctgtcac 1650

ccaggcttga gtgcagtggt gcgatcacag cccagtgcag cctcgaccac 1700

ctgtgctcaa gcaatcctcc catctccatc tcccaaagtg ctgggatgac 1750

aggcgtgagc cacagctccc agcctaggcc cttaatcttg ctgttatttt 1800

ccatggacta aaggtctggt catctgagct cacgctggct cacacagctc 1850

taggggcctg ctcctctaac tcacagtggg ttttgtgagg ctctgtggcc 1900

cagagcagac ctgcatatct gagcaaaaat agcaaaagcc tctctcagcc 1950

cactggcctg aatctacact ggaagccaac ttgctggcac ccccgctccc 2000

caacccttct tgcctgggta ggagaggcta aagatcaccc taaatttact 2050

catctctcta gtgctgcctc acattgggcc tcagcagctc cccagcacca 2100

1385
1386

attcacaggt caccctctc ttcttgcact gtccccaaac ttgctgtcaa 2150

ttccgagatc taatctcccc ctacgctctg ccaggaattc tttcagacct 2200

cactagcaca agcccggttg ctccttgtca ggagaatttg tagatcattc 2250

tcacttcaaa ttcctggggc tgatacttct ctcatcttgc accccaacct 2300

ctgtaaatag atttaccgca tttacggctg cattctgtaa gtgggcatgg 2350

tctcctaatg gaggagtgtt cattgtataa taagttattc acctgagtat 2400

gcaataaaga tgtggtggcc actctttcat ggtggtggca gcaaaaaaaa 2450

aaaaaa 2456

<210> 41
<211> 3213
<212> DNA
<213> Homo sapiens

<400> 41
agagactcaa gatgattccc tttttaccca tgttttctct actattgctg 50

cttattgtta accctataaa cgccaacaat cattatgaca agatcttggc 100

tcatagtcgt atcaggggtc gggaccaagg cccaaatgtc tgtgcccttc 150

aacagatttt gggcaccaaa aagaaatact tcagcacttg taagaactgg 200

tataaaaagt ccatctgtgg acagaaaacg actgttttat atgaatgttg 250

ccctggttat atgagaatgg aaggaatgaa aggctgccca gcagttttgc 300

ccattgacca tgtttatggc actctgggca tcgtgggagc caccacaacg 350

cagcgctatt ctgacgcctc aaaactgagg gaggagatcg agggaaaggg 400

atccttcact tactttgcac cgagtaatga ggcttgggac aacttggatt 450

ctgatatccg tagaggtttg gagagcaacg tgaatgttga attactgaat 500

gctttacata gtcacatgat taataagaga atgttgacca aggacttaaa 550

aaatggcatg attattcctt caatgtataa caatttgggg cttttcatta 600

accattatcc taatggggtt gtcactgtta attgtgctcg aatcatccat 650

gggaaccaga ttgcaacaaa tggtgttgtc catgtcattg accgtgtgct 700

tacacaaatt ggtacctcaa ttcaagactt cattgaagca gaagatgacc 750

tttcatcttt tagagcagct gccatcacat cggacatatt ggaggccctt 800

ggaagagacg gtcacttcac actctttgct cccaccaatg aggcttttga 850

gaaacttcca cgaggtgtcc tagaaaggtt catgggagac aaagtggctt 900

ccgaagctct tatgaagtac cacatcttaa atactctcca gtgttctgag 950

tctattatgg gaggagcagt ctttgagacg ctggaaggaa atacaattga 1000

gataggatgt gacggtgaca gtataacagt aaatggaatc aaaatggtga 1050

acaaaaagga tattgtgaca aataatggtg tgatccattt gattgatcag 1100

gtcctaattc ctgattctgc caaacaagtt attgagctgg ctggaaaaca 1150

gcaaaccacc ttcacggatc ttgtggccca attaggcttg gcatctgctc 1200

tgaggccaga tggagaatac actttgctgg cacctgtgaa taatgcattt 1250

tctgatgata ctctcagcat ggttcagcgc ctccttaaat taattctgca 1300

gaatcacata ttgaaagtaa aagttggcct taatgagctt tacaacgggc 1350

1386
1387

aaatactgga aaccatcgga ggcaaacagc tcagagtctt cgtatatcgt 1400

acagctgtct gcattgaaaa ttcatgcatg gagaaaggga gtaagcaagg 1450

gagaaacggt gcgattcaca tattccgcga gatcatcaag ccagcagaga 1500

aatccctcca tgaaaagtta aaacaagata agcgctttag caccttcctc 1550

agcctacttg aagctgcaga cttgaaagag ctcctgacac aacctggaga 1600

ctggacatta tttgtgccaa ccaatgatgc ttttaaggga atgactagtg 1650

aagaaaaaga aattctgata cgggacaaaa atgctcttca aaacatcatt 1700

ctttatcacc tgacaccagg agttttcatt ggaaaaggat ttgaacctgg 1750

tgttactaac attttaaaga ccacacaagg aagcaaaatc tttctgaaag 1800

aagtaaatga tacacttctg gtgaatgaat tgaaatcaaa agaatctgac 1850

atcatgacaa caaatggtgt aattcatgtt gtagataaac tcctctatcc 1900

agcagacaca cctgttggaa atgatcaact gctggaaata cttaataaat 1950

taatcaaata catccaaatt aagtttgttc gtggtagcac cttcaaagaa 2000

atccccgtga ctgtctatac aactaaaatt ataaccaaag ttgtggaacc 2050

aaaaattaaa gtgattgaag gcagtcttca gcctattatc aaaactgaag 2100

gacccacact aacaaaagtc aaaattgaag gtgaacctga attcagactg 2150

attaaagaag gtgaaacaat aactgaagtg atccatggag agccaattat 2200

taaaaaatac accaaaatca ttgatggagt gcctgtggaa ataactgaaa 2250

aagagacacg agaagaacga atcattacag gtcctgaaat aaaatacact 2300

aggatttcta ctggaggtgg agaaacagaa gaaactctga agaaattgtt 2350

acaagaagag gtcaccaagg tcaccaaatt cattgaaggt ggtgatggtc 2400

atttatttga agatgaagaa attaaaagac tgcttcaggg agacacaccc 2450

gtgaggaagt tgcaagccaa caaaaaagtt caaggttcta gaagacgatt 2500

aagggaaggt cgttctcagt gaaaatccaa aaaccagaaa aaaatgttta 2550

tacaacccta agtcaataac ctgaccttag aaaattgtga gagccaagtt 2600

gacttcagga actgaaacat cagcacaaag aagcaatcat caaataattc 2650

tgaacacaaa tttaatattt ttttttctga atgagaaaca tgagggaaat 2700

tgtggagtta gcctcctgtg gtaaaggaat tgaagaaaat ataacacctt 2750

acacccttt tcatcttgac attaaaagtt ctggctaact ttggaatcca 2800

ttagagaaaa atccttgtca ccagattcat tacaattcaa atcgaagagt 2850

tgtgaactgt tatcccattg aaaagaccga gccttgtatg tatgttatgg 2900

atacataaaa tgcacgcaag ccattatctc tccatgggaa gctaagttat 2950

aaaaataggt gcttggtgta caaaactttt tatatcaaaa ggctttgcac 3000

atttctatat gagtgggttt actggtaaat tatgttattt tttacaacta 3050

attttgtact ctcagaatgt ttgtcatatg cttcttgcaa tgcatatttt 3100

ttaatctcaa acgtttcaat aaaaccattt ttcagatata aagagaatta 3150

cttcaaattg agtaattcag aaaaactcaa gatttaagtt aaaaagtggt 3200

ttggacttgg gaa 3213

<210> 42
<211> 2450
<212> DNA
<213> Homo sapiens

<400> 42
ccettgagac tcaagatgat tccctttta cccatgtttt ctctactatt 50

gctgcttatt gttaacccta taaacgccaa caatcattat gacaagatct 100

tggctcatag tcgtatcagg ggtcgggacc aaggcccaaa tgtctgtgcc 150

cttcaacaga ttttgggcac caaaaagaaa tacttcagca cttgtaagaa 200

ctggtataaa aagtccatct gtggacagaa aacgactgtg ttatatgaat 250

gttgccctgg ttatatgaga atggaaggaa tgaaaggctg cccagcagtt 300

ttgcccattg accatgttta tggcactctg ggcatcgtgg gagccaccac 350

aacgcagcgc tattctgacg cctcaaaact gagggaggag atcgagggaa 400

agggatcctt cacttacttt gcaccgagta atgaggcttg ggacaacttg 450

gattctgata tccgtagagg tttggagagc aacgtgaatg ttgaattact 500

gaatgcttta catagtcaca tgattaataa gagaatgttg accaaggact 550

taaaaaatgg catgattatt ccttcaatgt ataacaattt ggggcttttc 600

attaaccatt atcctaatgg ggttgtcact gttaattgtg ctcgaatcat 650

ccatgggaac cagattgcaa caaatggtgt tgtccatgtc attgaccgtg 700

tgcttacaca aattggtacc tcaattcaag acttcattga agcagaagat 750

gacctttcat cttttagagc agctgccatc acatcggaca tattggaggc 800

ccttggaaga gacggtcact tcacactctt tgctcccacc aatgaggctt 850

ttgagaaact tccacgaggt gtcctagaaa ggatcatggg agacaaagtg 900

gcttccgaag ctcttatgaa gtaccacatc ttaaatactc tccagtgttc 950

tgagtctatt atgggaggag cagtctttga gacgctggaa ggaaatacaa 1000

ttgagatagg atgtgacggt gacagtataa cagtaaatgg aatcaaaatg 1050

gtgaacaaaa aggatattgt gacaaataat ggtgtgatcc atttgattga 1100

tcaggtccta attcctgatt ctgccaaaca agttattgag ctggctggaa 1150

aacagcaaac caccttcacg gatcttgtgg cccaattagg cttggcatct 1200

gctctgaggc cagatggaga atacactttg ctggcacctg tgaataatgc 1250

attttctgat gatactctca gcatggatca gcgcctcctt aaattaattc 1300

tgcagaatca catattgaaa gtaaaagttg gccttaatga gctttacaac 1350

gggcaaatac tggaaaccat cggaggcaaa cagctcagag tcttcgtata 1400

tcgtacagct gtctgcattg aaaattcatg catggagaaa gggagtaagc 1450

aagggagaaa cggtgcgatt cacatattcc gcgagatcat caagccagca 1500

gagaaatccc tccatgaaaa gttaaaacaa gataagcgct ttagcacctt 1550

cctcagccta cttgaagctg cagacttgaa agagctcctg acacaacctg 1600

gagactggac attatttgtg ccaaccaatg atgcttttaa gggaatgact 1650

agtgaagaaa aagaaattct gatacgggac aaaaatgctc ttcaaaacat 1700

cattctttat cacctgacac caggagtttt cattggaaaa ggatttgaac 1750

ctggtgttac taacatttta aagaccacac aaggaagcaa aatctttctg 1800

aaagaagtaa atgatacact tctggtgaat gaattgaaat caaaagaatc 1850

tgacatcatg acaacaaatg gtgtaattca tgttgtagat aaactcctct 1900

atccagcaga cacacctgtt ggaaatgatc aactgctgga aatacttaat 1950

aaattaatca aatacatcca aattaagttt gttcgtggta gcaccttcaa 2000

agaaatcccc gtgactgtct atagacccac actaacaaaa gtcaaaattg 2050

aaggtgaacc tgaattcaga ctgattaaag aaggtgaaac aataactgaa 2100

gtgatccatg gagagccaat tattaaaaaa tacaccaaaa tcattgatgg 2150

agtgcctgtg gaaataactg aaaaagagac acgagaagaa cgaatcatta 2200

caggtcctga aataaaatac actaggattt ctactggagg tggagaaaca 2250

gaagaaactc tgaagaaatt gttacaagaa gaggtcacca aggtcaccaa 2300

attcattgaa ggtggtgatg gtcatttatt tgaagatgaa gaaattaaaa 2350

gactgcttca gggagacaca cccgtgagga agttgcaagc caacaaaaaa 2400

gttcaaggat ctagaagacg attaagggaa ggtcgttctc agtgaagggc 2450

<210> 43
<211> 2276
<212> DNA
<213> Homo sapiens

<400> 43
cccttgagac tcaagatgat tccctttta cccatgtttt ctctactatt 50

gctgcttatt gttaacccta taaacgccaa caatcattat gacaagatct 100

tggctcatag tcgtatcagg ggtcgggacc aaggcccaaa tgtctgtgcc 150

cttcaacaga ttttgggcac caaaaagaaa tacttcagca cttgtaagaa 200

ctggtataaa aagtccatct gtggacagaa aacgactgtg ttatatgaat 250

gttgccctgg ttatatgaga atggaaggaa tgaaaggctg cccagcagtt 300

ttgcccattg accatgttta tggcactctg ggcatcgtgg gagccaccac 350

aacgcagcgc tattctgacg cctcaaaact gagggaggag atcgagggaa 400

agggatcctt cacttacttt gcaccgagta atgaggcttg ggacaacttg 450

gattctgata tccgtagagg tttggagagc aacgtgaatg ttgaattact 500

gaatgcttta catagtcaca tgattaataa gagaatgttg accaaggact 550

taaaaaatgg catgattatt ccttcaatgt ataacaattt ggggcttttc 600

attaaccatt atcctaatgg ggttgtcact gttaattgtg ctcgaatcat 650

ccatgggaac cagattgcaa caaatggtgt tgtccatgtc attgaccgtg 700

tgcttacaca aattggtacc tcaattcaag acttcattga agcagaagat 750

gacctttcat cttttagagc agctgccatc acatcggaca tattggaggc 800

ccttggaaga gacggtcact tcacactctt tgctcccacc aatgaggctt 850

ttgagaaact tccacgaggt gtcctagaaa ggatcatggg agacaaagtg 900

gcttccgaag ctcttatgaa gtaccacatc ttaaatactc tccagtgttc 950

tgagtctatt atgggaggag cagtctttga gacgctggaa ggaaatacaa 1000

ttgagatagg atgtgacggt gacagtataa cagtaaatgg aatcaaaatg 1050

1389
1390

gtgaacaaaa aggatattgt gacaaataat ggtgtgatcc atttgattga 1100

tcaggtccta attcctgatt ctgccaaaca agttattgag ctggctggaa 1150

aacagcaaac caccttcacg gatcttgtgg cccaattagg cttggcatct 1200

gctctgaggc cagatggaga atacactttg ctggcacctg tgaataatgc 1250

attttctgat gatactctca gcatggatca gcgcctcctt aaattaattc 1300

tgcagaatca catattgaaa gtaaaagttg gccttaatga gctttacaac 1350

gggcaaatac tggaaaccat cggaggcaaa cagctcagag tcttcgtata 1400

tcgtacagct gtctgcattg aaaattcatg catggagaaa gggagtaagc 1450

aagggagaaa cggtgcgatt cacatattcc gcgagatcat caagccagca 1500

gagaaatccc tccatgaaaa gttaaaacaa gataagcgct ttagcacctt 1550

cctcagccta cttgaagctg cagacttgaa agagctcctg acacaacctg 1600

gagactggac attatttgtg ccaaccaatg atgcttttaa gggaatgact 1650

agtgaagaaa aagaaattct gatacgggac aaaaatgctc ttcaaaacat 1700

cattctttat cacctgacac caggagtttt cattggaaaa ggatttgaac 1750

ctggtgttac taacatttta aagaccacac aaggaagcaa aatctttctg 1800

aaagaagtaa atgatacact tctggtgaat gaattgaaat caaaagaatc 1850

tgacatcatg acaacaaatg gtgtaattca tgttgtagat aaactcctct 1900

atccagcaga cacacctgtt ggaaatgatc aactgctgga aatacttaat 1950

aaattaatca aatacatcca aattaagttt gttcgtggta gcaccttcaa 2000

agaaatcccc gtgactgtct ataagccaat tattaaaaaa tacaccaaaa 2050

tcattgatgg agtgcctgtg gaaataactg aaaaagagac acgagaagaa 2100

cgaatcatta caggtcctga aataaaatac actaggattt ctactggagg 2150

tggagaaaca gaagaaactc tgaagaaatt gttacaagaa gacacacccg 2200

tgaggaagtt gcaagccaac aaaaaagttc aaggatctag aagacgatta 2250

agggaaggtc gttctcagtg aagggc 2276

<210> 44
<211> 2366
<212> DNA
<213> Homo sapiens

<400> 44
cccttgagac tcaagatgat tccctttta cccatgtttt ctctactatt 50

gctgcttatt gttaacccta taaacgccaa caatcattat gacaagatct 100

tggctcatag tcgtatcagg ggtcgggacc aaggcccaaa tgtctgtgcc 150

cttcaacaga ttttgggcac caaaaagaaa tacttcagca cttgtaagaa 200

ctggtataaa aagtccatct gtggacagaa aacgactgtg ttatatgaat 250

gttgccctgg ttatatgaga atggaaggaa tgaaaggctg cccagcagtt 300

ttgcccattg accatgttta tggcactctg ggcatcgtgg gagccaccac 350

aacgcagcgc tattctgacg cctcaaaact gagggaggag atcgagggaa 400

agggatcctt cacttacttt gcaccgagta atgaggcttg ggacaacttg 450

gattctgata tccgtagagg tttggagagc aacgtgaatg ttgaattact 500

gaatgcttta catagtcaca tgattaataa gagaatgttg accaaggact 550

taaaaaatgg catgattatt ccttcaatgt ataacaattt ggggcttttc 600

attaaccatt atcctaatgg ggttgtcact gttaattgtg ctcgaatcat 650

ccatgggaac cagattgcaa caaatggtgt tgtccatgtc attgaccgtg 700

tgcttacaca aattggtacc tcaattcaag acttcattga agcagaagat 750

gacctttcat cttttagagc agctgccatc acatcggaca tattggaggc 800

ccttggaaga gacggtcact tcacactctt tgctcccacc aatgaggctt 850

ttgagaaact tccacgaggt gtcctagaaa ggatcatggg agacaaagtg 900

gcttccgaag ctcttatgaa gtaccacatc ttaaatactc tccagtgttc 950

tgagtctatt atgggaggag cagtctttga gacgctggaa ggaaatacaa 1000

ttgagatagg atgtgacggt gacagtataa cagtaaatgg aatcaaaatg 1050

gtgaacaaaa aggatattgt gacaaataat ggtgtgatcc atttgattga 1100

tcaggtccta attcctgatt ctgccaaaca agttattgag ctggctggaa 1150

aacagcaaac caccttcacg gatcttgtgg cccaattagg cttggcatct 1200

gctctgaggc cagatggaga atacactttg ctggcacctg tgaataatgc 1250

attttctgat gatactctca gcatggatca gcgcctcctt aaattaattc 1300

tgcagaatca catattgaaa gtaaaagttg gccttaatga gctttacaac 1350

gggcaaatac tggaaaccat cggaggcaaa cagctcagag tcttcgtata 1400

tcgtacagct gtctgcattg aaaattcatg catggagaaa gggagtaagc 1450

aagggagaaa cggtgcgatt cacatattcc gcgagatcat caagccagca 1500

gagaaatccc tccatgaaaa gttaaaacaa gataagcgct ttagcacctt 1550

cctcagccta cttgaagctg cagacttgaa agagctcctg acacaacctg 1600

gagactggac attatttgtg ccaaccaatg atgcttttaa gggaatgact 1650

agtgaagaaa aagaaattct gatacgggac aaaaatgctc ttcaaaacat 1700

cattctttat cacctgacac caggagtttt cattggaaaa ggatttgaac 1750

ctggtgttac taacatttta aagaccacac aaggaagcaa aatctttctg 1800

aaagaagtaa atgatacact tctggtgaat gaattgaaat caaaagaatc 1850

tgacatcatg acaacaaatg gtgtaattca tgttgtagat aaactcctct 1900

atccagcaga cacacctgtt ggaaatgatc aactgctgga aatacttaat 1950

aaattaatca aatacatcca aattaagttt gttcgtggta gcaccttcaa 2000

agaaatcccc gtgactgtct atagacccac actaacaaaa gtcaaaattg 2050

aaggtgaacc tgaattcaga ctgattaaag aaggtgaaac aataactgaa 2100

gtgatccatg gagagccaat tattaaaaaa tacaccaaaa tcattgatgg 2150

agtgcctgtg gaaataactg aaaaagagac acgagaagaa cgaatcatta 2200

caggtcctga aataaaatac actaggattt ctactggagg tggagaaaca 2250

gaagaaactc tgaagaaatt gttacaagaa gacacacccg tgaggaagtt 2300

gcaagccaac aaaaaagttc aaggatctag aagacgatta agggaaggtc 2350

gttctcagtg aagggc 2366

<210> 45
<211> 2360
<212> DNA
<213> Homo sapiens

<400> 45
ccccttgagac tcaagatgat tccctttta cccatgtttt ctctactatt 50

gctgcttatt gttaacccta taaacgccaa caatcattat gacaagatct 100

tggctcatag tcgtatcagg ggtcgggacc aaggcccaaa tgtctgtgcc 150

cttcaacaga ttttgggcac caaaaagaaa tacttcagca cttgtaagaa 200

ctggtataaa aagtccatct gtggacagaa aacgactgtg ttatatgaat 250

gttgccctgg ttatatgaga atggaaggaa tgaaaggctg cccagcagtt 300

ttgcccattg accatgttta tggcactctg ggcatcgtgg gagccaccac 350

aacgcagcgc tattctgacg cctcaaaact gagggaggag atcgagggaa 400

agggatcctt cacttacttt gcaccgagta atgaggcttg ggacaacttg 450

gattctgata tccgtagagg tttggagagc aacgtgaatg ttgaattact 500

gaatgcttta catagtcaca tgattaataa gagaatgttg accaaggact 550

taaaaaatgg catgattatt ccttcaatgt ataacaattt ggggcttttc 600

attaaccatt atcctaatgg ggttgtcact gttaattgtg ctcgaatcat 650

ccatgggaac cagattgcaa caaatggtgt tgtccatgtc attgaccgtg 700

tgcttacaca aattggtacc tcaattcaag acttcattga agcagaagat 750

gacctttcat cttttagagc agctgccatc acatcggaca tattggaggc 800

ccttggaaga gacggtcact tcacactctt tgctcccacc aatgaggctt 850

ttgagaaact tccacgaggt gtcctagaaa ggatcatggg agacaaagtg 900

gcttccgaag ctcttatgaa gtaccacatc ttaaatactc tccagtgttc 950

tgagtctatt atgggaggag cagtctttga gacgctggaa ggaaatacaa 1000

ttgagatagg atgtgacggt gacagtataa cagtaaatgg aatcaaaatg 1050

gtgaacaaaa aggatattgt gacaaataat ggtgtgatcc atttgattga 1100

tcaggtccta attcctgatt ctgccaaaca agttattgag ctggctggaa 1150

aacagcaaac caccttcacg gatcttgtgg cccaattagg cttggcatct 1200

gctctgaggc cagatggaga atacactttg ctggcacctg tgaataatgc 1250

attttctgat gatactctca gcatggatca gcgcctcctt aaattaattc 1300

tgcagaatca catattgaaa gtaaaagttg gccttaatga gctttacaac 1350

gggcaaatac tggaaaccat cggaggcaaa cagctcagag tcttcgtata 1400

tcgtacagct gtctgcattg aaaattcatg catggagaaa gggagtaagc 1450

aagggagaaa cggtgcgatt cacatattcc gcgagatcat caagccagca 1500

gagaaatccc tccatgaaaa gttaaaacaa gataagcgct ttagcacctt 1550

cctcagccta cttgaagctg cagacttgaa agagctcctg acacaacctg 1600

gagactggac attatttgtg ccaaccaatg atgcttttaa gggaatgact 1650

agtgaagaaa aagaaattct gatacgggac aaaaatgctc ttcaaaacat 1700

cattctttat cacctgacac caggagtttt cattggaaaa ggatttgaac 1750

ctggtgttac taacatttta aagaccacac aaggaagcaa aatctttctg 1800

aaagaagtaa atgatacact tctggtgaat gaattgaaat caaaagaatc 1850

tgacatcatg acaacaaatg gtgtaattca tgttgtagat aaactcctct 1900

atccagcaga cacacctgtt ggaaatgatc aactgctgga aatacttaat 1950

aaattaatca aatacatcca aattaagttt gttcgtggta gcaccttcaa 2000

agaaatcccc gtgactgtct ataagccaat tattaaaaaa tacaccaaaa 2050

tcattgatgg agtgcctgtg gaaataactg aaaaagagac acgagaagaa 2100

cgaatcatta caggtcctga aataaaatac actaggattt ctactggagg 2150

tggagaaaca gaagaaactc tgaagaaatt gttacaagaa gaggtcacca 2200

aggtcaccaa attcattgaa ggtggtgatg gtcatttatt tgaagatgaa 2250

gaaattaaaa gactgcttca gggagacaca cccgtgagga agttgcaagc 2300

caacaaaaaa gttcaaggat ctagaagacg attaagggaa ggtcgttctc 2350

agtgaagggc 2360

<210> 46
<211> 2186
<212> DNA
<213> Homo sapiens

<400> 46
acacggacca aggagtctaa cacgtgcgcg agtcgggggc tcgcacgaaa 50

gccgccgtgg cgcaatgaag gtgaaggccg gcgcgctcgc cggccgaggt 100

gggatcccga ggcctctcca gtccgccgag ggcgcaccac cggcccgtct 150

cgcccgccgc gccggggagg tggagcacga gcgcacgtgt taggacccga 200

aagatggtga actatgcctg ggcagggcga agccagagga aactctggtg 250

gaggtccgta gcggtcctga cgtgcaaatc ggtcgtccga cctgggtata 300

ggggcgggct ccaggcgagg cggtcgacgc tcctgaaaac ttgcgcgcgc 350

gctcgcgcca ctgcgcccgg agcgatgaag atggtcgcgc cctggacgcg 400

gttctactcc aacagctgct gcttgtgctg ccatgtccgc accggcacca 450

tcctgctcgg cgtctggtat ctgatcatca atgctgtggt actgttgatt 500

ttattgagtg ccctggctga tccggatcag tataactttt caagttctga 550

actgggaggt gactttgagt tcatggatga tgccaacatg tgcattgcca 600

ttgcgatttc tcttctcatg atcctgatat gtgctatggc tacttacgga 650

gcgtacaagc aacgcgcagc ctggatcatc ccattcttct gttaccagat 700

ctttgacttt gccctgaaca tgttggttgc aatcactgtg cttatttatc 750

caaactccat tcaggaatac atacggcaac tgcctcctaa ttttccctac 800

agagatgatg tcatgtcagt gaatcctacc tgtttggtcc ttattattct 850

tctgtttatt agcattatct tgacttttaa gggttacttg attagctgtg 900

tttggaactg ctaccgatac atcaatggta ggaactcctc tgatgtcctg 950

gtttatgtta ccagcaatga cactacggtg ctgctacccc cgtatgatga 1000

tgccactgtg aatggtgctg ccaaggagcc accgccacct tacgtgtctg 1050

cctaagcctt caagtgggcg gagctgaggg cagcagcttg actttgcaga 1100

catctgagca atagttctgt tatttcactt ttgccatgag cctctctgag 1150

cttgtttgtt gctgaaatgc tacttttaa aatttagatg ttagattgaa 1200

aactgtagtt ttcaacatat gctttgctag aacactgtga tagattaact 1250

gtagaattct tcctgtacga ttggggatat aatgggcttc actaaccttc 1300

cctaggcatt gaaacttccc ccaaatctga tggacctaga agtctgcttt 1350

tgtacctgct gggccccaaa gttgggcatt tttctctctg ttccctctct 1400

tttgaaaatg taaaataaaa ccaaaaatag acaacttttt cttcagccat 1450

tccagcatag agaacaaaac cttatggaaa caggaatgtc aattgtgtaa 1500

tcattgttct aattaggtaa atagaagtcc ttatgtatgt gttacaagaa 1550

tttcccccac aacatccttt atgactgaag ttcaatgaca gtttgtgttt 1600

gggtggtaaa ggattttctc catggcctga attaagacca ttagaaagca 1650

ccaggccgtg ggagcagtga ccatctgctg actgttcttg tggatcttgt 1700

gtccagggac atggggtgac atgcctcgta tgtgttagag ggtggaatgg 1750

atgtgtttgg cgctgcatgg gatctggtgc ccctcttctc ctggattcac 1800

atccccaccc agggcccgct tttactaagt gttctgccct agattggttc 1850

aaggaggtca tccaactgac tttatcaagt ggaattggga tatatttgat 1900

atacttctgc ctaacaacat ggaaaagggt tttcttttcc ctgcaagcta 1950

catcctactg ctttgaactt ccaagtatgt ctagtcacct tttaaaatgt 2000

aaacattttc agaaaaatga ggattgcctt ccttgtatgc gctttttacc 2050

ttgactacct gaattgcaag ggatttttat atattcatat gttacaaagt 2100

cagcaactct cctgttggtt cattattgaa tgtgctgtaa attaagttgt 2150

ttgcaattaa aacaaggttt gcccacaaaa aaaaaa 2186

<210> 47
<211> 851
<212> DNA
<213> Homo sapiens

<400> 47
attcgcccct ccgaaagatg gtgaactatg cctgggcagg gcgggctcca 50

ggcgaggcgg tcgacgctcc tgaaaacttg cgcgcgcgct cgcgccactg 100

cgcccggagc gatgaagatg gtcgcgccct ggactcggtt ctactccaac 150

agctgctgct tgtgctgcca tgtccgcacc ggcaccatcc tgctcggcgt 200

ctggtatctg atcatcaatg ctgtggtact gttgatttta ttgagtgccc 250

tggctgatcc ggatcagtat aacttttcaa gttctgaact gggaggtgac 300

tttgagttca tggatgatgc caacatgtgc attgccattg cgatttctct 350

tctcatgatc ctgatatgtg ctatggctac ttacggagcg tacaagcaac 400

gcgcagcctg gatcatccca ttcttctgtt accagatctt tgactttgcc 450

ctgaacatgt tggttgcaat cactgtgctt atttatccaa actccattca 500

ggaatacata cggcaactgc ctcctaattt tccctacaga gatgatgtca 550

tgtcagtgaa tcctacctgt ttggtcctta ttattcttct gtttattagc 600

attatcttga cttttaaggg ttacttgatt agctgtgttt ggaactgcta 650

ccgatacatc aatggtagga actcctctga tgtcctggtt tatgttacca 700

gcaatgacac tacggtgctg ctacccccgt atgatgatgc cactgtgaat 750

ggtgctgcca aggagccacc gccaccttac gtgtctgcct aagccttcaa 800

gtgggcggag ctgagggcag cagcttgact ttgcagacat ctgagaaggg 850

c 851

<210> 48
<211> 1728
<212> DNA
<213> Homo sapiens

<400> 48
agcgccatgg ggcgagccgg gggcgggggc ccgggccggg ggccgccgcc 50

actgctgctg tttctggggg ccgcgctggt cctggcctct ggggccgtgc 100

cggcgcgtga ggcgggcagc gcggttgagg ccgaagagct ggtgaagggc 150

agcccggcgt gggagccgcc tgccaacgac acgcgggaag aagccggccc 200

accagcggct ggggaagatg aggcgtcgtg gacggcgccc ggtggcgagc 250

tggccgggcc agaagaggtg ctgcaggagt cggctgcggt gaccggcacc 300

gcctggctgg aagctgacag cccaggcctg ggaggagtga ccgcagaggc 350

gggcagcggc gatgcccagg cccttccagc tacgctccag gctccccacg 400

aggtcctcgg gcagtcaatc atgccccctg ccattcctga ggctacagag 450

gccagcgggc caccctcccc caccccggc gacaagctga gcccagcttc 500

tgaactcccc aaggagagcc ccttggaggt ttggctgaac ctgggggggca 550

gcacacccga ccctcaagtg ccagagctga cttacccatt tcagggcacc 600

ctggagcccc aaccggcatc agatatcatt gacatcgact acttcgaagg 650

actggatggt gagggtcgtg gcgcagatct ggggagcttc ccagggtcac 700

caggaacctc agagaaccac cctgatactg agggagagac cccttcctgg 750

agcctgcttg acttatacga tgatttcacc cccttcgatg aatctgattt 800

ctaccccacc acatcctttt atgatgactt ggatgaagag gaggaggaag 850

aggaggatga caaagatgca gtaggaggtg gagacctaga agatgaaaat 900

gagcttctag tgcccactgg gaagcctggt ctgggggcccg ggacaggcca 950

gcccaccagt cggtggcatg ctgtccctcc acagcacact ctggggtcgg 1000

tccccggcag cagcatcgcc ctcaggcccc gcccaggaga gccaggcagg 1050

gacttggcct ccagtgaaaa tggcactgag tgccgcagtg gctttgtgcg 1100

gcataacggc tcctgccggt cagtgtgcga cctcttccca agttactgtc 1150

acaatggcgg ccagtgctac ctggtggaga acataggggc cttctgcagg 1200

tgcaacacgc aggactacat ctggcacaag gggatgcgct gcgagtccat 1250

catcaccgac ttccaggtga tgtgcgtggc cgtgggctcg gctgccctcg 1300

tcctgctcct gctcttcatg atgacggtgt tctttgccaa gaagctctac 1350

ctgctcaaga cggagaatac caagctgcgt aggaccaaca aattccggac 1400

cccatctgag ctccacaatg ataacttctc cctctccacc attgccgagg 1450

gctctcaccc aaatgatgat cctagtgctc cccacaaaat ccaggaggtt 1500

ctcaagtcct gcctgaaaga ggaggagtca tttaacatcc agaactccat 1550

gtcgcccaaa cttgagggtg gcaaaggtga ccaggctgac ttggatgtga 1600

actgtcttca gaataattta acctaaagca gagcaagaag agaggaagcg 1650

gggtagtggg tgggggtagg gaagaaacat tatctcctct tgtacagagt 1700

ctatttcttg taaccatttg ttaaactc 1728

<210> 49
<211> 4231
<212> DNA
<213> Homo sapiens

<400> 49
ggacaggcgt ggcggccgga gccccagcat ccctgcttga ggtccaggag 50

cggagcccgc ggccaccgcc gcctgatcag cgcgaccccg gcccgcgccc 100

gccccgcccg gcaagatgct gcccgtgtac caggaggtga agcccaaccc 150

gctgcaggac gcgaacatct gctcacgcgt gttcttctgg tggctcaatc 200

ccttgtttaa aattggccat aaacggagat tagaggaaga tgatatgtat 250

tcagtgctgc cagaagaccg ctcacagcac cttggagagg agttgcaagg 300

gttctgggat aaagaagttt taagagctga gaatgacgca cagaagcctt 350

ctttaacaag agcaatcata aagtgttact ggaaatctta tttagttttg 400

ggaattttta cgttaattga ggaaagtgcc aaagtaatcc agcccatatt 450

tttgggaaaa attattaatt attttgaaaa ttatgatccc atggattctg 500

tggctttgaa cacagcgtac gcctatgcca cggtgctgac tttttgcacg 550

ctcattttgg ctatactgca tcacttatat ttttatcacg ttcagtgtgc 600

tgggatgagg ttacgagtag ccatgtgcca tatgatttat cggaaggcac 650

ttcgtcttag taacatggcc atggggaaga caaccacagg ccagatagtc 700

aatctgctgt ccaatgatgt gaacaagttt gatcaggtga cagtgttctt 750

acacttcctg tgggcaggac cactgcaggc gatcgcagtg actgccctac 800

tctggatgga gataggaata tcgtgccttg ctgggatggc agttctaatc 850

attctcctgc ccttgcaaag ctgttttggg aagttgttct catcactgag 900

gagtaaaact gcaactttca cggatgccag gatcaggacc atgaatgaag 950

ttataactgg tataaggata ataaaaatgt acgcctggga aaagtcattt 1000

tcaaatctta ttaccaattt gagaaagaag gagatttcca agattctgag 1050

aagttcctgc ctcaggggga tgaatttggc ttcgtttttc agtgcaagca 1100

aaatcatcgt gtttgtgacc ttcaccacct acgtgctcct cggcagtgtg 1150

atcacagcca gccgcgtgtt cgtggcagtg acgctgtatg gggctgtgcg 1200

gctgacggtt accctcttct tcccctcagc cattgagagg gtgtcagagg 1250

caatcgtcag catccgaaga atccagacct ttttgctact tgatgagata 1300

tcacagcgca accgtcagct gccgtcagat ggtaaaaaga tggtgcatgt 1350

gcaggatttt actgcttttt gggataaggc atcagagacc ccaactctac 1400

aaggcctttc ctttactgtc agacctggcg aattgttagc tgtggtcggc 1450

cccgtgggag cagggaagtc atcactgtta agtgccgtgc tcggggaatt 1500

ggccccaagt cacgggctgg tcagcgtgca tggaagaatt gcctatgtgt 1550

ctcagcagcc ctgggtgttc tcgggaactc tgaggagtaa tattttattt 1600

gggaagaaat atgaaaagga acgatatgaa aaagtcataa aggcttgtgc 1650

tctgaaaaag gatttacagc tgttggagga tggtgatctg actgtgatag 1700

gagatcgggg aaccacgctg agtggagggc agaaagcacg ggtaaacctt 1750

gcaagagcag tgtatcaaga tgctgacatc tatctcctgg acgatcctct 1800

cagtgcagta gatgcggaag ttagcagaca cttgttcgaa ctgtgtattt 1850

gtcaaatttt gcatgagaag atcacaattt tagtgactca tcagttgcag 1900

tacctcaaag ctgcaagtca gattctgata ttgaaagatg gtaaaatggt 1950

gcagaagggg acttacactg agttcctaaa atctggtata gattttggct 2000

ccctttaaa gaaggataat gaggaaagtg aacaacctcc agttccagga 2050

actcccacac taaggaatcg taccttctca gagtcttcgg tttggtctca 2100

acaatcttct agaccctcct tgaaagatgg tgctctggag agccaagata 2150

cagagaatgt cccagttaca ctatcagagg agaaccgttc tgaaggaaaa 2200

gttggttttc aggcctataa gaattacttc agagctggtg ctcactggat 2250

tgtcttcatt ttccttattc tcctaaacac tgcagctcag gttgcctatg 2300

tgcttcaaga ttggtggctt tcatactggg caaacaaaca aagtatgcta 2350

aatgtcactg taaatggagg aggaaatgta accgagaagc tagatcttaa 2400

ctggtactta ggaatttatt caggtttaac tgtagctacc gttctttttg 2450

gcatagcaag atctctattg gtattctacg tccttgttaa ctcttcacaa 2500

actttgcaca acaaaatgtt tgagtcaatt ctgaaagctc cggtattatt 2550

ctttgataga aatccaatag gaagaatttt aaatcgtttc tccaaagaca 2600

ttggacactt ggatgatttg ctgccgctga cgtttttaga tttcatccag 2650

acattgctac aagtggttgg tgtggtctct gtggctgtgg ccgtgattcc 2700

ttggatcgca ataccccttgg ttccccttgg aatcattttc atttttcttc 2750

ggcgatattt tttggaaacg tcaagagatg tgaagcgcct ggaatctaca 2800

actcggagtc cagtgttttc ccacttgtca tcttctctcc aggggctctg 2850

gaccatccgg gcatacaaag cagaagagag gtgtcaggaa ctgtttgatg 2900

cacaccagga tttacattca gaggcttggt tcttgttttt gacaacgtcc 2950

cgctggttcg ccgtccgtct ggatgccatc tgtgccatgt ttgtcatcat 3000

cgttgccttt gggtccctga ttctggcaaa aactctggat gccgggcagg 3050

ttggtttggc actgtcctat gccctcacgc tcatggggat gtttcagtgg 3100

tgtgttcgac aaagtgctga agttgagaat atgatgatct cagtagaaag 3150

ggtcattgaa tacacagacc ttgaaaaaga agcaccttgg gaatatcaga 3200

aacgcccacc accagcctgg ccccatgaag gagtgataat ctttgacaat 3250

gtgaacttca tgtacagtcc aggtgggcct ctggtactga agcatctgac 3300

agcactcatt aaatcacaag aaaaggttgg cattgtggga agaaccggag 3350

ctggaaaaag ttccctcatc tcagcccttt ttagattgtc agaacccgaa 3400

ggtaaaattt ggattgataa gatcttgaca actgaaattg gacttcacga 3450

tttaaggaag aaaatgtcaa tcatacctca ggaacctgtt ttgttcactg 3500

gaacaatgag gaaaaacctg gatcccttta aggagcacac ggatgaggaa 3550

ctgtggaatg ccttacaaga ggtacaactt aaagaaacca ttgaagatct 3600

tcctggtaaa atggatactg aattagcaga atcaggatcc aattttagtg 3650

ttggacaaag acaactggtg tgccttgcca gggcaattct caggaaaaat 3700

cagatattga ttattgatga agcgacggca aatgtggatc caagaactga 3750

tgagttaata caaaaaaaaa tccgggagaa atttgcccac tgcaccgtgc 3800

taaccattgc acacagattg aacaccatta ttgacagcga caagataatg 3850

gttttagatt caggaagact gaaagaatat gatgagccgt atgttttgct 3900

gcaaaataaa gagagcctat tttacaagat ggtgcaacaa ctgggcaagg 3950

cagaagccgc tgccctcact gaaacagcaa aacaggtata cttcaaaaga 4000

aattatccac atattggtca cactgaccac atggttacaa acacttccaa 4050

tggacagccc tcgaccttaa ctattttcga gacagcactg tgaatccaac 4100

caaaatgtca agtccgttcc gaaggcattt tccactagtt tttggactat 4150

gtaaaccaca ttgtactttt ttttactttg gcaacaaata tttatacata 4200

caagatgcta gttcatttga atatttctcc c 4231

<210> 50
<211> 2861
<212> DNA
<213> Homo sapiens

<400> 50
ctccgtccca ggggagaagg agaggcgtct gcaggggggca gagaccgcag 50

ctacctgccg ggtgcgcccc ccacccagga gcgctcgctt cgccccctttt 100

cctcccccccg cccccacctc cttattggtg ctagtttgca gcgcccagct 150

cctgcgcctt cgcttcgcgt ttgaatctgg ctcgcccctt cgtattatgt 200

ctgcactccg aaggaaattt ggggacgatt atcaggtagt gaccacatcg 250

tccagcggct cgggcttgca gccccagggg ccaggccagg accctcagca 300

gcagcttgtg cccaagaaga agcggcagcg gttcgtggac aagaacggcc 350

ggtgcaatgt acagcacggc aacctgggca gcgagacaag ccgctacctc 400

tcggacctct tcaccacgct ggtggacctc aagtggcgct ggaacctctt 450

catcttcatt ctcacctaca ccgtggcctg gcttttcatg gcgtccatgt 500

ggtgggtgat cgcctacact cggggcgacc tgaacaaagc ccacgtcggt 550

aactacacgc cttgcgtggc caatgtctat aacttccctt ctgccttcct 600

cttcttcatc gagacggagg ccaccatcgg ctatggctac cgatacatca 650

cagacaagtg ccccgagggc atcatcctct tcctcttcca gtccatcctg 700

ggctccatcg tggacgcctt cctcatcggc tgcatgttca tcaagatgtc 750

ccagcccaag aagcgcgccg agaccctcat gttcagcgag cacgcggtga 800

tctccatgag ggacggaaaa ctcacgctta tgttccgggt gggcaacctg 850

cgcaacagcc acatggtctc cgcgcagatt cgctgcaagc tgctcaaatc 900

tcggcagaca cctgagggtg agttccttcc ccttgaccaa cttgaactgg 950

atgtaggttt tagtacaggg gcagatcaac tttttcttgt gtccccctc 1000

acaatttgcc acgtgatcga tgccaaaagc ccctttttatg acctatccca 1050

gcgaagcatg caaactgaac agttcgagat tgtcgtcatc ctagaaggca 1100

ttgtggaaac aactgggatg acttgtcaag ctcgaacatc atatactgaa 1150

gatgaagttc tttggggtca tcgtttttt cctgtaattt ccttagaaga 1200

gggattcttt aaagttgatt actcccagtt ccacgcaaca tttgaagtcc 1250

ccaccccacc ttacagtgtg aaagagcagg aggaaatgct tctcatgtcg 1300

tccccttaa tagcaccagc cataactaac agcaaagaaa gacataattc 1350

tgtggaatgc ttagatggac tagatgatat tactacaaaa ctaccatcta 1400

agctgcagaa aattactgga agagaagact ttcccaaaaa actcttgagg 1450

atgagttcta caacttcaga aaaagcctac agcttgggag acttgcccat 1500

gaaacttcaa cgaataagtt cagttccggg caactcagaa gaaaaactgg 1550

tatctaaaac caccaagatg ttatctgatc ccatgagcca gtctgtggct 1600

gatttgccac caaagcttca aaagatggct ggaggagcag ctaggatgga 1650

agggaacctt ccagccaaat taagaaaaat gaactctgat cgcttcacat 1700

aacaaagcac tcccttaggc attatttaat gtttgattta gtaatagtcc 1750

aatatttggc gatgaggtaa ttctccctaa ggaatctgaa agtatatttt 1800

cctcccagtt ctacagcata tttcgagaac ccttcctttc ccaagtattg 1850

cgaatgtgca gaaagcaaca gttacggagg gaggacatca taaggaagtt 1900

attaacgggc atgtattatc acatcaagca tgcaataatg tgcaaatttt 1950

gcatttagtt ttatggcatg atttatatat ggcatattta tattgtatat 2000

tctggaaaaa aaaatatata tatatattta aaggggagat actctccctg 2050

acatttctaa catatgtatt aagccaaaca tgagtgaata gctttcaggg 2100

cgataaaact aaatatatgt ctgtgtgtgt gtgtgtatgt atacacacat 2150

atacatatat atatacacat acatacacat acatacatac atacatatat 2200

atctgataaa attgtgatgt tttgttcaaa gttgtagttc ttgtgcatgt 2250

ttactttatt agagtaggaa ggctactggc attaattatt aataccaaat 2300

attttagcct taaatttttg tcattttaaa atctgattta atgttttctg 2350

ctgtttaagg tcttgggagg ctttcaattg tattttatat gagagaatca 2400

cacaagtttg tgctatctat ggccctgcaa aaatataacc attacatgtt 2450

taaattgtaa attttagagc ataccagtac tcagtatagc attgaacatt 2500

tcttatgatt tttaaaagtt gctagtactg gggagaaata attgttgatt 2550

aatttgagaa ttattccttt cctagactaa ttaaaatctg gaaatctgtt 2600

ttgtatatga tctaatacaa agatgagctc tgaacaaaca ctgaatcatg 2650

ttaatagaca gtagccaagt tatattgaat atatcagaat ctgtgtgaag 2700

ttacacaatt aattgtccct gtttcaaact gagtaaattg gaaacatttt 2750

ctttcttttt ctggaaattt tgtccatttt aaaaaccaat cattttaaga 2800

1399
1400

agacatgaca atgcaatgaa acagatgata aatatttatg cttaaaataa 2850

aaaaaaaaaa a 2861

<210> 51
<211> 3557
<212> DNA
<213> Homo sapiens

<400> 51
gagagggtcc ttcagggtct gcttatgccc ttgttcaaga acaccagtgt 50

cagctctctg tactctggtt gcagactgac cttgctcagg cctgagaagg 100

atggggcagc caccagagtg gatgctgtct gcacccatcg tcctgacccc 150

aaaagccctg gactggacag agagcggctg tactggaagc tgagccagct 200

gacccacggc atcactgagc tgggccccta caccctggac aggcacagtc 250

tctatgtcaa tggtttcacc catcagagct ctatgacgac caccagaact 300

cctgatacct ccacaatgca cctggcaacc tcgagaactc cagcctccct 350

gtctggacct acgaccgcca gccctctcct ggtgctattc acaattaact 400

tcaccatcac taacctgcgg tatgaggaga acatgcatca ccctggctct 450

agaaagttta acaccacgga gagagtcctt cagggtctgc tcaggcctgt 500

gttcaagaac accagtgttg gccctctgta ctctggctgc agactgacct 550

tgctcaggcc caagaaggat ggggcagcca ccaaagtgga tgccatctgc 600

acctaccgcc ctgatcccaa aagccctgga ctggacagag agcagctata 650

ctgggagctg agccagctaa cccacagcat cactgagctg ggcccctaca 700

ccctggacag ggacagtctc tatgtcaatg gtttcacaca gcggagctct 750

gtgcccacca ctagcattcc tgggaccccc acagtggacc tgggaacatc 800

tgggactcca gtttctaaac ctggtccctc ggctgccagc cctctcctgg 850

tgctattcac tctcaacttc accatcacca acctgcggta tgaggagaac 900

atgcagcacc ctggctccag gaagttcaac accacggaga gggtccttca 950

gggcctgctc aggtccctgt tcaagagcac cagtgttggc cctctgtact 1000

ctggctgcag actgactttg ctcaggcctg aaaaggatgg gacagccact 1050

ggagtggatg ccatctgcac ccaccaccct gaccccaaaa gccctaggct 1100

ggacagagag cagctgtatt gggagctgag ccagctgacc cacaatatca 1150

ctgagctggg ccactatgcc ctggacaacg acagcctctt tgtcaatggt 1200

ttcactcatc ggagctctgt gtccaccacc agcactcctg ggacccccac 1250

agtgtatctg ggagcatcta agactccagc ctcgatattt ggcccttcag 1300

ctgccagcca tctcctgata ctattcaccc tcaacttcac catcactaac 1350

ctgcggtatg aggagaacat gtggcctggc tccaggaagt tcaacactac 1400

agagagggtc cttcagggcc tgctaaggcc cttgttcaag aacaccagtg 1450

ttggccctct gtactctggc tccaggctga ccttgctcag gccagagaaa 1500

gatggggaag ccaccggagt ggatgccatc tgcacccacc gccctgaccc 1550

cacaggccct gggctggaca gagagcagct gtatttggag ctgagccagc 1600

tgacccacag catcactgag ctgggccccc acacactgga cagggacagt 1650

1400
1401

ctctatgtca atggtttcac ccatcggagc tctgtaccca ccaccagcac 1700

cggggtggtc agcgaggagc cattcacact gaacttcacc atcaacaacc 1750

tgcgctacat ggcggacatg ggccaacccg gctccctcaa gttcaacatc 1800

acagacaacg tcatgaagca cctgctcagt cctttgttcc agaggagcag 1850

cctgggtgca cggtacacag gctgcagggt catcgcacta aggtctgtga 1900

agaacggtgc tgagacacgg gtggacctcc tctgcaccta cctgcagccc 1950

ctcagcggcc caggtctgcc tatcaagcag gtgttccatg agctgagcca 2000

gcagacccat ggcatcaccc ggctgggccc ctactctctg gacaaagaca 2050

gcctctacct taacggttac aatgaacctg gtctagatga gcctcctaca 2100

actcccaagc cagccaccac attcctgcct cctctgtcag aagccacaac 2150

agccatgggg taccacctga agaccctcac actcaacttc accatctcca 2200

atctccagta ttcaccagat atgggcaagg gctcagctac attcaactcc 2250

accgaggggg tccttcagca cctgctcaga cccttgttcc agaagagcag 2300

catgggcccc ttctacttgg gttgccaact gatctccctc aggcctgaga 2350

aggatggggc agccactggt gtggacacca cctgcaccta ccaccctgac 2400

cctgtgggcc ccgggctgga catacagcag ctttactggg agctgagtca 2450

gctgacccat ggtgtcaccc aactgggctt ctatgtcctg gacagggata 2500

gcctcttcat caatggctat gcaccccaga atttatcaat ccggggcgag 2550

taccagataa atttccacat tgtcaactgg aacctcagta atccagaccc 2600

cacatcctca gagtacatca ccctgctgag ggacatccag gacaaggtca 2650

ccacactcta caaaggcagt caactacatg acacattccg cttctgcctg 2700

gtcaccaact tgacgatgga ctccgtgttg gtcactgtca aggcattgtt 2750

ctcctccaat ttggacccca gcctggtgga gcaagtcttt ctagataaga 2800

ccctgaatgc ctcattccat tggctgggct ccacctacca gttggtggac 2850

atccatgtga cagaaatgga gtcatcagtt tatcaaccaa caagcagctc 2900

cagcacccag cacttctacc cgaatttcac catcaccaac ctaccatatt 2950

cccaggacaa agcccagcca ggcaccacca attaccagag gaacaaaagg 3000

aatattgagg atgcgctcaa ccaactcttc cgaaacagca gcatcaagag 3050

ttatttttct gactgtcaag tttcaacatt caggtctgtc cccaacaggc 3100

accacaccgg ggtggactcc ctgtgtaact tctcgccact ggctcggaga 3150

gtagacagag ttgccatcta tgaggaattt ctgcggatga cccggaatgg 3200

tacccagctg cagaacttca ccctggacag gagcagtgtc cttgtggatg 3250

ggtattctcc caacagaaat gagcccttaa ctgggaattc tgaccttccc 3300

ttctgggctg tcatcttcat cggcttggca ggactcctgg gactcatcac 3350

atgcctgatc tgcggtgtcc tggtgaccac ccgccggcgg aagaaggaag 3400

gagaatacaa cgtccagcaa cagtgcccag gctactacca gtcacaccta 3450

gacctggagg atctgcaatg actggaactt gccggtgcct ggggtgcctt 3500

tcccccagcc agggtccaaa gaagcttggc tggggcagaa ataaaccata 3550

1401
1402

ttggtcg 3557

<210> 52
<211> 3866
<212> DNA
<213> Homo sapiens

<400> 52
gcccgggcag gtgctgatag cacagttctg tccagagaag gaaggcggaa 50

taaacttatt cattcccagg aactcttggg gtaggtgtgt gtttttcaca 100

tcttaaaggc tcacagaccc tgcgctggac aaatgttcca ttcctgaagg 150

acctctccag aatccggatt gctgaatctt ccctgttgcc tagaagggct 200

ccaaaccacc tcttgacaat gggaaactgg gtggttaacc actggttttc 250

agttttgttt ctggttgttt ggttagggct gaatgttttc ctgtttgtgg 300

atgccttcct gaaatatgag aaggccgaca aatactacta cacaagaaaa 350

atccttgggt caacattggc ctgtgcccga gcgtctgctc tctgcttgaa 400

ttttaacagc acgctgatcc tgcttcctgt gtgtcgcaat ctgctgtcct 450

tcctgagggg cacctgctca ttttgcagcc gcacactgag aaagcaattg 500

gatcacaacc tcaccttcca caagctggtg gcctatatga tctgcctaca 550

tacagctatt cacatcattg cacacctgtt taactttgac tgctatagca 600

gaagccgaca ggccacagat ggctcccttg cctccattct ctccagccta 650

tctcatgatg agaaaaaggg gggttcttgg ctaaatccca tccagtcccg 700

aaacacgaca gtggagtatg tgacattcac cagcgttgct ggtctcactg 750

gagtgatcat gacaatagcc ttgattctca tggtaacttc agctactgag 800

ttcatccgga ggagttattt tgaagtcttc tggtatactc accacctttt 850

tatcttctat atccttggct tagggattca cggcattggt ggaattgtcc 900

ggggtcaaac agaggagagc atgaatgaga gtcatcctcg caagtgtgca 950

gagtcttttg agatgtggga tgatcgtgac tcccactgta ggcgccctaa 1000

gtttgaaggg catccccctg agtcttggaa gtggatcctt gcaccggtca 1050

ttctttatat ctgtgaaagg atcctccggt tttaccgctc ccagcagaag 1100

gttgtgatta ccaaggttgt tatgcaccca tccaaagttt tggaattgca 1150

gatgaacaag cgtggcttca gcatggaagt ggggcagtat atctttgtta 1200

attgcccctc aatctctctc ctggaatggc atccttttac tttgacctct 1250

gctccagagg aagatttctt ctccattcat atccgagcag caggggactg 1300

gacagaaaat ctcataaggg ctttcgaaca acaatattca ccaattccca 1350

ggattgaagt ggatggtccc tttggcacag ccagtgagga tgttttccag 1400

tatgaagtgg ctgtgctggt tggagcagga attggggtca ccccctttgc 1450

ttctatcttg aaatccatct ggtacaaatt ccagtgtgca gaccacaacc 1500

tcaaaacaaa aaagatctat ttctactgga tctgcaggga gacaggtgcc 1550

ttttcctggt tcaacaacct gttgacttcc ctggaacagg agatggagga 1600

attaggcaaa gtgggttttc taaactaccg tctcttcctc accggatggg 1650

acagcaatat tgttggtcat gcagcattaa actttgacaa ggccactgac 1700

1402
1403

atcgtgacag gtctgaaaca gaaaacctcc tttgggagac caatgtggga 1750

caatgagttt tctacaatag ctacctccca ccccaagtct gtagtgggag 1800

ttttcttatg tggccctcgg actttggcaa agagcctgcg caaatgctgt 1850

caccgatatt ccagtctgga tcctagaaag gttcaattct acttcaacaa 1900

agaaaatttt tgagttatag gaataaggac ggtaatctgc attttgtctc 1950

tttgtatctt cagtaattga gttataggaa taaggacggt aatctgcatt 2000

ttgtctcttt gtatcttcag taatttactt ggtctcgtca ggtttgagca 2050

gtcactttag gataagaatg tgcctctcaa gccttgactc cctggtattc 2100

tttttttgat tgcattcaac ttcgttactt gagcttcagc aacttaagaa 2150

cttctgaagt tcttaaagtt ctgaagttct taaagcccat ggatcctttc 2200

tcagaaaaat aactgtaaat ctttctggac agccatgact gtagcaaggc 2250

ttgatagcag aggtttggtg gttcagagtt atacaactaa tcccaggtga 2300

ttttatcaat tccagtgtta ccatctcctg agttttggtt tgtaatcttt 2350

tgtccctccc acccccacag aagatttcta agtagggtga ctttttaaat 2400

aaaaatttat tgaataatta atgataaaac ataataataa acataaataa 2450

taaacaaaat taccgagaac cccatcccca tataacacca acagtgtaca 2500

tgtttactgt cacttttgat atggtcttat ccagtgtgaa cagcaattta 2550

ttatttttgc tcatcaaaaa ataaaggatt ttcttcttca cttgatgaat 2600

gagtcttttg ttacttcttt ttgggctctc catttgggaa agaaaatcta 2650

gaaagtatgc tcattgtaat tgaaattgta ttataattgc aagtcagcta 2700

tactgtcatc ttgtattatt ctgttaaaag cgattgtgaa aacttgaact 2750

ccatgttagg atgacttatt tagtcaagag gagctcttta tcactattaa 2800

gggagtttat taatgtctag tttaaaaaat aacaatttat aattttcata 2850

gatttttgtg gactcctacc tatatcgcta aggaatcatt cagctaataa 2900

cataagcaaa atgccttcat tctctatcag aatgcttttt atttctgcca 2950

tctacccaac ctcggcctac aaggaccaca ctggctctgg gatgttagtc 3000

gtcacagtga atccctagaa ggagagaagc ccagggctct taagagagaa 3050

acctcacact tcattgtgcc cagacaattc aactcattat agattagaag 3100

aatggttgta ctgaattagc attatattat ctatttttc taccattaag 3150

tgaccacgaa tgaacaaatg gttgtttcta atagtttagg aacctctacg 3200

gcttgaggga taaaactgtt tttggttttt ttttgtcact agtctgagca 3250

caataaaggt aggtcatgaa ctcatggtgg ctaatcatgt ttcatttatt 3300

acaacagtaa aacaatactg atggttgttt acatagggag ggataccact 3350

gctgttgcca atttgttcac ccggatttca gctaattcag gtttaataag 3400

gctttgggaa aatgcacaat gaaaccttga agacatgatt cttattagac 3450

aactactgac caaatgacaa cttcatcttc aatgaccttg atttggaatc 3500

ttgaaactta caacgagttc ctaatgagaa aggcatcact cttggtctcc 3550

atgtctaaat atattgacac actcacttca agatatgatg gctgccaaaa 3600

1403
1404

tcaatatccc aaggcttgaa tatttatggg agtaggcata aattcgtgat 3650

tatattcatg gtaatgaata taataaatcc tatgcatagc caacatccaa 3700

gttgaactgg ggatgagcag gatacacaga aagctcctca cctgctttga 3750

gcaagttttt cttgcaaagt tgtaatatga aataccttaa tcaacaaatt 3800

tcttataaat gtttatatga gtagaaaatc cctcaatttc attgtcactc 3850

aaaaagttga ttatct 3866

<210> 53
<211> 1817
<212> DNA
<213> Homo sapiens

<400> 53
cccttgggct ccaaaccacc tcttgacaat gggaaactgg gtggttaacc 50

actggttttc agttttgttt ctggttgttt ggttagggct gaatgttttc 100

ctgtttgtgg atgccttcct gaaatatgag aaggccgaca aatactacta 150

cacaagaaaa atccttgggt caacattggc ctgtgcccga gcgtctgctc 200

tctgcttgaa ttttaacagc acgctgatcc tgcttcctgt gtgtcgcaat 250

ctgctgtcct tcctgagggg cacctgctca ttttgcagcc gcacactgag 300

aaagcaattg gatcacaacc tcaccttcca caagctggtg gcctatatga 350

tctgcctaca tacagctatt cacatcattg cacacctgtt taactttgac 400

tgctatagca gaagccgaca ggccacagat ggctcccttg cctccattct 450

ctccagccta tctcatgatg agaaaaaggg gggttcttgg ctaaatccca 500

tccagtcccg aaacacgaca gtggagtatg tgacattcac cagcattgct 550

ggtctcactg gagtgatcat gacaatagcc ttgattctca tggtaacttc 600

agctactgag ttcatccgga ggagttattt tgaagtcttc tggtatactc 650

accacctttt tatcttctat atccttggct tagggattca cggcattggt 700

ggaattgtcc ggggtcaaac agaggagagc atgaatgaga gtcatcctcg 750

caagtgtgca gagtcttttg agatgtggga tgatcgtgac tcccactgta 800

ggcgccctaa gtttgaaggg catccccctg agtcttggaa gtggatcctt 850

gcaccggtca ttctttatat ctgtgaaagg atcctccggt tttaccgctc 900

ccagcagaag gttgtgatta ccaaggttgt tatgcaccca tccaaagttt 950

tggaattgca gatgaacaag cgtggcttca gcatggaagt ggggcagtat 1000

atctttgtta attgcccctc aatctctctc ctggaatggc atccttttac 1050

tttgacctct gctccagagg aagatttctt ctccattcat atccgagcag 1100

caggggactg gacagaaaat ctcataaggg ctttcgaaca acaatattca 1150

ccaattccca ggattgaagt ggatggtccc tttggcacag ccagtgagga 1200

tgttttccag tatgaagtgg ctgtgctggt tggagcagga attggggtca 1250

ccccctttgc ttctatcttg aaatccatct ggtacaaatt ccagtgtgca 1300

gaccacaacc tcaaaacaaa aaagatctat ttctactgga tctgcaggga 1350

gacaggtgcc ttttcctggt tcaacaacct gttgacttcc ctggaacagg 1400

agatggagga attaggcaaa gtgggttttc taaactaccg tctcttcctc 1450

accggatggg acagcaatat tgttggtcat gcagcattaa actttgacaa 1500

ggccactgac atcgtgacag gtctgaaaca gaaaacctcc tttgggagac 1550

caatgtggga caatgagttt tctacaatag ctacctccca ccccaagtct 1600

gtagtgggag ttttcttatg tggccctcgg actttggcaa agagcctgcg 1650

caaatgctgt caccgatatt ccagtctgga tcctagaaag gttcaattct 1700

acttcaacaa agaaaatttt tgagttatag gaataaggac ggtaatctgc 1750

attttgtctc tttgtatctt cagtaattta cttggtctcg tcaggtttga 1800

gcagtcactt taggaag 1817

<210> 54
<211> 1000
<212> DNA
<213> Homo sapiens

<400> 54
gggctccaaa ccacctcttg acaatgggaa actgggtggt taaccactgg 50

ttttcagttt tgtttctggt tgtttggtta gggctgaatg ttttcctgtt 100

tgtggatgcc ttcctgaaat atgagaaggc cgacaaatac tactacacaa 150

gaaaaatcct tgggtcttgg aagtggatcc ttgcaccggt cattctttat 200

atctgtgaaa ggatcctccg gttttaccgc tcccagcaga aggttgtgat 250

taccaaggtt gttatgcacc catccaaagt tttggaattg cagatgaaca 300

agcgtggctt cagcatggaa gtggggcagt atatctttgt taattgcccc 350

tcaatctctc tcctggaatg gcatcctttt actttgacct ctgctccaga 400

ggaagatttc ttctccattc atatccgagc agcaggggac tggacagaaa 450

atctcataag ggctttcgaa caacaatatt caccaattcc caggattgaa 500

gtggatggtc cctttggcac agccagtgag gatgttttcc agtatgaagt 550

ggctgtgctg gttggagcag gaattggggt cacccccttt gcttctatct 600

tgaaatccat ctggtacaaa ttccagtgtg cagaccacaa cctcaaaaca 650

aaaaaggttg gtcatgcagc attaaacttt gacaaggcca ctgacatcgt 700

gacaggtctg aaacagaaaa cctcctttgg gagaccaatg tgggacaatg 750

agttttctac aatagctacc tcccacccca agtctgtagt gggagttttc 800

ttatgtggcc ctcggacttt ggcaaagagc ctgcgcaaat gctgtcaccg 850

atattccagt ctggatccta gaaaggttca attctacttc aacaaagaaa 900

atttttgagt tataggaata aggacggtaa tctgcatttt gtctctttgt 950

atcttcagta atttacttgg tctcgtcagg tttgagcagt cactttagga 1000

<210> 55
<211> 1665
<212> DNA
<213> Homo sapiens

<400> 55
gggctccaaa ccacctcttg acaatgggaa actgggtggt taaccactgg 50

ttttcagttt tgtttctggt tgtttggtta gggctgaatg ttttcctgtt 100

tgtggatgcc ttcctgaaat atgagaaggc cgacaaatac tactacacaa 150

gaaaaatcct tgggtcaaca ttggcctgtg cccgagcgtc tgctctctgc 200

ttgaatttta acagcacgct gatcctgctt cctgtgtgtc gcaatctgct 250

gtccttcctg aggggcacct gctcattttg cagccgcaca ctgagaaagc 300

aattggatca caacctcacc ttccacaagc tggtggccta tatgatctgc 350

ctacatacag ctattcacat cattgcacac ctgtttaact ttgactgcta 400

tagcagaagc cgacaggcca cagatggctc ccttgcctcc attctctcca 450

gcctatctca tgatgagaaa aaggggggtt cttggctaaa tcccatccag 500

tcccgaaaca cgacagtgga gtatgtgaca ttcaccagca ttgctggtct 550

cactggagtg atcatgacaa tagccttgat tctcatggta acttcagcta 600

ctgagttcat ccggaggagt tattttgaag tcttctggta tactcaccac 650

ctttttatct tctatatcct tggcttaggg attcacggca ttggtggaat 700

tgtccggggt caaacagagg agagcatgaa tgagagtcat cctcgcaagt 750

gtgcagagtc ttttgagatg tgggatgatc gtgactccca ctgtaggcgc 800

cctaagtttg aagggcatcc ccctgagtct tggaagtgga tccttgcacc 850

ggtcattctt tatatctgtg aaaggatcct ccggttttac cgctcccagc 900

agaaggttgt gattaccaag gttgttatgc acccatccaa agttttggaa 950

ttgcagatga acaagcgtgg cttcagcatg gaagtggggc agtatatctt 1000

tgttaattgc ccctcaatct ctctcctgga atggcatcct tttactttga 1050

cctctgctcc agaggaagat ttcttctcca ttcatatccg agcagcaggg 1100

gactggacag aaaatctcat aagggctttc gaacaacaat attcaccaat 1150

tcccaggatt gaagtggatg gtcccctttgg cacagccagt gaggatgttt 1200

tccagtatga agtggctgtg ctggttggag caggaattgg ggtcaccccc 1250

tttgcttcta tcttgaaatc catctggtac aaattccagt gtgcagacca 1300

caacctcaaa acaaaaaagg ttggtcatgc agcattaaac tttgacaagg 1350

ccactgacat cgtgacaggt ctgaaacaga aaacctcctt tgggagacca 1400

atgtgggaca atgagttttc tacaatagct acctcccacc ccaagtctgt 1450

agtgggagtt ttcttatgtg gccctcggac tttggcaaag agcctgcgca 1500

aatgctgtca ccgatattcc agtctggatc ctagaaaggt tcaattctac 1550

ttcaacaaag aaaatttttg agttatagga ataaggacgg taatctgcat 1600

tttgtctctt tgtatcttca gtaatttact tggtctcgtc aggtttgagc 1650

agtcacttta ggaag 1665

<210> 56
<211> 2400
<212> DNA
<213> Homo sapiens

<400> 56
ggactgcttt gtaactgcta agattgcaga cagaaatagc acacaaccac 50

tgtgagctgt atgcgattca gaaaccaaga ccaaattttg ctcactttca 100

ttaatcagtt gctcagatag aaggaaatga catctggttc tgtcttcttc 150

tacatcttaa tttttggaaa atatttttct catgggggtg gacaggatgt 200

caagtgctcc cttggctatt tcccctgtgg gaacatcaca aagtgcttgc 250

ctcagctcct gcactgtaac ggtgtggacg actgcgggaa tcaggccgat 300

gaggacaact gtggagacaa caatggatgg tccatgcaat ttgacaaata 350

ttttgccagt tactacaaaa tgacttccca atatcctttt gaggcagaaa 400

cacctgaatg tttggtcggt tctgtgccag tgcaatgtct ttgccaaggt 450

ctggagcttg actgtgatga aaccaattta cgagctgttc catcggtttc 500

ttcaaatgtg actgcaatgt cacttcagtg gaacttaata agaaagcttc 550

ctcctgattg cttcaagaat tatcatgatc ttcagaagct gtacctgcaa 600

aacaataaga ttacatccat ctccatctat gctttcagag gactgaatag 650

ccttactaaa ctgtatctca gtcataacag aataaccttc ctgaagccgg 700

gtgtttttga agatcttcac agactagaat ggctgataat tgaagataat 750

cacctcagtc gaatttcccc accaacattt tatggactaa attctcttat 800

tctcttagtc ctgatgaata acgtcctcac ccgtttacct gataaacctc 850

tctgtcaaca catgccaaga ctacattggc tggaccttga aggcaaccat 900

atccataatt taagaaattt gacttttatt tcctgcagta atttaactgt 950

tttagtgatg aggaaaaaca aaattaatca cttaaatgaa aatacttttg 1000

cacctctcca gaaactggat gaattggatt taggaagtaa taagattgaa 1050

aatcttccac cgcttatatt caaggacctg aaggagctgt cacaattgaa 1100

tctttcctat aatccaatcc agaaaattca agcaaaccaa tttgattatc 1150

ttgtcaaact caagtctctc agcctagaag ggattgaaat ttcaaatatc 1200

caacaaagga tgtttagacc tcttatgaat ctctctcaca tatattttaa 1250

gaaattccag tactgtgggt atgcaccaca tgttcgcagc tgtaaaccaa 1300

acactgatgg aatttcatct ctagagaatc tcttggcaag cattattcag 1350

agagtatttg tctgggttgt atctgcagtt acctgctttg gaaacatttt 1400

tgtcatttgc atgcgacctt atatcaggtc tgagaacaag ctgtatgcca 1450

tgtcaatcat ttctctctgc tgtgccgact gcttaatggg aatatattta 1500

ttcgtgatcg gaggctttga cctaaagttt cgtggagaat acaataagca 1550

tgcgcagctg tggatggaga gtactcattg tcagcttgta ggatctttgg 1600

ccattctgtc cacagaagta tcagttttac tgttaacatt tctgacattg 1650

gaaaaataca tctgcattgt ctatcctttt agatgtgtga gacctggaaa 1700

atgcagaaca attacagttc tgattctcat ttggattact ggttttatag 1750

tggctttcat tccattgagc aataaggaat ttttcaaaaa ctactatggc 1800

accaatggag tatgcttccc tcttcattca gaagatacag aaagtattgg 1850

agcccagatt tattcagtgg caatttttct tggtattaat ttggccgcat 1900

ttatcatcat agttttttcc tatggaagca tgttttatag tgttcatcaa 1950

agtgccataa cagcaactga aatacggaat caagttaaaa aagagatgat 2000

ccttgccaaa cgtttttttct ttatagtatt tactgatgca ttatgctgga 2050

tacccatttt tgtagtgaaa tttctttcac tgcttcaggt agaaatacca 2100

ggtaccataa cctcttgggt agtgattttt attctgccca ttaacagtgc 2150

1406
1408

tttgaaccca attctctata ctctgaccac aagaccattt aaagaaatga 2200

ttcatcggtt ttggtataac tacagacaaa gaaaatctat ggacagcaaa 2250

ggtcagaaaa catatgctcc atcattcatc tgggtggaaa tgtggccact 2300

gcaggagatg ccacctgagt taatgaagcc ggacctttte acataccect 2350

gtgaaatgtc actgatttct caatcaacga gactcaattc ctattcatga 2400

<210> 57
<211> 2967
<212> DNA
<213> Homo sapiens

<400> 57
tagagatccc tcgacctcga cccacgcgtc cgaggaaaga aaaaaagagg 50

aatggaaaga gacagagaaa ggaaatggga gtggaaggag ggaggactgc 100

tttgtaactg ctaagattgc agacagaaat agcacacaac cactgtgagc 150

tgtatgcgat tcagaaacca agaccaaatt ttgctcactt tcattaatca 200

gttgctcaga tagaaggaaa tgacatctgg ttctgtcttc ttctacatct 250

taatttttgg aaaatatttt tctcatgggg gtggacagga tgtcaagtgc 300

tcccttggct atttcccctg tgggaacatc acaaagtgct tgcctcagct 350

cctgcactgt aacggtgtgg acgactgcgg gaatcaggcc gatgaggaca 400

actgtggaga caacaatgga tggtctctgc aatttgacaa atattttgcc 450

agttactaca aaatgacttc ccaatatcct tttgaggcag aaacacctga 500

atgtttggtc ggttctgtgc cagtgcaatg tctttgccaa ggtctggagc 550

ttgactgtga tgaaaccaat ttacgagctg ttccatcggt ttcttcaaat 600

gtgactgcaa tgtcacttca gtggaactta ataagaaagc ttcctcctga 650

ttgcttcaag aattatcatg atcttcagaa gctgtacctg caaaacaata 700

agattacatc catctccatc tatgctttca gaggactgaa tagccttact 750

aaactgtatc tcagtcataa cagaataacc ttcctgaagc cgggtgtttt 800

tgaagatctt cacagactag aatggctgat aattgaagat aatcacctca 850

gtcgaatttc cccaccaaca ttttatggac taaattctct tattctctta 900

gtcctgatga ataacgtcct cacccgttta cctgataaac ctctctgtca 950

acacatgcca agactacatt ggctggacct tgaaggcaac catatccata 1000

atttaagaaa tttgacttt atttcctgca gtaatttaac tgttttagtg 1050

atgaggaaaa acaaaattaa tcacttaaat gaaaatactt ttgcacctct 1100

ccagaaactg gatgaattgg atttaggaag taataagatt gaaaatcttc 1150

caccgcttat attcaaggac ctgaaggagc tgtcacaatt gaatctttcc 1200

tataatccaa tccagaaaat tcaagcaaac caatttgatt atcttgtcaa 1250

actcaagtct ctcagcctag aagggattga aatttcaaat atccaacaaa 1300

ggatgtttag acctcttatg aatctctctc acatatattt taagaaattc 1350

cagtactgtg ggtatgcacc acatgttcgc agctgtaaac caaacactga 1400

tggaatttca tctctagaga atctcttggc aagcattatt cagagagtat 1450

ttgtctgggt tgtatctgca gttacctgct ttggaaacat ttttgtcatt 1500

1408
1409

tgcatgcgac cttatatcag gtctgagaac aagctgtatg ccatgtcaat 1550

catttctctc tgctgtgccg actgcttaat gggaatatat ttattcgtga 1600

tcggaggctt tgacctaaag tttcgtggag aatacaataa gcatgcgcag 1650

ctgtggatgg agagtactca ttgtcagctt gtaggatctt tggccattct 1700

gtccacagaa gtatcagttt tactgttaac atttctgaca ttggaaaaat 1750

acatctgcat tgtctatcct tttagatgtg tgagacctgg aaaatgcaga 1800

acaattacag ttctgattct catttggatt actggtttta tagtggcttt 1850

cattccattg agcaataagg aatttttcaa aaactactat ggcaccaatg 1900

gagtatgctt ccctcttcat tcagaagata cagaaagtat tggagcccag 1950

atttattcag tggcaatttt tcttggtatt aatttggccg catttatcat 2000

catagttttt tcctatggaa gcatgtttta tagtgttcat caaagtgcca 2050

taacagcaac tgaaatacgg aatcaagtta aaaaagagat gatccttgcc 2100

aaacgttttt tctttatagt atttactgat gcattatgct ggatacccat 2150

ttttgtagtg aaatttcttt cactgcttca ggtagaaata ccaggtacca 2200

taacctcttg ggtagtgatt tttattctgc ccattaacag tgctttgaac 2250

ccaattctct atactctgac cacaagacca tttaaagaaa tgattcatcg 2300

gttttggtat aactacagac aaagaaaatc tatggacagc aaaggtcaga 2350

aaacatatgc tccatcattc atctgggtgg aaatgtggcc actgcaggag 2400

atgccacctg agttaatgaa gccggacctt ttcacatacc cctgtgaaat 2450

gtcactgatt tctcaatcaa cgagactcaa ttcctattca tgactgactc 2500

tgaaattcat ttcttcgcag agaatactgt gggggtgctt catgagggat 2550

ttactggtat gaaatgaata ccacaaaatt aatttataat aatagctaag 2600

ataaatattt tacaaggaca tgaggaaaaa taaaaatgac taatgctctt 2650

acaaagggaa gtaattatat caataatgta tatatattag tagacatttt 2700

gcataagaaa ttaagagaaa tctacttcag taacattcat tcatttttct 2750

aacatgcatt tattgagtac ccactactat gtgcatagca ttgcaatata 2800

gtcctggaag tagacagtgc agaacctttc aatctgtaga tggtgtttaa 2850

tgacaaaaga ctatacaaag tccatctgca gttcctagtt taaagtagag 2900

ctttacctgt catgtgcatc agcaagaatc ataggcactt ttaaataaag 2950

gtttaaagtt ttggaat 2967

<210> 58
<211> 3032
<212> DNA
<213> Homo sapiens

<400> 58
gatccattgt gctggcgcgg attctttatc actggaagga ggaggagggg 50

gagagtcgct cccgccgggc gagcatgggg cgcctggcct cgaggccgct 100

gctgctggcg ctcctgtcgt tggctctttg ccgagggcgt gtggtgagag 150

tccccacagc gaccctggtt cgagtggtgg gcactgagct ggtcatcccc 200

tgcaacgtca gtgactatga tggccccagc gagcaaaact ttgactggag 250

1408
1410

cttctcatct ttggggagca gctttgtgga gcttgcaagc acctgggagg 300

tggggttccc agcccagctg taccaggagc ggctgcagag gggcgagatc 350

ctgttaaggc ggactgccaa cgacgccgtg gagctccaca taaagaacgt 400

ccagccttca gaccaaggcc actacaaatg ttcaaccccc agcacagatg 450

ccactgtcca gggaaactat gaggacacag tgcaggttaa agtgctggcc 500

gactccctgc acgtgggccc cagcgcgcgg cccccgccga gcctgagcct 550

gcgggagggg gagcccttcg agctgcgctg caccgccgcc tccgcctcgc 600

cgctgcacac gcacctggcg ctgctgtggg aggtgcaccg cggcccggcc 650

aggcggagcg tcctcgccct gacccacgag ggcaggttcc acccgggcct 700

ggggtacgag cagcgctacc acagtgggga cgtgcgcctc gacaccgtgg 750

gcagcgacgc ctaccgcctc tcagtgtccc gggctctgtc tgccgaccag 800

ggctcctaca ggtgtatcgt cagcgagtgg atcgccgagc agggcaactg 850

gcaggaaatc caagaaaagg ccgtggaagt tgccaccgtg gtgatccagc 900

cgacagttct gcgagcagct gtgcccaaga atgtgtctgt ggctgaagga 950

aaggaactgg acctgacctg taacatcaca acagaccgag ccgatgacgt 1000

ccggcccgag gtgacgtggt ccttcagcag gatgcctgac agcaccctac 1050

ctggctcccg cgtgttggcg cggcttgacc gtgattccct ggtgcacagc 1100

tcgcctcatg ttgctttgag tcatgtggat gcacgctcct accatttact 1150

ggttcgggat gttagcaaag aaaactctgg ctactattac tgccacgtgt 1200

ccctgtgggc acccggacac aacaggagct ggcacaaagt ggcagaggcc 1250

gtgtcttccc cagctggtgt gggtgtgacc tggctagaac cagactacca 1300

ggtgtacctg aatgcttcca aggtccccgg gtttgcggat gaccccacag 1350

agctggcatg ccgggtggtg gacacgaaga gtggggaggc gaatgtccga 1400

ttcacggttt cgtggtacta caggatgaac cggcgcagcg acaatgtggt 1450

gaccagcgag ctgcttgcag tcatggacgg ggactggacg ctaaaatatg 1500

gagagaggag caagcagcgg gcccaggatg gagactttat tttttctaag 1550

gaacatacag acacgttcaa tttccggatc caaaggacta cagaggaaga 1600

cagaggcaat tattactgtg ttgtgtctgc ctggaccaaa cagcggaaca 1650

acagctgggt gaaaagcaag gatgtcttct ccaagcctgt taacatattt 1700

tgggcattag aagattccgt gcttgtggtg aaggcgaggc agccaaagcc 1750

tttctttgct gccggaaata catttgagat gacttgcaaa gtatcttcca 1800

agaatattaa gtcgccacgc tactctgttc tcatcatggc tgagaagcct 1850

gtcggcgacc tctccagtcc caatgaaacg aagtacatca tctctctgga 1900

ccaggattct gtggtgaagc tggagaattg gacagatgca tcacgggtgg 1950

atggcgttgt tttagaaaaa gtgcaggagg atgagttccg ctatcgaatg 2000

taccagactc aggtctcaga cgcagggctg taccgctgca tggtgacagc 2050

ctggtctcct gtcaggggca gcctttggcg agaagcagca accagtctct 2100

ccaatcctat tgagatagac ttccaaacct caggtcctat atttaatgct 2150

tctgtgcatt cagacacacc atcagtaatt cggggagatc tgatcaaatt 2200

gttctgtatc atcactgtcg agggagcagc actggatcca gatgacatgg 2250

cctttgatgt gtcctggttt gcggtgcact cttttggcct ggacaaggct 2300

cctgtgctcc tgtcttccct ggatcggaag ggcatcgtga ccacctcccg 2350

gagggactgg aagagcgacc tcagcctgga gcgcgtgagt gtgctggaat 2400

tcttgctgca agtgcatggc tccgaggacc aggactttgg caactactac 2450

tgttccgtga ctccatgggt gaagtcacca acaggttcct ggcagaagga 2500

ggcagagatc cactccaagc ccgtttttat aactgtgaag atggatgtgc 2550

tgaacgcctt caagtatccc ttgctgatcg gcatcggtct gtccacggtc 2600

atcgggctcc tgtcctgtct catcgggtac tgcagctccc actggtgttg 2650

taagaaggag gttcaggaga cacggcgcga gcgccgcagg ctcatgtcga 2700

tggagatgga ctaggctggc ccgggagggg agtgacagag ggacgttcta 2750

ggagcaattg gggcaagaag aggacagtga tattttaaaa caaagtgtgt 2800

tacactaaaa accagtcctc tctaatctca ggtgggactt ggcgctctct 2850

cttttctgca tgtcaagttc tgagcgcgga catgtttacc agcacacggc 2900

tcttcttccc acggcacttt ctgatgtaac aatcgagtgt gtgttttccc 2950

aactgccagt gataaagaat ccgcgccagc acaatggatc tcgagggatc 3000

ttccatacct accagttctg cgcctgcagg tc 3032

<210> 59
<211> 2771
<212> DNA
<213> Homo sapiens

<400> 59
cccacgcgtc cgcggtaaat cacttcttgg aggtgccctg cacgccggtc 50

ctgggagcag gcggcctccc gggggtgcgg gagccccact cctccgtggt 100

gtgttccatt tgcttcccac atctggagga gctgacgtgc cagcctcccc 150

cagcaccacc cagggacggg aggcatgagc cggtcaaggc acctgggcaa 200

aatccggaag cgtctggaag atgtcaagag ccagtgggtc cggccagcca 250

gggctgactt tagtgacaac gagagtgccc ggctggccac ggacgccctc 300

ttggatgggg gttctgaagc ctactggcgg gtgctcagcc aggaaggcga 350

ggtggacttc ttgtcctcgg tggaggccca gtacatccag gcccaggcca 400

gggagccccc gtgtccccca gacaccctgg gaggggcgga agcaggccct 450

aagggactgg actccagctc cctacagtcc ggcacctact tccctgtggc 500

ctcagagggc agcgagccgg ccctactgca cagctgggcc tcagctgaga 550

agccctacct gaaggaaaaa tccagcgcca ctgtgtactt ccagaccgtc 600

aagcacaaca acatcagaga cctcgtccgc cgctgcatca cccggactag 650

ccaggtcctg gtcatcctga tggatgtgtt cacggatgtg gagatcttct 700

gtgacattct agaggcagcc aacaagcgtg gggtgttcgt ttgtgtgctc 750

ctggaccagg gaggtgtgaa gctcttccag gagatgtgtg acaaagtcca 800

gatctctgac agtcacctca agaacatttc catccggagt gtggaaggag 850

1410
1412

agatatactg tgccaagtca ggcaggaaat tcgctggcca aatccgggag 900

aagttcatca tctcggactg gagatttgtc ctgtctggat cttacagctt 950

cacctggctc tgcggacacg tgcaccggaa catcctctcc aagttcacag 1000

gccaggcggt ggagctgttt gacgaggagt tccgccacct ctacgcctcc 1050

tccaagcctg tgatgggcct gaagtccccg cggctggtcg cccccgtccc 1100

gcccggagca gccccggcca atggccgcct tagcagcagc agtggctccg 1150

ccagtgaccg cacgtcctcc aacccccttca gcggccgctc ggcaggcagc 1200

caccccggta cccgaagtgt gtccgcgtct tcagggccct gtagccccgc 1250

ggccccacac ccgcctccac cgccccggtt ccagccccac caaggccctt 1300

ggggagcccc gagtccccag gcccacctct ccccgcggcc ccacgacggc 1350

ccgcccgccg ctgtctacag caacctgggg gcctacaggc ccacgcggct 1400

gcagctggag cagctgggcc tggtgccgag gctgactcca acctggaggc 1450

ccttcctgca ggcctccccc cacttctgaa ggtcccatcc cctgctgccc 1500

tccgcaggcc cagggctggg cactccctga gacccaaaga cccacctcaa 1550

cgacgagtgg cgttgagcca cttccctttg aaaagacact caaaatcact 1600

gccatggttc aatgttccca ggccccaggc catccacttg ccggcccccа 1650

ccagttcttg ggttccccgc tctagtttga cctgtgcagc acattccaga 1700

aggttccagg ggaggttgtg gggcagctag aggacaaaat catgaaaaca 1750

gagtccctgt cttccagaga tcatccgggg ctttaatatt aatggccccc 1800

aaaactccgt aagaagcagg aaatgcagcc caagttttac aaatgggtaa 1850

acagaggcac tgagagatag atggtagttt ggtacttctg gttcccagtg 1900

cccaggaatg gtccactccc aagaaattca ggaaagaaag actgaggaga 1950

aggtgtggga acattctgga tgtttcggga gagttgggga aactcctcct 2000

cttaggaaag gctaatacta gggtatcctt gggcccaatg aattaggggt 2050

gaggccccag aacccgttat ctatgagttg tatgggggag ccatctgaag 2100

ctgtagccac cagggatgca gctagctgag gagtttgggg tgttgggttg 2150

gacaaggcag gttagtagac tcagattctt gcttcaaaga gccttgggct 2200

ggcctggagg tccctggagt ctagactgga cctaggagct tgagttgtca 2250

ggggccagga ctggccccac tgcagtgccc aggccagtct tgagcagcag 2300

ggagggctca gctgtcccca gatccaggtg cctctgacca gcctggtcac 2350

ctcctgagga ataaatgctg aacctcacaa gccccatcat tcatttcttc 2400

tcaattcaca gtgcccctct ttgtttctgg ggtggaacta ggtcctgagg 2450

gcacagccta gctgagtgca aagaaatata ggatgcttag aaagcataca 2500

ggaggggcca ggcgtggtgg ctcatgcctg taatcccaga actttgggat 2550

gccaaggtgg ttggattacc tgagatcagg tggattacct ggtctcgaga 2600

ccagcctgac caatatggtg aaaccccgtc tctactaaaa atacaaaaat 2650

taggctgaga caggagaatt gcttgaaccc aggaagcaga ggttgcaatg 2700

agctgagatt gcatcactgc actccagcat gggcaacaaa gcaagactcc 2750

1411
1413

gtcacagaaa aaaaaaaaaa a 2771

<210> 60
<211> 4060
<212> DNA
<213> Homo sapiens

<400> 60
atggctctgc cgcagccgcc gccgcccgcc ggggcgcgtg acccggtgac 50

cccggagcac tgggcatccg gacccccctc gcaggcgcag ccgctattaa 100

ggcaggaggc taaagaggag gaggaagggg aggagacagg cgtccagggc 150

gcctggggaa ccggcacggc ggagcagcgg cggcggggct ggggggaggc 200

cgccgagtcg gccgcggccg aggaggggca ggcggaggtc gggggcgccg 250

cggccgcggg gtcaggttcc ccggcgggag gcgcaggtgg cggcctgggg 300

agctggcggc cgctcctcgc ctggctgcag cggcggcagc cccagtgctg 350

cccctgtgcg gcgccccttt cccgctccgc cgcgcactgt tgtcatggag 400

gaaccaagat ggcggctctg gcctacaacc tgggcaagcg ggagatcaac 450

cactacttca gcgtgaggag cgccaaggtg ctggcgctgg tggccgtgct 500

gctgctcgca gcgtgccacc tcgcctcccg ccgctaccga ggcaatgatt 550

cgtgtgaata ccttctctca agtggcagat ttcttggaga gaaagtttgg 600

caacctcaca gttgtatgat gcataaatac aaaatcagtg aagcaaagaa 650

ctgccttgta gataaacata ttgcatttat tggagattcc agaattcgtc 700

aattgtttta ttcttttgta aaaataatta atccccaatt caaagaagaa 750

ggaaataagc atgaaaacat tccttttgaa gacaagactg catcagttaa 800

agtggatttt ctgtggcatc ctgaagttaa tggttctatg aaacagtgta 850

tcaaagtgtg gactgaggat tccattgcaa agccacatgt gattgtagca 900

ggagctgcca catggtccat caagattcac aatggtagca gtgaagcgct 950

ttctcaatat aaaatgaaca tcacctccat agcaccactt ttagaaaaat 1000

tggcaaagac tagtgatgtt tattgggtct tacaagatcc tgtttatgaa 1050

gatctattaa gtgaaaatag gaagatgatc actaatgaga agatagatgc 1100

ttacaatgaa gctgcagtca gtattttgaa tagtagcacc agaaattcta 1150

aatcaaatgt taagatgttc agtgtttcca aattaattgc tcaagaaacc 1200

atcatggaat ctttggatgg cttacatctt cctgaatcga gcagagaaac 1250

tactgcaatg attcttatga atgtgtattg caataagatt ttgaagcctg 1300

tagatgggtc ctgttgtcaa cctcggcctc ctgttactct catacagaag 1350

ctagctgctt gtttttttcac tttatctatt atcggatatt taatttttta 1400

cataattcat cgtaatgctc atcggaagaa taagccgtgt actgatttgg 1450

aaagtggaga ggaaaagaaa aatattatca ataccccctgt gtcttcatta 1500

gaaatacttt tacaatcttt ctgcaaactt ggcctgatta tggcatattt 1550

ctatatgtgt gaccgtgcaa atctgttcat gaaggaaaac aaattttata 1600

cacattcatc tttctttatt ccaattatct acattttggt tttgggagta 1650

ttttataatg aaaatactaa agagactaaa gtattaaata gagaacaaac 1700

agacgaatgg aaaggctgga tgcaacttgt gattttgatt tatcacattt 1750

ctggagcaag tacatttttg cctgtataca tgcacattcg agttctggtt 1800

gctgcatatt tatttcagac agggtatggg catttctcat acttttggat 1850

aaaaggagat tttggaatct atagagtatg tcaggtttta tttcgtctca 1900

atttcctggt agtggtgtta tgtatagtaa tggatcgacc ttatcaattc 1950

tattactttg tcccccttggt cactgtatgg ttcatggtca tatatgttac 2000

tttagcacta tggccacaaa taatccaaaa aaaagcaaac ggaaattgtt 2050

tctggcattt tggcttactg ttgaaaactag gctttttgct gttattcata 2100

tgtttttttgg catattctca gggtgcattt gagaagatct tttctctttg 2150

gccattgtcc aagtgttttg aactgaaagg gaatgtatat gaatggtggt 2200

tcagatggag gttagaccgt tatgtagttt tccacggaat gctgtttgct 2250

tttatttatc tggctttgca gaagcgtcaa atactttctg aaggaaaggg 2300

tgaacctctt ttttcaaaca aaatttcaaa ttttctgttg tttatttcag 2350

tagtttcttt cttgacctat tccatctggg ctagcagttg taaaaacaaa 2400

gcagagtgca atgaactcca tccgtctgtt tctgtggtac agattttagc 2450

cttcatccta ataagaaaca tccctggata tgcccgttca gtttacagtt 2500

catttttgc ttggtttgga aaaatttcat tagagctatt tatttgccag 2550

tatcacatat ggctggcagc ggacacaagg ggtatcttgg tactgatacc 2600

tggaaaccct atgctcaaca tcattgtcag cactttcata tttgtttgtg 2650

tggcacatga aatttctcag atcactaatg atcttgcaca gattattatt 2700

cctaaagata actcatctct cttgaaaagg ttggcatgta tagctgcatt 2750

tttttgtgga ctcctcatct tatcatccat tcaagataaa tcaaaacatt 2800

aggttccaaa aattctaaaa aacctaaact cttcaggcta cctttgtgtg 2850

tctctagaag agaaaagcat ctatctggag atataaatgt gtwtgtaaat 2900

ataaacgttt gtggcaagag gacagttctg tgacatctgt tgaacatatg 2950

tggttgtata tattggaaat gtacatatcc aatatgaaat actaaaacaa 3000

acaaacaaac aaaaaaccag aatgcattgt ataggattgc atgtgaagtc 3050

ttttctactg aatctatatt tccatttgta agtgatttta agttaacata 3100

tgaaggcagg gaaatgatta cctttccagt aaaaagtata gataatttaa 3150

ttaacttagt gacaccacca agtgttttga tataactaaa tttgtggtaa 3200

taagactgtc tgcacctgta ttcattgtgg aacttcctct ttcattggaa 3250

actttcttac tcaagaatga cggcagtatt gttttcttat atgtgcaatg 3300

aagtggaatg ataaacagta tgcctttaat ttatatgtgt tcttgttctg 3350

atgttgtttc ctgaaatgat ttttcttcct aactgtggtt ttcgggtatg 3400

caagcctaaa tctttgtaca ctttgtctca cagaatagtt ctgaggctcc 3450

atgacagggt tttgtcattg ttgatgttat tgttgcttcg ttttataaaa 3500

aagccaaaat ttttttttcca atccaaacgt tcacctgttt cctttcctca 3550

agctatacca gtgtaatacc agttaccctg tggatccatt taatatgtta 3600

tccccactaa ttaattttcg tatattattt ccaatatttg gaaagctctt 3650

tatagccatt tggtatttcc tattacccac ctcctatttt aaatatttat 3700

cagtctaaac ttgtgcagtg tagtaaacat gcaagttgtt acgattgagc 3750

tgtattacca taagtagaat tttaagtaaa ctggtgaatt tgggcaataa 3800

atgtttttgc tttttgtttg attttttttt acaagctaac tgttagaggt 3850

atacatttat ttatctgttg tacagatttg attatgattt taatgtttga 3900

aagattgcac ttgtttgctt ttactatatg tggggtaaaa tatattttct 3950

gttcacagta tatgaaaata tggagtaatt taaacagtaa ataaacattc 4000

tgtggatgct tatttttgta ttggcaaagt atcaattaaa ctatatgtgt 4050

tctttttcaa 4060

<210> 61
<211> 5472
<212> DNA
<213> Homo sapiens

<400> 61
gaaagacata cacacttcat gtaatgctac ctgcaagtct ccctagaaaa 50

gcagtttttg taggtgaaaa caatgaagcc aggtaatatt gcaaggaggc 100

tgtaatttta gcagacctac caacaacact gatgtaggaa gctcattatt 150

ttaatttctg gagcctttta atttttttctt tagaaagtgt ataaataatt 200

gcagtgctgc tttgcttcca aaactgggca gtgagttcaa caacaacgac 250

aacaacagcc gcagctcatc ctggccgtca tggagtttct tgaaagaacg 300

tatcttgtga atgataaagc tgccaagatg tatgctttca cactagaaag 350

aaggagctgc aaatgaacac ttcatagcaa tgtggaactc caacagaaac 400

cggtgaataa agatcagtgt cccagagaga gaccagagga gctggagtca 450

ggaggcatgt accactgcca cagtggctcc aagcccacag aaaaggggc 500

gaatgagtac gcctatgcca agtggaaact ctgttctgct tcagcaatat 550

gcttcatttt catgattgca gaggtcgtgg gtgggcacat tgctgggagt 600

cttgctgttg tcacagatgc tgcccacctc ttaattgacc tgaccagttt 650

cctgctcagt ctcttctccc tgtggttgtc atcgaagcct ccctctaagc 700

ggctgacatt tggatggcac cgagcagaga tccttggtgc cctgctctcc 750

atcctgtgca tctgggtggt gactggcgtg ctagtgtacc tggcatgtga 800

gcgcctgctg tatcctgatt accagatcca ggcgactgtg atgatcatcg 850

tttccagctg cgcagtggcg gccaacattg tactaactgt ggttttgcac 900

cagagatgcc ttggccacaa tcacaaggaa gtacaagcca atgccagcgt 950

cagagctgct tttgtgcatg cccttggaga tctatttcag agtatcagtg 1000

tgctaattag tgcacttatt atctacttta agccagagta taaaatagcc 1050

gacccaatct gcacattcat cttttccatc ctggtcttgg ccagcaccat 1100

cactatctta aaggacttct ccatcttact catggaaggt gtgccaaaga 1150

gcctgaatta cagtggtgtg aaagagctta ttttagcagt cgacggggtg 1200

ctgtctgtgc acagcctgca catctggtct ctaacaatga atcaagtaat 1250

~~1414~~
1416

tctctcagct catgttgcta cagcagccag ccgggacagc caagtggttc 1300

ggagagaaat tgctaaagcc cttagcaaaa gctttacgat gcactcactc 1350

accattcaga tggaatctcc agttgaccag gaccccgact gccttttctg 1400

tgaagacccc tgtgactagc tcagtcacac cgtcagtttc ccaaatttga 1450

caggccacct tcaaacatgc tgctatgcag tttctgcatc atagaaaata 1500

aggaaccaaa ggaagaaatt catgtcatgg tgcaatgcac attttatcta 1550

tttatttagt tccattcacc atgaaggaag aggcactgag atccatcaat 1600

caattggatt atatactgat cagtagctgt gttcaattgc aggaatgtgt 1650

atatagatta ttcctgagtg gagccgaagt aacagctgtt tgtaactatc 1700

ggcaatacca aattcatctc ccttccaata atgcatcttg agaacacata 1750

ggtaaatttg aactcaggaa agtcttacta gaaatcagtg gaagggacaa 1800

atagtcacaa aattttacca aaacattaga aacaaaaaat aaggagagcc 1850

aagtcaggaa taaaagtgac tctgtatgct aacgccacat tagaacttgg 1900

ttctctcacc aagctgtaat gtgatttttt tttctactct gaattggaaa 1950

tatgtatgaa tatacagaga agtgcttaca actaattttt atttacttgt 2000

cacattttgg caataaatcc ctcttatttc taaattctaa cttgtttatt 2050

tcaaaacttt atataatcac tgttcaaaag gaaatatttt cacctaccag 2100

agtgcttaaa cactggcacc agccaaagaa tgtggttgta gagacccaga 2150

agtcttcaag aacagccgac aaaaacattc gagttgaccc caccaagttg 2200

ttgccacaga taatttagat atttacctgc aagaaggaat aaagcagatg 2250

caaccaattc attcagtcca cgagcatgat gtgagcactg ctttgtgcta 2300

gacattgggc ttagcattga aactataaag aggaatcaga cgcagcaagt 2350

gcttctgtgt tctggtagca actcaacact atctgtggag agtaaactga 2400

agatgtgcag gccaacattc tggaaatcct atgtcactgg gtttggtttg 2450

gaaacctgga cttctgcatt tttaaaagtt acccagagat gcttctaaag 2500

atgagccata gtctagaaga ttgtcaacca caggagttca ttgagtggga 2550

cagctagaya catacattgg cagctacaat agtatcatga attgcaatga 2600

tgtagtgggg tataaaagga aagcgatgga tattgccgga tgggcatggc 2650

cagtgatgtt tcacgtcatt gaggtgacag ctctgctgga ctttgaatta 2700

catatggagg ctctccagga agacgaagaa gagaaggaca ttctaggcaa 2750

aaagaagact aggcacaagg cacacttatg tttgtctgtt agcttttagt 2800

tgaaaaagca aaatacatga tgcaaagaaa cctctccacg ctgtgatttt 2850

taaaactaca tacttttgc aactttatgg ttatgagtat tgtagagaac 2900

aggagatagg tcttagatga tttttatgtt gttgtcagac tctagcaagg 2950

tactagaaac ctagcaggca ttaataattg ttgaggcaat gactctgagg 3000

ctatatctgg gccttgtcat tatttatcat ttatatttgt atttttttct 3050

gaaatttgag ggccaagaaa acattgactt tgactgagga ggtcacatct 3100

gtgccatctc tgcaaatcaa tcagcaccac tgaaataact acttagcatt 3150

ctgctgagct ttccctgctc agtagagaca aatatactca tcccccacct 3200

cagtgagctt gtttaggcaa ccaggattag agctgctcag gttcccaacg 3250

tctcctgcca catcgggttc tcaaaatgga aagaatggtt tatgccaaat 3300

cacttttcct gtctgaagga ccactgaatg gttttgtttt tccatatttt 3350

gcataggacg ccctaaagac taggtgactt ggcaaacaca caagtgttag 3400

tataattctt tgcttctgct tctttttgaa aatcatgttt agatttgatt 3450

ttaagtcaga aattcactga atgtcaggta atcattatgg agggagattt 3500

gtgtgtcaac caaagtaatt gtcccatggc cccagggtat ttctgttgtt 3550

tccctgaaat tctgcttttt tagtcagcta gattgaaaac tctgaacagt 3600

agatgtttat atggcaaaat gcaagacaat ctataaggga gattttaagg 3650

attttgagat gaaaaaacag atgctactca ggggctttat ggaccatcca 3700

tcaattctga agttctgact ctcccattac cctttccctg gtgtggtcag 3750

aactccaggt cactggaagt tagtggaatc atgtagttga attctttact 3800

tcaagacatt gtattctctc cagctatcaa aacattaatg atcttttatg 3850

tctttttttt gttattgtta tactttaagt tctggggtac atgtgcggaa 3900

catgtaggtt tgttacatag gtatacatgt gccatggtgg tttgctgcac 3950

tcatcaacct gtcatctaca ttcttttatg tctgtctttc aaagcaacac 4000

tctgttcttc tgagtagtga aatcaggtca actttaccac cagcctccat 4050

ttttaatatg cttcaccatc atccagcacc tacttaagat ttatctaggg 4100

ctctgtggtg atgttaggac ccataaaaga aatttatgcc ttccatatgt 4150

ttggttacag atgggaaatg ggaatgttga aggacatgaa agaaaggatg 4200

tttacacatt aagcatcagt tctgaagcta gattgtctga gtttgaatct 4250

tagctcttcc ctttattagc tctgtgacct cgagctagtt acttaaatgc 4300

tctgatcctc tatttcctga tcagtgaaac ctccctattc aaatgtgtga 4350

gagtttaata aattaggaca cttaaaaatg ttggagcagt gcatagcatg 4400

tagtgttcag tacatgttaa atgttgtttt ttattatgta caaacatgtg 4450

tgggcacaga attttaaatc atctcaactt ttgagaaatt ttgagttatc 4500

aacaccgttc ccacaagaca gtggcaaaat tattggtgag aattaaacag 4550

ctgtttctca gaggaagcaa tggaggcttg ctgggataaa ggcatttact 4600

gagaggctgt tacctagtga gagtgatgaa ttaattaaaa tagtcgaatc 4650

cctttctgac tgtctctgaa agcttccgct tttatctttg aagagcagaa 4700

ttgtcacccc aaggacattt attaataaaa agaacaactg tccagtgcaa 4750

tgaaggcaaa gtcataggtc tcccaagtct taccccattc ctgtgaaata 4800

tcaagttctt ggcttttctc tgtcatgtag cctcaacttt ctccgaccgg 4850

gtgcatttct ttctctggtt tctaaattgc cagtggcaaa tttggatcac 4900

ttacttaata tctgttaaat tttgtgaccc aacaaagtct tttagcactg 4950

tggtgtcaaa aagaaaaaca cctcccaggc atatacattt tatagattcc 5000

tggagaatgt tgctctccag ctccatcccc acccaatgaa atatgatcca 5050

gagagtcttg caaagagaca agcctcattt tccacaatta gctctaaagt 5100

gcctccagga aatgattttc tcagctcatc tctctgtatt ccctgttttg 5150

gatcacaggg caatctgttt aaatgactaa ttacagaaat cattaaaggc 5200

accaagcaaa tgtcatctct gaatacacac atcccaagct ttacaaatcc 5250

tgcctggctt gacagtgatg aggccactta acagtccagc gcaggcggat 5300

gttaaaaaaa ataaaaaggt gaccatctgc ggtttagttt tttaactttc 5350

tgatttcaca cttaacgtct gtcattctgt tactgggcac ctgtttaaat 5400

tctattttaa aatgttaatg agtgttgttt aaaataaaat caggaaagag 5450

agaaaaaaaa aaaaaaaaaa aa 5472

<210> 62
<211> 5170
<212> DNA
<213> Homo sapiens

<400> 62
gcggcgggag agctgctggc tcgcccggat cccgggagct gcctggaggc 50

gggcccggcc cggggaaggt gagcggctgc gggacccagc ccctcgccgg 100

gagcgggcac catggtgctg tcggtgcctg tgatcgcgct gggcgccacg 150

ctgggcacag ccaccagcat cctcgcgttg tgcggggtca cctgcctgtg 200

tcggcacatg caccccaaga aggggctgct gccgcgggac caggaccccg 250

acctggagaa ggcgaagccc agcttgctcg ggtctgcaca acagttcaat 300

gttaaaaagt ccacggaacc tgttcagccc cgtgccctcc tcaagttccc 350

agacatctat ggacccaggc cagctgtgac ggctccagag gtcatcaact 400

atgcagacta ttcactgagg tctacggagg agcccactgc acctgccagc 450

ccccaacccc cgaatgacag tcgcctcaag aggcaggtca cagaggagct 500

gttcatcctc cctcagaatg gtgtggtgga ggatgtctgt gtcatggaga 550

cctggaaccc agagaaggct gccagttgga accaggcccc caaactccac 600

tactgcctgg actatgactg tcagaaggca gaattgtttg tgactcgcct 650

ggaagctgtg accagcaacc acgacggagg ctgtgactgc tacgtccaag 700

ggagtgtggc caataggacc ggctctgtgg aggctcagac agccctaaag 750

aagcggcagc tgcacaccac ctgggaggag ggcctggtgc tccccctggc 800

ggaggaggag ctccccacag ccaccctgac gctgaccttg aggacctgcg 850

accgcttctc ccgtcacagc gtggccgggg agctccgcct gggcctggac 900

gggacatctg tgcctctagg ggctgcccag tggggcgagc tgaagacttc 950

agcgaaggag ccatctgcag gagctggaga ggtcctacta tccatcagct 1000

acctcccggc tgccaaccgc ctcctggtgg tgctgattaa agccaagaac 1050

ctccactcta accagtccaa ggagctcctg gggaaggatg tctctgtcaa 1100

ggtgaccttg aagcaccagg ctcggaagct gaagaagaag cagactaaac 1150

gagctaagca caagatcaac cccgtgtgga acgagatgat catgtttgag 1200

ctgcctgacg acctgctgca ggcctccagt gtggagctgg aagtgctggg 1250

ccaggacgat tcagggcaga gctgtgcgct tggccactgc agcctgggcc 1300

tgcacacctc gggctctgag cgcagccact gggaggagat gctcaaaaac 1350

cctcgccggc agattgccat gtggcaccag ctgcacctgt aaccagctgc 1400

ccagctgcct cccttcttgg acagccctga cccgtcctct gcaacctcct 1450

ttctgtgccc ctttcctcat tctgacaccc agaagacagt gacagatgtg 1500

tttgcaaggc tgggatggct ctctcatcat actcttgttt cttagaaata 1550

agcaagacag agcaggaaat ggaatatgcg ggtcacactg aggaatgcat 1600

tttgctcatc tgtgttattg aaggaggtgc ttattaaata cagttcctat 1650

gcctgtttta taggtggggt taggccagat gcagagaaag ctaaatgtgg 1700

gaatcatgga tgcaaagaag aatttggctt tttgaaaaac aagcatttca 1750

aaaatgatga aggaagtgaa agtatcctgg atcaactcct agagttagag 1800

attgcccagg tggaaagaaa ccttagccag cgttcaatca agctcaccat 1850

gcagggcagt cacccggcag ttctcaaact ttagcatgtg aagagtcacc 1900

agcagattcc tgggctcgcc tggagacatt cctagtcggt attcctggtc 1950

gaagcccagg agccttcctt tttaacaagc tgatgtagag ggtggagcac 2000

tgtatgtgga gaaattcctt ctacaatatt ccacacaggt ttttggccac 2050

agtccttgat ggagtcccaa aaccatggtg cagccagttc caatgctgga 2100

cacctcaacc atcagggtga aatctggggc ctcagctttt taatttaatt 2150

attttaattc ttaatacttt aatttgtgca tttcataagc cccctgctct 2200

tggactgaat tttgtgcttt ttattgaaga attttattgt ttttatctta 2250

aaatcagttt ctattatcct tggggagacc atccctaaca aagtacaggt 2300

gggatctcct gtgagtcatt ggctgggttc tgattgctag atgtcacacc 2350

caccagcatc accaaagtga ctctgagata gaccggtccc ttctcagcgt 2400

tccagtcact tcaggaggaa tttagttatt gacttagtct atgacatctg 2450

gctacatgta ggtagagaag aaagacaatt ttaaaaagga aatcaggtct 2500

tttgcaactg tgcctccctc tgtctgtttt cacttgaatg ggtaaataac 2550

cagcagctag gttttgaatt cctaccttgt tattctaaac agatgtccac 2600

attgttaatt aaatctaaat tatgagcctt gctgagtgga tacggtactt 2650

acacctgaac caggattcct gggttctgtt gttgacattg cccttcagca 2700

cctgtttggc cagctgtata agataggact aatgactagg aagcctaccc 2750

caatgaatga tatactagat gaaatagtgt tcaaaacctg taggcactct 2800

ctggctaaaa acaaactctg aggccaccag cagatcatct ttaagctaag 2850

ttactatttt tcacctttt ttttagacgg agttttgctc tttgttgccc 2900

aggctggagt gcagtggcac gatctcggct cactgcaacc tccgcctccc 2950

aagttcaagc gattctcctg tctcagcctc ctgggtagct gggattacag 3000

gtgcccacca acatgcctgg ctaatttttg tacttttagt agagatgggg 3050

tttcaccatg ttggccaggc tggtcttcaa ctccagatct caggtgatct 3100

accctcctcg gcctcccaaa gtactgggat tacaggcctg agccaccgcg 3150

cccggcctat ttttcacttt aatttggcag ctgagaatgc ccaaaaagtg 3200

ccagaagcat cgtggcattt ccagaaccat ggattctgcc tttggacccc 3250

tctctattaa tattaaaact ctgggccttc agatgtcacc ctaatccact 3300

gccctaagac agaatttctg gacaagatgg gtaagggctt cattccttca 3350

acaagtcaag tcatacttgg cctctccctg agaatctgag caggagcctt 3400

ataacctgtg gtcattattt tttctttctg tacagaaata gaaaagcatt 3450

agaaataact tctaaccatc ctctgaaaaa acagaaaaaa tatcgaatcc 3500

ctctttcatg agaagtcttt tggataattg gaaaccttca tcactgaggt 3550

tggccagccc ctgccaagtg ttgtgtaggc aaagcacttg ttagtggctt 3600

cctatgaaat gttttagaga tctcttcacc atactggttt cttctctttg 3650

gttggtgtgg gtaaaagaaa acaaaacatt tcctataagc tgaaagctga 3700

ccagcattct cttcttggta acatctacta ctccaaccta gaaaatttgg 3750

attctagacc aaaaatcagg aaacatggct ccttataaat ctgtgcagct 3800

gccttatagt accatcaaag gaatttcagg tgggctgggc ggggccccga 3850

tcccagaatt atcaactcca cccatcatca tttggtcatg aagcatcctt 3900

tcattcttct tcttcttttt tttgggggggg gcggggcggg ggagggatct 3950

caaagtttta gtcttccaga atccaaatta aaggttgccc ctgatggggg 4000

ccaggttccg ccacagaaca tcttagatgt cagccttgac ctcacttagc 4050

agggattaca gaaatgagat acattttgaa ggagagttgt ctgttatgtt 4100

cactgtattc taagtgcctg ggataaagct gtctcatggg tgctccatat 4150

atattcatat atatttgttg agtgaattaa tgaattaaga gtggctggca 4200

gagtaggcag aaaaagcacc tgcaaatggc ataaaaatta aagtcctagc 4250

tgagttctca atggtaaagg catcagatgt cttagcagtc aagctagaaa 4300

ttcatgacaa tgagtattac tatttgccta atgacaactc attgctctcc 4350

atgtaaatgt aatcaacaga tgaagagaat ataattgctc tgcttttcca 4400

ctaaaactcc atcttagtga attttaaatt atccagagat gtcaaactgc 4450

caaataaaaa tatttcagta gtctttgcat cagcttacct tgtaccagaa 4500

acatttccaa tttactatca aattatagta actgagcctg tgtgaagtat 4550

ctcatcattt tcgaaaggaa caccttgtgt gatgccagtg agcatttcta 4600

aaaagggtgt gaggtagagg taaaaataag gtgagagacc atttcagaat 4650

gcactgttgc tcaaaaaggt gatctggttc tttcttcaga gatttctacg 4700

gggatagaaa atcgggagtc tgccctcatt aatctgtgac tccacctctt 4750

gcatcaaatc aatatctatt tgttgagcac ttattgatta agaccttgca 4800

tatgtctgtc cattttgatt tgagatacaa ctttttgtgt gggttgaatg 4850

acaaatcact ccaaacaaaa ctgggcacag agaatcagct aggagaccag 4900

ttattcaggg tccatttctc ttggatgtaa aggagtcctg ggtaaaatgt 4950

ggctgtaacc taaaccaact agtccttgtg atttgtttct gccctctgtg 5000

tttcctgttg tcaaatgcta agtgtgtgtt ttgcagtcat gaactaaagc 5050

acaaaaagat gcatgagaca ttgtagtcat atgtctggtg tgacactttg 5100

gagcaaaaac cttgcagtgg taaataaaaa atttccaaca gggaaaaaaa 5150

aaaaaaaaaa aaaaaaaaaa 5170

<210> 63
<211> 5563
<212> DNA
<213> Homo sapiens

<400> 63
agcggcagct ccgctgaaaa ctgcattcag ccagtcctcc ggacttctgg 50

agcggggaca gggcgcaggg catcagcagc caccagcagg acctgggaaa 100

tagggattct tctgcctcca cttcaggttt tagcagcttg gtgctaaatt 150

gctgtctcaa aatgcagagg atctaatttg cagaggaaaa cagccaaaga 200

aggaagagga ggaaaaggaa aaaaaaaggg gtatattgtg gatgctctac 250

ttttcttgga aatgcaaaag attatgcata tttctgtcct cctttctcct 300

gttttatggg gactgatttt tggtgtctct tctaacagca tacagatagg 350

ggggctattt cctaggggcg ccgatcaaga atacagtgca tttcgagtag 400

ggatggttca gttttccact tcggagttca gactgacacc ccacatcgac 450

aatttggagg tggcaaacag cttcgcagtc actaatgctt tctgctccca 500

gttttcgaga ggagtctatg ctatttttgg attttatgac aagaagtctg 550

taaataccat cacatcattt tgcggaacac tccacgtctc cttcatcact 600

cccagcttcc caacagatgg cacacatcca tttgtcattc agatgagacc 650

cgacctcaaa ggagctctcc ttagcttgat tgaatactat caatgggaca 700

agtttgcata cctctatgac agtgacagag gcttatcaac actgcaagct 750

gtgctggatt ctgctgctga aaagaaatgg caagtgactg ctatcaatgt 800

gggaaacatt aacaatgaca agaaagatga gatgtaccga tcacttttc 850

aagatctgga gttaaaaaag gaacggcgtg taattctgga ctgtgaaagg 900

gataaagtaa acgacattgt agaccaggtt attaccattg gaaaacacgt 950

taaagggtac cactacatca ttgcaaatct ggaatttact gatggagacc 1000

tattaaaaat ccagtttgga ggtgcaaatg tctctggatt tcagatagtg 1050

gactatgatg attcgttggt atctaaattt atagaaagat ggtcaacact 1100

ggaagaaaaa gaataccctg gagctcacac aacaacaatt aagtatactt 1150

ctgctctgac ctatgatgcc gttcaagtga tgactgaagc cttccgcaac 1200

ctaaggaagc aaagaattga aatctcccga agggggaatg caggagactg 1250

tctggcaaac ccagcagtgc cctggggaca aggtgtagaa atagaaaggg 1300

ccctcaaaca ggttcaggtt gaaggtctct caggaaatat aaagtttgac 1350

cagaatggaa aaagaataaa ctatacaatt aacatcatgg agctcaaaac 1400

taatgggccc cggaagattg gctactggag tgaagtggac aaaatggttg 1450

ttacccttac tgagctccct tctggaaatg acacctctgg gcttgagaat 1500

aagactgttg ttgtcaccac aattttggaa tctccgtatg ttatgatgaa 1550

gaaaaatcat gaaatgcttg aaggcaatga gcgctatgag ggctactgtg 1600

ttgacctggc tgcagaaatc gccaaacatt gtgggttcaa gtacaagttg 1650

1440
1422

acaattgttg gtgatggcaa gtatggggcc agggatgcag acacgaaaat 1700

ttggaatggg atggttggag aacttgtata tgggaaagct gatattgcaa 1750

ttgctccatt aactattacc cttgtgagag aagaggtgat tgacttctca 1800

aagcccttca tgagcctcgg gatatctatc atgatcaaga agcctcagaa 1850

gtccaaacca ggagtgtttt cctttcttga tcctttagcc tatgagatct 1900

ggatgtgcat tgtttttgcc tacattgggg tcagtgtagt tttattcctg 1950

gtcagcagat ttagccccta cgagtggcac actgaggagt ttgaagatgg 2000

aagagaaaca caaagtagtg aatcaactaa tgaatttggg atttttaata 2050

gtctctggtt ttccttgggt gcctttatgc ggcaaggatg cgatatttcg 2100

ccaagatccc tctctgggcg cattgttgga ggtgtgtggt ggttctttac 2150

cctgatcata atctcctcct acacggctaa cttagctgcc ttcctgactg 2200

tagagaggat ggtgtctccc atcgaaagtg ctgaggatct ttctaagcaa 2250

acagaaattg cttatggaac attagactct ggctccacta aagagttttt 2300

caggagatct aaaattgcag tgtttgataa aatgtggacc tacatgcgga 2350

gtgcggagcc ctctgtgttt gtgaggacta cggccgaagg ggtggctaga 2400

gtgcggaagt ccaaagggaa atatgcctac ttgttggagt ccacgatgaa 2450

cgagtacatt gagcaaagga agccttgcga caccatgaaa gttggtggaa 2500

acctggattc caaaggctat ggcatcgcaa cacctaaagg atcctcatta 2550

ggaaccccag taaatcttgc agtattgaaa ctcagtgagc aaggcgtctt 2600

agacaagctg aaaaacaaat ggtggtacga taaaggtgaa tgtggagcca 2650

aggactctgg aagtaaggaa aagaccagtg ccctcagtct gagcaacgtt 2700

gctggagtat tctacatcct tgtcgggggc cttggtttgg caatgctggt 2750

ggctttgatt gagttctgtt acaagtcaag ggccgaggcg aaacgaatga 2800

aggtggcaaa gaatgcacag aatattaacc catcttcctc gcagaattca 2850

cagaattttg caacttataa ggaaggttac aacgtatatg gcatcgaaag 2900

tgttaaaatt taggggatga ccttgaatga tgccatgagg aacaaggcaa 2950

ggctgtcaat tacaggaagt actggagaaa atggacgtgt tatgactcca 3000

gaatttccca aagcagtgca tgctgtccct tacgtgagtc ctggcatggg 3050

aatgaatgtc agtgtgactg atctctcgtg attgataaga acctttgag 3100

tgccttacac aatggttttc ttgtgtgttt attgtcaaag tggtgagagg 3150

catccagtat cttgaagact tttctttcag ccaagaattc ttaaatatgt 3200

ggagttcatc ttgaattgta aggaatgatt aattaaaaca caacatcttt 3250

ttctactcga gttacagaca aagcgtggtg gacatgcaca gctaacatgg 3300

aagtactata atttacctga agtctttgta cagacaacaa acctgtttct 3350

gcagccacta ttgttagtct cttgattcat aatgacttaa gcacacttga 3400

catcaactgc atcaagatgt gacatgtttt ataaaaaaag gaaaaaaaac 3450

atttaaaact aaaaaatatt tttaggtatt ttcacaaaca aactggcttt 3500

taaataaatt tgcttccata ttggttgaat aagacaaaaa caattaaact 3550

gagtgggaag tgaataaaaa aaggctttag gtatcgattc catatttttc 3600

aaagccaaat atgtaaatgc taaggaaagt aaacaaagag gagattccaa 3650

tcttgtaatt taatattgtt attaaaactt taatgtatcc tattctttaa 3700

catttggtgt taatataaaa ttacttggca atgcttgaca tttgaaataa 3750

acatttttct attgttttat tgcaagtggt ccaattaatt ttgcttagct 3800

acagtttggt cataaatcaa gtgagtttaa agacactacc aagttgttag 3850

gtgcccagag aaaatttctc ccttttaaaa aggccaggtg atttttcaaa 3900

tgtaatcttg cccccaaagt aatatccgaa tatctttttg acatgtctaa 3950

atatatatat atataaagaa atatttgtta acacaaaagc atttgatcta 4000

tgtagataaa tgctaataga tttaaaaagc taatattaac aaataccaga 4050

atacgtgaag ttccattttt aaagtgtttg agcttacaga agagaaacat 4100

tcattttaaa tgaagtaaaa aatgccttga aagtaattct ttagatagtt 4150

gcccattgat taaattccaa aaactaaata tgtttttagc tttaaaatta 4200

taaaagctgt cataaacttt atatattatg aattttaaaa tatgtttgag 4250

tctcctgcaa tatagtttca tcccattgac atcaattaaa aataacccta 4300

atatattatt tttatattta ttcctcaggt ggaatggcta ttttaatatg 4350

cccagtgtgg ataaaatgtc acatttctgt aacttttgac taaagagcct 4400

atatttatct agttaatgaa tttaaaggat ctatctttcc cttcataaaa 4450

tacctcttat ttccattaaa gccccccaag tttaattaat ttaggatttt 4500

gaatgattat tgacatccaa tagttatttt taatatttgt attcttgtta 4550

tttctggaag aaagcctttg tgtagcactt ggtattttgc aaagtgcttt 4600

taaaacattc ttacttaccg tatttcatag aagggaagga aaaatgtaag 4650

gtttaacagt aagcacttgc attgaacatg gaggcatgtg gtatcatgat 4700

attcttcact aaatttagct gtccctaatc acagatccta aggtaatata 4750

atataatttt agtgcatttc tcctcatcag gaatgctgga ggtgcatttt 4800

aagttttaat aataagtgct agaatgacca aattgcagac taattgtttc 4850

catattgtac ttaaaatgag tttttaaaag tgaaaaagaa atgactatat 4900

acaatcaatg ctatttattg tacctctggg cctactcttc taaaaattgt 4950

agcttatcga tttttctctg tcaagcttga actaatgtaa ataattgaaa 5000

taatgtaaag ttatattttc atgtttttat agatacaaca tgacaagaat 5050

acataatgta agagtatttc aactatggat aatgttgatt ggataatgca 5100

catctcagtt acaagcagta ctcatagttt aatatccatg taacggtgca 5150

tcaatatatt gctatataaa tatgtctgtg tgcatataag tgaaaagtgg 5200

tcaaacaaga gtgatgacag ctgtctaaag gttttttttat tcattttata 5250

taaaaactgt tatggaaaga ccaaaatgtt tatgaactat tcttatgtaa 5300

atttacaatt gtcctttact gtacttttt gtttacagta tagtaccttа 5350

ttttctgctg tgttaagtgg gtgtcaaact ccaagaagac atacactttc 5400

tataacttct attgaagata ttggaatttc caatttttca tgtgtactat 5450

1420
1424

gtcagaaaat gctttcgatt ttatttttaa atctaacatc ggatggcttt 5500

tccggagtgt tgtaaaaact tcaatcatac ataaaacatg ttcttacaaa 5550

aggcaaagat ctt 5563

<210> 64
<211> 1460
<212> DNA
<213> Homo sapiens

<400> 64
gccaacactg gccaaaaggg atagctgtcc aaggtctccc ccagcactga 50

ggagctcgcc tgctgccctc ttgcgcgcgg gaagcagcac caagttcacg 100

gccaacgcct tggcactagg gtccagaatg gctacaacag tccctgatgg 150

ttgccgcaat ggcctgaaat ccaagtacta cagactttgt gataaggctg 200

aagcttgggg catcgtccta gaaacggtgg ccacagccgg ggttgtgacc 250

tcggtggcct tcatgctcac tctcccgatc ctcgtctgca aggtgcagga 300

ctccaacagg cgaaaaatgc tgcctactca gtttctcttc ctcctgggtg 350

tgttgggcat ctttggcctc accttcgcct tcatcatcgg actggacggg 400

agcacagggc ccacacgctt cttcctcttt gggatcctct tttccatctg 450

cttctcctgc ctgctggctc atgctgtcag tctgaccaag ctcgtccggg 500

ggaggaagcc cctttccctg ttggtgattc tgggtctggc cgtgggcttc 550

agcctagtcc aggatgttat cgctattgaa tatattgtcc tgaccatgaa 600

taggaccaac gtcaatgtct tttctgagct ttccgctcct cgtcgcaatg 650

aagactttgt cctcctgctc acctacgtcc tcttcttgat ggcgctgacc 700

ttcctcatgt cctccttcac cttctgtggt tccttcacgg gctggaagag 750

acatggggcc cacatctacc tcacgatgct cctctccatt gccatctggg 800

tggcctggat caccctgctc atgcttcctg actttgaccg caggtgggat 850

gacaccatcc tcagctccgc cttggctgcc aatggctggg tgttcctgtt 900

ggcttatgtt agtcccgagt tttggctgct cacaaagcaa cgaaacccca 950

tggattatcc tgttgaggat gctttctgta aacctcaact cgtgaagaag 1000

agctatggtg tggagaacag agcctactct caagaggaaa tcactcaagg 1050

ttttgaagag acaggggaca cgctctatgc cccctattcc acacattttc 1100

agctgcagaa ccagcctccc caaaaggaat tctccatccc acgggcccac 1150

gcttggccga gcccttacaa agactatgaa gtaaagaaag agggcagcta 1200

actctgtcct gaagagtggg acaaatgcag ccgggcggca gatctagcgg 1250

gagctcaaag ggatgtgggc gaaatcttga gtcttctgag aaaactgtac 1300

aagacactac gggaacagtt tgcctccctc ccagcctcaa ccacaattct 1350

tccatgctgg ggctgatgtg ggctagtaag actccagttc ttagaggcgc 1400

tgtagtattt ttttttttt gtctcatcct ttagatactt cttttaagtg 1450

ggagtctcag 1460

<210> 65
<211> 1304
<212> DNA

<213> Homo sapiens

<220>
<221> unsure
<222> 1144-1145
<223> unknown base

<400> 65
tcattccatt aaacatattt ctaaataata gtaagtggta ctaacaaaat 50

aaataataat ttaatagcct tagaaataaa tgactgtata cttatacagg 100

ttgaaaaaaa ctcggtagga ataagttacc tttttgttta ctaatgttgg 150

tttcaaaaat actcagattc attttagttg gctgacatct ggaagtagtt 200

aacaactaac cagtggactt caacaatcat ttgctcccag gcttccccca 250

tcatcaccct caccacatat cctgctaata tccaacaaca aacaaatatt 300

taatattgaa atagcccatt gcctgagaat gaacacaagc taaaatacat 350

gcaagggtac ttaatggaag ccaaaccatg ttctatacct aaggagaaaa 400

catggacatg tagaatgctt ttattcatgt attcaaaatc agaacaaatc 450

agtgtatatc actagaacat cagatggagg ataacacaag aagtgataca 500

gatcagggtt cacttctctt accctctctc tgttaggacc acattcctat 550

tttagccaaa tgtttctggt acgggccatc ttttcaccat aaatggcatt 600

atgtttcaaa tggctaaaag cttattatgg gtatgctcaa aggagtaaaa 650

cccaattcac agagatgtgg cttttctaaa gaaccaaatt aagggaatac 700

ttagtatcta caactaaagt attagtcatt ctgaggatta tttgttgctt 750

taagtattag ctccctacta catccataac accctactta ctaaatttaa 800

ttacacacaa cttttcaaga attcgtattt tatttgaagg gaggtacctg 850

tctactttat ctacaataaa aacaaagggt attggctttc tctaatccat 900

gcaaactaca aattccatcg ggagtcctac atcactaaca gtggtatgaa 950

caaaactaag aaagtacttc ctatccacat gtgaatgttt taaaaaagtt 1000

tttgccataa aacctaagtg taatttagca taccacagtg ctctgaagat 1050

gggtcattga cgatgtacca tttgtatata ggtaatacac atgtaaatca 1100

ctaattgtta attaaatata agaagtaggt ttttatcttt ttannaaaaa 1150

aacaaaaaag gtgactccct ttctcattct gttctgttgt attgcgtcca 1200

aaagttgtgt gtaatttttt taaggtccaa gaaatgtaaa gaaatttgtt 1250

ggaatacctg aatttctgta aaaaaataaa aataaaaata agattttgtt 1300

actt 1304

<210> 66
<211> 3268
<212> DNA
<213> Homo sapiens

<400> 66
gggaggctgt gccaggcgag ccggaggggt gctccgcgct cccccgccct 50

ccttccggga gcgaggatgc agactctgaa actggtgctg ctgggctgag 100

gcggaggcag gggagttgca gcgcgcgagg ctccgtgagt gtgtctcctg 150

cgcgctgaga ggcgggggga ggcggaggac caggaggagg aggaggagga 200

1426
~~1429~~

ggaggagggg gagaatgccc ggagccgccg ccgctgccgc cgccgccgcc 250

gccgcgatgc tcccggctca ggaggctgcc aagctgtacc acaccaacta 300

tgtgcggaac tcgcgggcca tcggcgtgct gtgggccatc ttcaccatct 350

gctttgccat cgtcaacgtg gtgtgcttca tccagcccta ctggataggc 400

gacggcgtgg acacccccgca agccggctat ttcgggctct tccactactg 450

catcggcaac ggcttctccc gggagctgac ctgcaggggc agcttcacgg 500

acttctccac gctgccctcg ggcgccttca aagccgcctc cttctttatc 550

ggcctctcca tgatgctcat cattgcctgc atcatttgct ttaccctctt 600

cttcttctgc aacacggcca ctgtgtacaa gatatgtgcc tggatgcagc 650

tcacctccgc tgcctgcctt gtgcttggct gtatgatttt ccctgatggc 700

tgggactcag atgaagtaaa acggatgtgt ggagaaaaga cagacaagta 750

cactcttggg gcttgctcag tccgctgggc atacatcctg gctattattg 800

gaattttgga tgccctgatc ctctcatttc tagcatttgt gcttggtaat 850

cgacaagaca gcttgatggc agaggaactg aaggcagaaa acaaagttct 900

gctaagccaa tattctctag aatgagcaca aaacaaatcg aataacagct 950

aaacaaatcg aataacagct aaacgaatcg aataacagct tttgtacatc 1000

aacatcaaga aggaatacgc ctgagagaga tcagagtata tagatgaata 1050

tgaacaagaa tggaacattc acttgtcaac gcactttcta aatctagatc 1100

agcagagatg ggagtgattt tctggaaaga gatgtgatca tggattaaac 1150

accagctcat tggaaactca ttggatgaga tcagaaaacg ttcatgaaaa 1200

atcatattca ggaaataagg aagaggaata taaatgctct agagttaaca 1250

tgtaaaatat atacgtactg aggtttgtaa actgtccttt ttaaatcaaa 1300

ctgaaaacaa aaagctttaa cctttcaaca gaatttttaa aaaggcagtt 1350

agttctaaat tattcctatc tcaatagcca agaggctgat caagcgtcat 1400

ttattgagga agcatcttag aaaatgcctc tgaatgtttt cataggagcc 1450

gtgacctttg gttcttcatc tctaccattc atttacttca ctgtgtaatt 1500

agttacaacc actcagttat taagagacgt aacgcttcaa actttttacc 1550

aagtctgtgt tctgtttaat ctgtccatac aagttattac tgagaaagtg 1600

tttatgccat atactattac tccatcaagc tgtatattac aggaagtaca 1650

tctttacatc ataggttccc aagcaacata gatttcccta tctttcagga 1700

aacagcatca aggaactctg aaaaatatag aaaaagttca ttttcacctt 1750

ggaagctcac gtgtaatatt ataggctact atcaaataaa cacttttttt 1800

ctaattctcc ctagtatatg cataggaatt taatatactt tataaataag 1850

tatctaaaat gtctcctact tttttcctat ttctttgcca tacatgttat 1900

cagaaatcca tgtcttctat ttcccttact gatgggcact catttttatt 1950

tttttaaaaa tcattccatt aaacatattt ctaaataata gtaagtggta 2000

ctaacaaaat aaataataat ttaatagcct tagaaataaa tgactgtata 2050

cttatacagg ttgaaaaaaa ctcggtagga ataagttacc tttttgttta 2100

ctaatgttgg tttcaaaaat actcagattc attttagttg gctgacatct 2150

ggaagtagtt aacaactaac cagtggactt caacaatcat ttgctcccag 2200

gcttccccca tcatcaccct caccacatat cctgctaata tccaacaaca 2250

aacaaatatt taatattgaa atagcccatt gcctgagaat gaacacaagc 2300

taaaatacat gcaagggtac ttaatggaag ccaaaccatg ttctatacct 2350

aaggagaaaa catggacatg tagaatgctt ttattcatgt attcaaaatc 2400

agaacaaatc agtgtatatc actagaacat cagatggagg ataacacaag 2450

aagtgataca gatcagggtt cacttctctt accctctctc tgttaggacc 2500

acattcctat tttagccaaa tgtttctggt acgggccatc ttttcaccat 2550

aaatggcatt atgtttcaaa tggctaaaag catattatgg gtatgctcaa 2600

aggagtaaaa cccaattcac agagatgtgg cttttctaaa gaaccaaatt 2650

aagggaatac ttagtatcta caactaaagt attagtcatt ctgaggatta 2700

tttgttgctt taagtattag ctccctacta catccataac accctactta 2750

ctaaatttaa ttacacacaa cttttcaaga attcgtattt tatttgaagg 2800

gaggtacctg tctactttat ctacaataaa aacaaagggt attggctttc 2850

tctaatccat gcaaactaca aattccatcg ggagtcctac atcactaaca 2900

gtggtatgaa caaaactaag aaagtacttc ctatccacat gtgaatgttt 2950

taaaaaagtt tttgccataa aacctaagtg taatttagca taccacagtg 3000

ctctgaagat gggtcattga cgatgtacca tttgtatata ggtaatacac 3050

atgtaaatca ctaattgtta attaaatata agaagtaggt ttttatcttt 3100

ttaaaaaaaa aacaaaaaag gtgactccct ttctcattct gttctgttgt 3150

attgcgtcca aaagttgtgt gtaatttttt taaggtccaa gaaatgtaaa 3200

gaaatttgtt ggaatacctg aatttctgta aaaaaataaa aataaaaata 3250

agattttgtt acttaaaa 3268

<210> 67
<211> 3268
<212> DNA
<213> Homo sapiens

<400> 67
gggaggctgt gccaggcgag ccggaggggt gctccgcgct cccccgccct 50

ccttccggga gcgaggatgc agactctgaa actggtgctg ctgggctgag 100

gcggaggcag gggagttgca gcgcgcgagg ctccgtgagt gtgtctcctg 150

cgcgctgaga ggcgggggga ggcggaggac caggaggagg aggaggagga 200

ggaggagggg gagaatgccc ggagccgccg ccgctgccgc cgccgccgcc 250

gccgcgatgc tcccggctca ggaggctgcc aagctgtacc acaccaacta 300

tgtgcggaac tcgcgggcca tcggcgtgct gtgggccatc ttcaccatct 350

gctttgccat cgtcaacgtg gtgtgcttca tccagcccta ctggataggc 400

gacggcgtgg acaccccgca agccggctat ttcgggctct tccactactg 450

catcggcaac ggcttctccc gggagctgac ctgcaggggc agcttcacgg 500

acttctccac gctgccctcg ggcgccttca aagccgcctc cttctttatc 550

1428
~~1424~~

ggcctctcca tgatgctcat cattgcctgc atcatttgct ttaccctctt 600

cttcttctgc aacacggcca ctgtgtacaa gatatgtgcc tggatgcagc 650

tcacctccgc tgcctgcctt gtgcttggct gtatgatttt ccctgatggc 700

tgggactcag atgaagtaaa acggatgtgt ggagaaaaga cagacaagta 750

cactcttggg gcttgctcag tccgctgggc atacatcctg gctattattg 800

gaattttgga tgccctgatc ctctcatttc tagcatttgt gcttggtaat 850

cgacaagaca gcttgatggc agaggaactg aaggcagaaa acaaagttct 900

gctaagccaa tattctctag aatgagcaca aaacaaatcg aataacagct 950

aaacaaatcg aataacagct aaacgaatcg aataacagct tttgtacatc 1000

aacatcaaga aggaatacgc ctgagagaga tcagagtata tagatgaata 1050

tgaacaagaa tggaacattc acttgtcaac gcactttcta aatctagatc 1100

agcagagatg ggagtgattt tctggaaaga gatgtgatca tggattaaac 1150

accagctcat tggaaactca ttggatgaga tcagaaaacg ttcatgaaaa 1200

atcatattca ggaaataagg aagaggaata taaatgctct agagttaaca 1250

tgtaaaatat atacgtactg aggtttgtaa actgtccttt ttaaatcaaa 1300

ctgaaaacaa aaagctttaa cctttcaaca gaatttttaa aaaggcagtt 1350

agttctaaat tattcctatc tcaatagcca agaggctgat caagcgtcat 1400

ttattgagga agcatcttag aaaatgcctc tgaatgtttt cataggagcc 1450

gtgacctttg gttcttcatc tctaccattc atttacttca ctgtgtaatt 1500

agttacaacc actcagttat taagagacgt aacgcttcaa acttttttacc 1550

aagtctgtgt tctgtttaat ctgtccatac aagttattac tgagaaagtg 1600

tttatgccat atactattac tccatcaagc tgtatattac aggaagtaca 1650

tctttacatc ataggttccc aagcaacata gatttcccta tctttcagga 1700

aacagcatca aggaactctg aaaaatatag aaaaagttca ttttcacctt 1750

ggaagctcac gtgtaatatt ataggctact atcaaataaa cacttttttt 1800

ctaattctcc ctagtatatg cataggaatt taatatactt tataaataag 1850

tatctaaaat gtctcctact ttttttcctat ttctttgcca tacatgttat 1900

cagaaatcca tgtcttctat ttcccttact gatgggcact catttttatt 1950

tttttaaaaa tcattccatt aaacatattt ctaaataata gtaagtggta 2000

ctaacaaaat aaataataat ttaatagcct tagaaataaa tgactgtata 2050

cttatacagg ttgaaaaaaa ctcggtagga ataagttacc tttttgttta 2100

ctaatgttgg tttcaaaaat actcagattc attttagttg gctgacatct 2150

ggaagtagtt aacaactaac cagtggactt caacaatcat ttgctcccag 2200

gcttcccccа tcatcaccct caccacatat cctgctaata tccaacaaca 2250

aacaaatatt taatattgaa atagcccatt gcctgagaat gaacacaagc 2300

taaaatacat gcaagggtac ttaatggaag ccaaaccatg ttctatacct 2350

aaggagaaaa catggacatg tagaatgctt ttattcatgt attcaaaatc 2400

agaacaaatc agtgtatatc actagaacat cagatggagg ataacacaag 2450

1429
1425

gctaagccaa tattctctag aatgagcaca aaacaaatcg aataacagct 950

aaacaaatcg aataacagct aaacgaatcg aataacagct tttgtacatc 1000

aacatcaaga aggaatacgc ctgagagaga tcagagtata tagatgaata 1050

tgaacaagaa tggaacattc acttgtcaac gcactttcta aatctagatc 1100

agcagagatg ggagtgattt tctggaaaga gatgtgatca tggattaaac 1150

accagctcat tggaaactca ttggatgaga tcagaaaacg ttcatgaaaa 1200

atcatattca ggaaataagg aagaggaata taaatgctct agagttaaca 1250

tgtaaaatat atacgtactg aggtttgtaa actgtccttt ttaaatcaaa 1300

ctgaaaacaa aaagctttaa cctttcaaca gaatttttaa aaaggcagtt 1350

agttctaaat tattcctatc tcaatagcca agaggctgat caagcgtcat 1400

ttattgagga agcatcttag aaaatgcctc tgaatgtttt cataggagcc 1450

gtgacctttg gttcttcatc tctaccattc atttacttca ctgtgtaatt 1500

agttacaacc actcagttat taagagacgt aacgcttcaa actttttacc 1550

aagtctgtgt tctgtttaat ctgtccatac aagttattac tgagaaagtg 1600

tttatgccat atactattac tccatcaagc tgtatattac aggaagtaca 1650

tctttacatc ataggttccc aagcaacata gatttcccta tctttcagga 1700

aacagcatca aggaactctg aaaaatatag aaaaagttca ttttcacctt 1750

ggaagctcac gtgtaatatt ataggctact atcaaataaa cactttttt 1800

ctaattctcc ctagtatatg cataggaatt taatatactt tataaataag 1850

tatctaaaat gtctcctact tttttcctat ttctttgcca tacatgttat 1900

cagaaatcca tgtcttctat ttcccttact gatgggcact catttttatt 1950

tttttaaaaa tcattccatt aaacatattt ctaaataata gtaagtggta 2000

ctaacaaaat aaataataat ttaatagcct tagaaataaa tgactgtata 2050

cttatacagg ttgaaaaaaa ctcggtagga ataagttacc tttttgttta 2100

ctaatgttgg tttcaaaaat actcagattc attttagttg gctgacatct 2150

ggaagtagtt aacaactaac cagtggactt caacaatcat ttgctcccag 2200

gcttccccca tcatcaccct caccacatat cctgctaata tccaacaaca 2250

aacaaatatt taatattgaa atagcccatt gcctgagaat gaacacaagc 2300

taaaatacat gcaagggtac ttaatggaag ccaaaccatg ttctatacct 2350

aaggagaaaa catggacatg tagaatgctt ttattcatgt attcaaaatc 2400

agaacaaatc agtgtatatc actagaacat cagatggagg ataacacaag 2450

aagtgataca gatcagggtt cacttctctt accctctctc tgttaggacc 2500

acattcctat tttagccaaa tgtttctggt acgggccatc ttttcaccat 2550

aaatggcatt atgtttcaaa tggctaaaag catattatgg gtatgctcaa 2600

aggagtaaaa cccaattcac agagatgtgg cttttctaaa gaaccaaatt 2650

aagggaatac ttagtatcta caactaaagt attagtcatt ctgaggatta 2700

tttgttgctt taagtattag ctccctacta catccataac accctactta 2750

ctaaatttaa ttacacacaa cttttcaaga attcgtattt tatttgaagg 2800

1430
1426

aagtgataca gatcagggtt cacttctctt accctctctc tgttaggacc 2500

acattcctat tttagccaaa tgtttctggt acgggccatc ttttcaccat 2550

aaatggcatt atgtttcaaa tggctaaaag catattatgg gtatgctcaa 2600

aggagtaaaa cccaattcac agagatgtgg cttttctaaa gaaccaaatt 2650

aagggaatac ttagtatcta caactaaagt attagtcatt ctgaggatta 2700

tttgttgctt taagtattag ctccctacta catccataac accctactta 2750

ctaaatttaa ttacacacaa cttttcaaga attcgtattt tatttgaagg 2800

gaggtacctg tctactttat ctacaataaa aacaaagggt attggctttc 2850

tctaatccat gcaaactaca aattccatcg ggagtcctac atcactaaca 2900

gtggtatgaa caaaactaag aaagtacttc ctatccacat gtgaatgttt 2950

taaaaaagtt tttgccataa aacctaagtg taatttagca taccacagtg 3000

ctctgaagat gggtcattga cgatgtacca tttgtatata ggtaatacac 3050

atgtaaatca ctaattgtta attaaatata agaagtaggt ttttatcttt 3100

ttaaaaaaaa aacaaaaaag gtgactccct ttctcattct gttctgttgt 3150

attgcgtcca aaagttgtgt gtaatttttt taaggtccaa gaaatgtaaa 3200

gaaatttgtt ggaatacctg aatttctgta aaaaaataaa aataaaaata 3250

agattttgtt acttaaaa 3268

<210> 68
<211> 3268
<212> DNA
<213> Homo sapiens

<400> 68
gggaggctgt gccaggcgag ccggaggggt gctccgcgct cccccgccct 50

ccttccggga gcgaggatgc agactctgaa actggtgctg ctgggctgag 100

gcggaggcag gggagttgca gcgcgcgagg ctccgtgagt gtgtctcctg 150

cgcgctgaga ggcgggggga ggcggaggac caggaggagg aggaggagga 200

ggaggagggg gagaatgccc ggagccgccg ccgctgccgc cgccgccgcc 250

gccgcgatgc tcccggctca ggaggctgcc aagctgtacc acaccaacta 300

tgtgcggaac tcgcgggcca tcggcgtgct gtgggccatc ttcaccatct 350

gctttgccat cgtcaacgtg gtgtgcttca tccagcccta ctggataggc 400

gacggcgtgg acaccccgca agccggctat ttcgggctct tccactactg 450

catcggcaac ggcttctccc gggagctgac ctgcaggggc agcttcacgg 500

acttctccac gctgccctcg ggcgccttca aagccgcctc cttctttatc 550

ggcctctcca tgatgctcat cattgcctgc atcatttgct ttaccctctt 600

cttcttctgc aacacggcca ctgtgtacaa gatatgtgcc tggatgcagc 650

tcacctccgc tgcctgcctt gtgcttggct gtatgatttt ccctgatggc 700

tgggactcag atgaagtaaa acggatgtgt ggagaaaaga cagacaagta 750

cactcttggg gcttgctcag tccgctgggc atacatcctg gctattattg 800

gaattttgga tgccctgatc ctctcatttc tagcatttgt gcttggtaat 850

cgacaagaca gcttgatggc agaggaactg aaggcagaaa acaaagttct 900

gctaagccaa tattctctag aatgagcaca aaacaaatcg aataacagct 950

aaacaaatcg aataacagct aaacgaatcg aataacagct tttgtacatc 1000

aacatcaaga aggaatacgc ctgagagaga tcagagtata tagatgaata 1050

tgaacaagaa tggaacattc acttgtcaac gcactttcta aatctagatc 1100

agcagagatg ggagtgattt tctggaaaga gatgtgatca tggattaaac 1150

accagctcat tggaaactca ttggatgaga tcagaaaacg ttcatgaaaa 1200

atcatattca ggaaataagg aagaggaata taaatgctct agagttaaca 1250

tgtaaaatat atacgtactg aggtttgtaa actgtccttt ttaaatcaaa 1300

ctgaaaacaa aaagctttaa cctttcaaca gaatttttaa aaaggcagtt 1350

agttctaaat tattcctatc tcaatagcca agaggctgat caagcgtcat 1400

ttattgagga agcatcttag aaaatgcctc tgaatgtttt cataggagcc 1450

gtgacctttg gttcttcatc tctaccattc atttacttca ctgtgtaatt 1500

agttacaacc actcagttat taagagacgt aacgcttcaa acttttacc 1550

aagtctgtgt tctgtttaat ctgtccatac aagttattac tgagaaagtg 1600

tttatgccat atactattac tccatcaagc tgtatattac aggaagtaca 1650

tctttacatc ataggttccc aagcaacata gatttcccta tctttcagga 1700

aacagcatca aggaactctg aaaaatatag aaaaagttca ttttcacctt 1750

ggaagctcac gtgtaatatt ataggctact atcaaataaa cacttttttt 1800

ctaattctcc ctagtatatg cataggaatt taatatactt tataaataag 1850

tatctaaaat gtctcctact tttttcctat ttctttgcca tacatgttat 1900

cagaaatcca tgtcttctat ttcccttact gatgggcact catttttatt 1950

tttttaaaaa tcattccatt aaacatattt ctaaataata gtaagtggta 2000

ctaacaaaat aaataataat ttaatagcct tagaaataaa tgactgtata 2050

cttatacagg ttgaaaaaaa ctcggtagga ataagttacc tttttgttta 2100

ctaatgttgg tttcaaaaat actcagattc attttagttg gctgacatct 2150

ggaagtagtt aacaactaac cagtggactt caacaatcat ttgctcccag 2200

gcttccccca tcatcaccct caccacatat cctgctaata tccaacaaca 2250

aacaaatatt taatattgaa atagcccatt gcctgagaat gaacacaagc 2300

taaaatacat gcaagggtac ttaatggaag ccaaaccatg ttctatacct 2350

aaggagaaaa catggacatg tagaatgctt ttattcatgt attcaaaatc 2400

agaacaaatc agtgtatatc actagaacat cagatggagg ataacacaag 2450

aagtgataca gatcagggtt cacttctctt accctctctc tgttaggacc 2500

acattcctat tttagccaaa tgtttctggt acgggccatc ttttcaccat 2550

aaatggcatt atgtttcaaa tggctaaaag catattatgg gtatgctcaa 2600

aggagtaaaa cccaattcac agagatgtgg cttttctaaa gaaccaaatt 2650

aagggaatac ttagtatcta caactaaagt attagtcatt ctgaggatta 2700

tttgttgctt taagtattag ctccctacta catccataac accctactta 2750

ctaaatttaa ttacacacaa cttttcaaga attcgtattt tatttgaagg 2800

1432
1424

gaggtacctg tctactttat ctacaataaa aacaaagggt attggctttc 2850

tctaatccat gcaaactaca aattccatcg ggagtcctac atcactaaca 2900

gtggtatgaa caaaactaag aaagtacttc ctatccacat gtgaatgttt 2950

taaaaaagtt tttgccataa aacctaagtg taatttagca taccacagtg 3000

ctctgaagat gggtcattga cgatgtacca tttgtatata ggtaatacac 3050

atgtaaatca ctaattgtta attaaatata agaagtaggt ttttatcttt 3100

ttaaaaaaaa aacaaaaaag gtgactccct ttctcattct gttctgttgt 3150

attgcgtcca aaagttgtgt gtaatttttt taaggtccaa gaaatgtaaa 3200

gaaatttgtt ggaatacctg aatttctgta aaaaaataaa aataaaaata 3250

agattttgtt acttaaaa 3268

<210> 69
<211> 1755
<212> DNA
<213> Homo sapiens

<400> 69
caggcctctg aggctccctt gccgagggcc ccgagctgca gggacagtga 50

gcagtgagtc ccttgggcat cccgctcctg ggcaggtcac caataggtcc 100

ccgcagttcc caatggaact gttccagtcc tccccgaggc ctccacttca 150

acctgtctgt gtctgcccag gcctggagtt gtgtgaccct ccccaccgcc 200

tggccttctc catgggggct ggccttttct cggtggtggg caccctgctg 250

ctgcccggcc tggctgcgct tgtgcaggac tggcgtcttc tgcaggggct 300

gggtgccctg atgagtggac tcttgctgct cttttggggg ttcccggccc 350

tgttccccga gtctccctgc tggctgctgg ccacaggtca ggtagctcga 400

gccaggaaga tcctgtggcg ctttgcagaa gccagtggcg tgggccccgg 450

ggacagttcc ttggaggaga actccctggc tacagagctg accatgctgt 500

ctgcacggag ccccccagccc cggtaccact ccccactggg gcttctgcgt 550

acccgagtca cctggagaaa cgggcttatc ttgggcttca gctcgctggt 600

tggtggaggc atcagagcta gcttccgccg cagcctggca cctcaggtgc 650

cgaccttcta cctgccctac ttcctggagg ccggcctgga ggcggcagcc 700

ttggtcttcc tgctcctgac ggcagattgc tgtggacgcc gccccgtgct 750

gctgctgggc accatggtca caggcctggc atccctgctg ctcctcgctg 800

gggcccagta tctgccaggc tggactgtgc tgttcctctc tgtcctgggg 850

ctcctggcct cccgggctgt gtccgcactc agcagcctct tcgcggccga 900

ggtcttcccc acggtgatca gggggccgg gctgggcctg gtgctggggg 950

ccgggttcct gggccaggca gccggccccc tggacaccct gcacggccgg 1000

cagggcttct tcctgcaaca agtcgtcttc gcctcccttg ctgtccttgc 1050

cctgctgtgt gtcctgctgc tgcctgagag ccgaagccgg gggctgcccc 1100

agtcactgca ggacgccgac cgcctgcgcc gctccccact cctgcggggc 1150

cgcccccgcc aggaccacct gcctctgctg ccgccctcca actcctactg 1200

ggccggccac acccccgagc agcactagtc ctgcctggtg gccctgggag 1250

ccaggatggg accaaagtca aggcctgggg catggctgag taccccagac 1300

gtctggtcca gggcagacac attcctctca gaagcccgtg tctcagtgca 1350

ggtggagccg tggggacagc gtgaaggtgt ctccagccag gccccaggca 1400

ctgggaggcc ctgggtctcc ccccagccac acccagtagg tgtggaggat 1450

aaaggcttct gtggaacttg gcgtcctccc tggttcttcc ctcaggccac 1500

acctggggcc tggaagggtc tcacctgtgc ctgctcccca cctctgggct 1550

caccagcaag gagccagggt gtcagcggct gctccatagc ccctgtccag 1600

gtagtggcag ccccccggccg cctcctgccc cagccaagca gggtggtatg 1650

ggggcctggc cctcaggtcc ctgccaggcc aggcctcact ggccccatcc 1700

ccctcatggt cagtgtggca gtgggcgccc gtcctgcatg aaagagacta 1750

gcacg 1755

<210> 70
<211> 4021
<212> DNA
<213> Homo sapiens

<400> 70
agcggggcgg gtccccgaag cccagtacca gacctcctgt gcgaagagta 50

acctgctctt cggactgata actggatgga atggccttta gaaagggagc 100

atcttcagaa tctgcagctg ggatcagatt gttcttttgt tgaatcaact 150

cagctctgcc gagccagcat cttgctaagg gttacaaaat aaatgccatt 200

tgaacagtgc catctgtcat ggaaaaacgt gagacgtttg tacaatgccg 250

tgtctaagga gctggttgga gagtttttgc aatttgttca acttgataaa 300

gaggcctctg atcctttcag cctaaatgaa ttactagatg aattatcaag 350

gaaacagaaa gaagaattat ggcaaaggct gaagaattta ttgacagatg 400

tgttgttaga aagcccagtg gatgggtggc aggtagtgga agcccagggt 450

gaagacaata tggaaaccga acatggctca aaaatgagaa aaagcataga 500

aataatttat gcaattacat ctgtgattct tgcttctgtg tctgtaataa 550

atgaaagtga gaactacgaa gccctactgg aatgtgttat tatattaaat 600

ggtattttat atgcattacc tgagtctgaa cgaaaactac agagttctat 650

tcaggatttg tgtgttacct ggtgggagaa aggcctgcct gccaaggaag 700

acacaggaaa gacagccttt gtcatgttac taaggaggag tctggagact 750

aagacaggtg cagacgtatg tcggctttgg cgtatccatc aagctttata 800

ttgctttgat tatgatttgg aggaaagtgg agaaattaaa gatatgttac 850

ttgagtgctt cataaatatt aattatatca agaaagaaga gggaagaaga 900

tttcttagtt gtctcttcaa ctggaatatc aacttcatca aaatgatcca 950

cgggaccatt aaaaaccagt tacagggatt acaaaagtct ttgatggtat 1000

acattgcaga aatttatttc agagcttgga aaaaggcttc agggaaaata 1050

ctggaggcga ttgaaaatga ttgcatccag gacttcatgt tccacgggat 1100

acaccttccg aggaggtctc cagtgcattc caaagtgcgg gaggttctga 1150

gttactttca ccatcaaaag aaagttcggc agggagtgga agagatgctt 1200

tatagattat ataagcccat cctttggaga ggattaaagg ccagaaactc 1250

tgaagttcga tcaaatgctg cattgttgtt tgttgaagca tttcctatta 1300

gggatccaaa ccttcatgct attgaaatgg atagtraaat ccagaaacag 1350

tttgaagagc tctatagcct tttagaagat ccttacccga tggtccgttc 1400

cacagggatc cttggtgttt gtaaaataac ttctaagtac tgggaaatga 1450

tgcccccgac cattcttatt gacctcctga agaaggtgac tggggaactg 1500

gcatttgaca cgagctcagc tgatgttcgt tgttctgtct ttaagtgtct 1550

gccaatgatt ttggacaaca aactgagcca cccattgtta gagcagctcc 1600

ttccagctct cagatacagt ctccacgaca attcggagaa agtgagggta 1650

gcttttgtgg acatgctgtt gaagatcaaa gctgtgaggg ctgctaagtt 1700

ttggaaaata tgtcccatgg agcacattct ggttcgtctg gaaactgatt 1750

ctcgacctgt gtctcggcgc ctggtgagcc tcatctttaa ttctttcctg 1800

cctgtgaatc agccggagga ggtctggtgc gagcgctgtg tcaccctggt 1850

gcagatgaac cacgccgctg ccaggaggtt ctatcagtac gcccacgaac 1900

acaccgcctg caccaacata gcaaagctga ttcacgttat tcgtcattgc 1950

ttaaatgcct gtatccagag ggcagtgaga gagcctccag aggacgagga 2000

ggaagaggac ggaagggaga aggagaatgt gactgttctg gacaaaacac 2050

tgtcagtaaa cgatgttgca tgcatggcag gtttactaga aatcattgtg 2100

attctctgga aaagtattga cagaagtatg gaaaataata aagaggccaa 2150

actttacacg attaacaagt ttgcctctgt gcttccagag tatctgaaag 2200

tatttaagga tgatcgctgc aagatccctt tattcatgct aatgtccttt 2250

atgccggcct ctgctgtccc cccattcagc tgtggtgtga tttccacgct 2300

gagaagccgg gaggagggcg ctgtggacaa gagctactgc actttgttgg 2350

attgcctctg ctcctggggg caggtggggc acattctgga gcttgttgac 2400

aactggctgc ccacagagca tgcccaggcc aagagcaaca cagcttctaa 2450

aggtagggtg cagatccatg acacacgccc agtcaaacct gaattggcat 2500

tggtctacat tgagtatctg ctgactcatc caaagaaccg cgagtgcttg 2550

ctctctgctc ctcggaagaa acttaaccat ctttgaaag ccctgaaac 2600

gtcaaaggca gatctggagt cacttctgca gacaccgggt gggaagcctc 2650

gtggcttcag tgaagcagct gccccgcgag cctttggtct ccactgtcgc 2700

ctgagcatcc atcttcagca caagttctgc tcagaaggaa aggtgtattt 2750

gtccatgttg gaagacactg gcttttggtt agaaagcaaa attttatctt 2800

ttattcaaga tcaagaagaa gactacctga agcttcatag ggtcatttat 2850

cagcaaatta tccagaccta cctgactgtg tgtaaagatg ttgttatggt 2900

aggccttggt gaccatcagt ttcagatgca actcttacag cggagtcttg 2950

gaatcatgca aacagtgaag ggatttttt atgtttcatt acttcttgac 3000

attctgaaag agataactgg aagttccttg attcagaaaa cagattcaga 3050

tgaagaagtt gcaatgctgt tggacacagt ccagaaagta tttcagaaaa 3100

tgttggaatg tattgcacgg agcttcagga agcagccgga agaaggcctg 3150

cggctgcttt attctgttca gaggcctctt catgagttca ttactgctgt 3200

tcagtctcgg cacacagaca cccctgtgca ccggggtgta ctttctactc 3250

tgatcgctgg gcctgtggtt gagataagtc accagctacg gaaggtttct 3300

gacgtagaag agcttacccc tccagagcat ctttctgatc ttccaccatt 3350

ttcaaggtgt ttaataggaa taataataaa gtcttcgaat gtggtcaggt 3400

catttttgga tgaattaaag gcatgtgtgg cttctaatga tattgaaggc 3450

attgtgtgcc tcacggctgc tgtgcatatt atcctggtta ttaatgcagg 3500

taaacataaa agctcaaaag tgagggaggt tgcagccact gttcacagaa 3550

aactaaagac attcatggaa attactttgg aagaggatag cattgaaaga 3600

tttctctatg aatcatcatc aagaactctg ggagaacttt tgaattcata 3650

accaagccaa catctccaga catgtaaaaa tagggaaaag tgattcaaat 3700

tgaaatgcct gtgtattttc ctattgtttt taatgttaat aacccatata 3750

atagggaaag ggtgggattt ttttgtggga atgtgggaag gtgggggtta 3800

tggaggagat aactcaaaac ttcttcaatt ttgcctagtg cctgcgtaaa 3850

taatatattt aatataaagg actccaggta tgaatggtgt agaaatccat 3900

gattccaaga aaaaacactt ttctagcaaa cctggttgtt tttaaaatga 3950

cttttatata tgtaatattg cttggaaact atgagtaata aagcaatgac 4000

aacaamaaaa aaaaaaaaaa a 4021

<210> 71
<211> 3580
<212> DNA
<213> Homo sapiens

<400> 71
gcccttcaga atctgcagct gggatcagat tgttcttttg ttgaatcaac 50

tcagctctgc cgagccagca tcttgctaag ggttacaaaa taaatgccat 100

ttgaacagtg ccatctgtca tggaaaaacg tgagacgttt gtacaagccg 150

tgtctaagga gctggttgga gagtttttgc aatttgttca acttgataaa 200

gaggcctctg atcctttcag cctaaatgaa ttactagatg aattatcaag 250

gaaacagaaa gaagaattat ggcaaaggct gaagaattta ttgacagatg 300

tgttgttaga aagcccagtg gatgggtggc aggtagtgga agcccagggt 350

gaagacaata tggaaaccga acatggctca aaaatgagaa aaagcataga 400

aataatttat gcaattacat ctgtgattct tgcttctgtg tctgtaataa 450

atgaaagtga gaactacgaa gccctactgg aatgtgttat tatattaaat 500

ggtattttat atgcattacc tgagtctgaa cgaaaactac agagttctat 550

tcaggatttg tgtgttacct ggtgggagaa aggcctgcct gccaaggaag 600

acacaggaaa gacagccttt gtcatgttac taaggaggag tctggagact 650

aagacaggtg cagacgtatg tcggctttgg cgtatccatc aagctttata 700

ttgctttgat tatgatttgg aggaaagtgg agaaattaaa gatatgttac 750

ttgagtgctt cataaatatt aattatatca agaaagaaga gggaagaaga 800

1436
1429

tttcttagtt gtctcttcaa ctggaatatc aacttcatca aaatgatcca 850

cgggaccatt aaaaaccagt tacagggatt acaaaagtct ttgatggtat 900

acattgcaga aatttatttc agagcttgga aaaaggcttc agggaaaata 950

ctggaggcga ttgaaaatga ttgcatccag gacttcatgt tccacgggat 1000

acaccttccg aggaggtctc cagtgcattc caaagtgcgg gaggttctga 1050

gttactttca ccatcaaaag aaagttcggc agggagtgga agagatgctt 1100

tatagattat ataagcccat cctttggaga ggattaaagg ccagaaactc 1150

tgaagttcga tcaaatgctg cattgttgtt tgttgaagca tttcctatta 1200

gggatccaaa ccttcatgct attgaaatgg atagtgaaat ccagaaacag 1250

tttgaagagc tctatagcct tttagaagat ccttacccga tggtccgttc 1300

cacagggatc cttggtgttt gtaaaataac ttctaagtac tgggaaatga 1350

tgcccccgac cattcttatt gacctcctga agaaggtgac tggggaactg 1400

gcatttgaca cgagctcagc tgatgttcgt tgttctgtct ttaagtgtct 1450

gccaatgatt ttggacaaca aactgagcca cccattgtta gagcagctcc 1500

ttccagctct cagatacagt ctccacgaca attcggagaa agtgagggta 1550

gcttttgtgg acatgctgtt gaagatcaaa gctgtgaggg ctgctaagtt 1600

ttggaaaata tgtcccatgg agcacattct ggttcgtctg gaaactgatt 1650

ctcgacctgt gtctcggcgc ctggtgagcc tcatctttaa ttctttcctg 1700

cctgtgaatc agccggagga ggtctggtgc gagcgctgtg tcaccctggt 1750

gcagatgaac cacgccgctg ccaggaggtt ctatcagtac gcccacgaac 1800

acaccgcctg caccaacata gcaaagctga ttcacgttat tcgtcattgc 1850

ttaaatgcct gtatccagag ggcagtgaga gagcctccag aggacgagga 1900

ggaagaggac ggaagggaga aggagaatgt gactgttctg gacaaaacac 1950

tgtcagtaaa cgatgttgca tgcatggcag gtttactaga aatcattgtg 2000

attctctgga aaagtattga cagaagtatg gaaaataata aagaggccaa 2050

actttacacg attaacaagt ttgcctctgt gcttccagag tatctgaaag 2100

tatttaagga tgatcgctgc aagatccctt tattcatgct aatgtccttt 2150

atgccggcct ctgctgtccc cccattcagc tgtggtgtga tttccacgct 2200

gagaagccgg gaggagggcg ctgtggacaa gagctactgc actttgttgg 2250

attgcctctg ctcctggggg caggtggggc acattctgga gcttgttgac 2300

aactggctgc ccacagagca tgcccaggcc aagagcaaca cagcttctaa 2350

aggtagggtg cagatccatg acacacgccc agtcaaacct gaattggcat 2400

tggtctacat tgagtatctg ctgactcatc caaagaaccg cgagtgcttg 2450

ctctctgctc ctcggaagaa acttaaccat cttttgaaag cccttgaaac 2500

gtcaaaggca gatctggagt cacttctgca gacaccgggt gggaagcctc 2550

gtggcttcag tgaagcagct gccccgcgag cctttggtct ccactgtcgc 2600

ctgagcatcc atcttcagca caagttctgc tcagaaggaa aggtgtattt 2650

gtccatgttg gaagacactg gcttttggtt agaaagcaaa attttatctt 2700

1437
1479

ttattcaaga tcaagaagaa gactacctga agcttcatag ggtcatttat 2750

cagcaaatta tccagaccta cctgactgtg tgtaaagatg ttgttatggt 2800

aggccttggt gaccatcagt ttcagatgca actcttacag cggagtcttg 2850

gaatcatgca aacagtgaag ggatttttt atgtttcatt acttcttgac 2900

attctgaaag agataactgg aagttccttg attcagaaaa cagattcaga 2950

tgaagaagtt gcaatgctgt tggacacagt ccagaaagta tttcagaaaa 3000

tgttggaatg tattgcacgg agcttcagga agcagccgga agaaggcctg 3050

cggctgcttt attctgttca gaggcctctt catgagttca ttactgctgt 3100

tcagtctcgg cacacagaca cccctgtgca ccggggtgta ctttctactc 3150

tgatcgctgg gcctgtggtt gagataagtc accagctacg gaaggtttct 3200

gacgtagaag agcttacccc tccagagcat ctttctgatc ttccaccatt 3250

ttcaaggtgt ttaataggaa taataataaa gtcttcgaat gtggtcaggt 3300

catttttgga tgaattaaag gcatgtgtgg cttctaatga tattgaaggc 3350

attgtgtgcc tcacggctgc tgtgcatatt atcctggtta ttaatgcagg 3400

taaacataaa agctcaaaag tgagggaggt tgcagccact gttcacagaa 3450

aactaaagac attcatggaa attactttgg aagaggatag cattgaaaga 3500

tttctctatg aatcatcatc aagaactctg ggagaacttt tgaattcata 3550

accaagccaa catctccaga catgaagggc 3580

<210> 72
<211> 1191
<212> DNA
<213> Homo sapiens

<400> 72
tccaaccaga ttttcaccgc tatgcctcct ttcactctgg gagtcttcca 50

gaggtcttgc actcgggaga gcatgctcag gtttccccag ctctacaaaa 100

tcacccagaa tgccaaagac ttcaacgtga ccagtggtca tgccactgcc 150

tgttgatttg ttgaaaatat tgtttacacg tatgttcttg ttactgattg 200

tcagaaagct ggttttgaga ctgcagcttg gactaaattc agtcatctgg 250

ctgtctgggg aagcatgctg accagtctgg tgttctttgg catctactca 300

gccatctggt ccaccattct cattgcccca aatatgagag gacagaagaa 350

tggtaccggt actgccaatg gagatggagg aaggagacag aaagaaacag 400

agcccagacc ctagggacca ccagcatttg cagaatggat aaacagcctt 450

cttcctaaca aaggaagcac agcaactgtg atcctgagct gtgcacactt 500

ctggttggga ttatttctgg tttctacttc ctgtttgaag atgtggcatg 550

gagagtgaac aagctgctgc ccaccacctg gcatcacagc cccagaactc 600

agctatttcc atgggaccac agcatctcat ctctgggctg agccagaaag 650

acccctactg aagtccagag gcacttttct gaaaggctct gctttgacct 700

gaagtatttt atctatcctc agtctcagga cactgttgat ggaattaagg 750

ccaagcacat ctgcaaaaaa gacattgctg gaggaggtgc aaagagctgg 800

aaaccaagtc tccagtcctg ggaaaagcag tggtatggaa aagcaatgga 850

1438
1430

aagagcattt tgaaaatgcc attccactgt tttctggcct ttatgatttc 900

tgctgagaaa tccactgtta gtctgatggg gtctccttca tagcaccaat 950

gacctgaaga gccttgttga aggaagactc catctgatga ctcagagcaa 1000

gtatttttta gtgtgttatt gttattagca gaaagagggc cataaaatac 1050

atggggcaag ctgaatatat cttaggcaaa agaagaaaat attcaaattc 1100

ttatgttatt ttatctaatt attttatctc tttttgtgtg tgacttataa 1150

tgtgtgtatt gtattaataa aagtatataa acatgtagtt t 1191

<210> 73
<211> 2752
<212> DNA
<213> Homo sapiens

<400> 73
cattcccacc cttccttctt taataagcag gagcgaaaaa gacaaattcc 50

aaagaggatt gttcagttca agggaatgaa gaattcagaa taattttggt 100

aaatggattc caatatcggg aataagaata agctgaacag ttgacctgct 150

ttgaagaaac atactgtcca tttgtctaaa ataatctata acaaccaaac 200

caatcaaaat gaattcaaca ttattttccc aggttgaaaa tcattcagtc 250

cactctaatt tctcagagaa gaatgcccag cttctggctt ttgaaaatga 300

tgattgtcat ctgcccttgg ccatgatatt taccttagct cttgcttatg 350

gagctgtgat cattcttggt gtctctggaa acctggcctt gatcataatc 400

atcttgaaac aaaaggagat gagaaatgtt accaacatcc tgattgtgaa 450

cctttccttc tcagacttgc ttgttgccat catgtgtctc ccctttacat 500

ttgtctacac attaatggac cactgggtct ttggtgaggc gatgtgtaag 550

ttgaatcctt ttgtgcaatg tgtttcaatc actgtgtcca ttttctctct 600

ggttctcatt gctgtggaac gacatcagct gataatcaac cctcgagggt 650

ggagaccaaa taatagacat gcttatgtag gtattgctgt gatttgggtc 700

cttgctgtgg cttcttcttt gcctttcctg atctaccaag taatgactga 750

tgagccgttc caaaatgtaa cacttgatgc gtacaaagac aaatacgtgt 800

gctttgatca atttccatcg gactctcata ggttgtctta taccactctc 850

ctcttggtgc tgcagtattt tggtccactt tgttttatat ttatttgcta 900

cttcaagata tatatacgcc taaaaaggag aaacaacatg atggacaaga 950

tgagagacaa taagtacagg tccagtgaaa ccaaaagaat caatatcatg 1000

ctgctctcca ttgtggtagc atttgcagtc tgctggctcc ctcttaccat 1050

ctttaacact gtgtttgatt ggaatcatca gatcattgct acctgcaacc 1100

acaatctgtt attcctgctc tgccacctca cagcaatgat atccacttgt 1150

gtcaacccca tattttatgg gttcctgaac aaaaacttcc agagagactt 1200

gcagttcttc ttcaactttt gtgatttccg gtctcgggat gatgattatg 1250

aaacaatagc catgtccacg atgcacacag atgtttccaa aacttctttg 1300

aagcaagcaa gcccagtcgc atttaaaaaa atcaacaaca atgatgataa 1350

tgaaaaaatc tgaaactact tatagcctat ggtcccggat gacatctgtt 1400

1439
1457

taaaaacaag cacaacctgc aacatacttt gattacctgt tctcccaagg 1450

aatggggttg aaatcatttg aaaatgacta agattttctt gtcttgcttt 1500

ttactgcttt tgttgtagtt gtcataatta catttggaac aaaaggtgtg 1550

ggctttgggg tcttctggaa atagttttga ccagacatct ttgaagtgct 1600

ttttgtgaat ttatgcatat aatataaaga cttttatact gtacttattg 1650

gaatgaaatt tctttaaagt attacgatgc gctgacttca gaagtacctg 1700

ccatccaata cggtcattag attgggtcat cttgattaga ttagattaga 1750

ttagattgtc aacagattgg gccatcctta ctttatgata ggcatcattt 1800

tagtgtgtta caatagtaac agtatgcaaa agcagcattc aggagccgaa 1850

agatagtctt gaagtcattc agaagtggtt tgaggtttct gttttttggt 1900

ggtttttgtt tgttttttt ttttttcacc ttaagggagg ctttcatttc 1950

ctcccgactg attgtcactt aaatcaaaat ttaaaaatga ataaaaagac 2000

atacttctca gctgcaaata ttatggagaa ttgggcaccc acaggaatga 2050

agagagaaag cagctcccca acttcaaaac cattttggta cctgacaaca 2100

agagcatttt agagtaatta atttaataaa gtaaattagt attgctgcaa 2150

atagctaaat tatatttatt tgaattgatg gtcaagagat tttccatttt 2200

ttttacagac tgttcagtgt ttgtcaagct tctggtctaa tatgtactcg 2250

aaagactttc cgcttacaat ttgtagaaac acaaatatcg ttttccatac 2300

agcagtgcct atatagtgac tgattttaac tttcaatgtc catctttcaa 2350

aggaagtaac accaaggtac aatgttaaag gaatattcac tttacctagc 2400

agggaaaaat acacaaaaac tgcagatact tcatatagcc cattttaact 2450

tgtataaact gtgtgacttg tggcgtctta taaataatgc actgtaaaga 2500

ttactgaata gttgtgtcat gttaatgtgc ctaatttcat gtatcttgta 2550

atcatgattg agcctcagaa tcatttggag aaactatatt ttaaagaaca 2600

agacatactt caatgtatta tacagataaa gtattacatg tgtttgattt 2650

taaaagggcg gacattttat taaaatcaat attgtttttg ctttttctga 2700

ggagtctctt tcagtttcat tttttctcat cccatgactt ccctccgatg 2750

gt 2752

<210> 74
<211> 1001
<212> DNA
<213> Homo sapiens

<400> 74
agtatgaggc gcccaccatg gtgcgatgcg accgcgggct gcagatgctg 50

ctgaccacgg ccggagcctt cgccgccttc tcgctcatgg ccatcgccat 100

cggcaccgac tactggctgt actccagcgc gcacatctgc aacggcacca 150

acctgaccat ggacgacggg ccccgcccc gccgcgcccg cggcgacctc 200

acccactctg gtctgtggcg ggtgtgctgc atcgaaggga tctataaagg 250

gcactgcttc cggatcaatc acttcccaga ggacaatgac tacgaccacg 300

acagctcgga gtacctcctc cgcatcgtgc gagcctccag cgtcttcccc 350

1440
2432

atcctcagca ccatcctgct cctgctgggt ggcctgtgca tcggtgctgg 400

caggatctac agccgcaaga acaacatcgt cctcagtgcc ggcatcctct 450

tcgtggctgc aggcctcagt aacatcatcg gtatcatcgt ctacatttcc 500

agcaacacag gtgacccgag tgacaagcgg gacgaagaca aaaagaacca 550

ttacaactac ggctggtctt tttactttgg agctctgtct ttcattgtgg 600

ctgagaccgt gggcgtcctg gctgtaaaca tttacattga gaaaaataaa 650

gagttgaggt ttaagaccaa acgggaattc cttaaggcgt cttcctcttc 700

tccttatgcc aggatgccga gctacaggta ccggcgacgg cgctcgaggt 750

ccagctcaag gtccaccgag gcctcgccct ccagggacgt gtcgcccatg 800

ggcctgaaga tcacaggggc catccccatg ggggagctgt ccatgtacac 850

gctgtccagg gagcccctca aggtgaccac cgcagccagc tacagccccg 900

accaggaggc cagcttcctg caggtgcatg acttttttcca gcaggacctg 950

aaggaaggtt tccacgtcag catgctgaac cgacggacga cccctgtgtg 1000

a 1001

<210> 75
<211> 2805
<212> DNA
<213> Homo sapiens

<400> 75
cgggtctgat agtccctacc tgtcaggact ggtgttagga tgagataatg 50

tttgtgaact gtaaacatat ataaacgtgt gctactgtga gaactggaac 100

aaagaagaga gggagtgaga gaaatcaagg gagggctggg gctgggaaag 150

aacgaaaagg gagtcgcgta tagaggagag gcgacagtcg cgagccacac 200

tttgcaatga aactctttag actttctgcc gggagagcgg cccagacgcg 250

ccaggtctgt agcaggaggc cgcgagggcg ggtccccaga agcctacagg 300

tgagtatcgg ttctcccctt cccggctttc ggtccggagg aggcgggagc 350

agcttccctg ttctgatcct atcgcgggcg gcgcagggcc ggcttggcct 400

tccgtgggac ggggaggggg gcgggatgtg tcacccaaat accagtgggg 450

acggtcggtg gtggaaccag ccgggcaggt cgggtagagt ataagagccg 500

gagggagcgg ccggggcgca gacgcctgca gaccatccca gacgccggag 550

cccgagcccc gacgagtccc cgcgcctcat ccgcccgcgt ccggtccgcg 600

ttcctccgcc ccaccatggc tcggggcccc ggcctcgcgc cgccaccgct 650

gcggctgccg ctgctgctgc tggtgctggc ggcggtgacc ggccacacgg 700

ccgcgcagga caactgcacg tgtcccacca acaagatgac cgtgtgcagc 750

cccgacggcc ccggcggccg ctgccagtgc cgcgcgctgg gctcgggcat 800

ggcggtcgac tgctccacgc tgacctccaa gtgtctgctg ctcaaggcgc 850

gcatgagcgc ccccaagaac gcccgcacgc tggtgcggcc gagtgagcac 900

gcgctcgtgg acaacgatgg cctctacgac cccgactgcg accccgaggg 950

ccgcttcaag gcgcgccagt gcaaccagac gtcggtgtgc tggtgcgtga 1000

actcggtggg cgtgcgccgc acggacaagg gcgacctgag cctacgctgc 1050

gatgagctgg tgcgcaccca ccacatcctc attgacctgc gccaccgccc 1100

caccgccggc gccttcaacc actcagacct ggacgccgag ctgaggcggc 1150

tcttccgcga gcgctatcgg ctgcacccca agttcgtggc ggccgtgcac 1200

tacgagcagc ccaccatcca gatcgagctg cggcagaaca cgtctcagaa 1250

ggccgccggt gaagtggata tcggcgatgc cgcctactac ttcgagaggg 1300

acatcaaggg cgagtctcta ttccagggcc gcggcggcct ggacttgcgc 1350

gtgcgcggag aacccctgca ggtggagcgc acgctcatct attacctgga 1400

cgagattccc ccgaagttct ccatgaagcg cctcaccgcc ggcctcatcg 1450

ccgtcatcgt ggtggtcgtg gtggccctcg tcgccggcat ggccgtcctg 1500

gtgatcacca accggagaaa gtcggggaag tacaagaagg tggagatcaa 1550

ggaactgggg gagttgagaa aggaaccgag cttgtaggta cccggcgggg 1600

caggggatgg ggtggggtac cggatttcgg tatcgtccca gacccaagtg 1650

agtcacgctt cctgattcct cggcgcaaag gagacgttta tcctttcaaa 1700

ttcctgcctt ccccctccct tttgcgcaca caccaggttt aatagatcct 1750

ggcctcaggg tctcctttct ttctcacttc tgtcttgagg gaagcatttc 1800

taaaatgtat ccccttccgg tccaacaaca ggaaaccctga ctggggcagt 1850

gaaggaaggg atggcacagc gttatgtgta aaaaacaagt atctgtatga 1900

caacccggga tcgtttgcaa gtaactgaat ccattgcgac attgtgaagg 1950

cttaaatgag tttagatggg aaatagcgtt gttatcgcct tgggtttaaa 2000

ttatttgatg agttccactt gtatcatggc ctacccgagg agaagaggag 2050

tttgttaact gggcctatgt agtagcctca tttaccatcg tttgtattac 2100

tgaccacata tgcttgtcac tgggaaagaa gcctgtttca gctgcctgaa 2150

cgcagtttgg atgtctttga ggacagacat tgcccggaaa ctcagtctat 2200

ttattcttca gcttgccctt actgccactg atattggtaa tgttcttttt 2250

tgtaaaatgt ttgtacatat gttgtctttg ataatgttgc tgtaattttt 2300

taaaataaaa cacgaattta ataaaatatg ggaaaggcac aaaccagaag 2350

tcggcatttg tgaaaagtcc ctccagattt ctatcacttt ggtctctaat 2400

ttcccaagac ttgtattttt tttttatttc aaattataac actttttttt 2450

cccccagaag tgggtgtttc atgttgctac tctggtgtgt cccaagatat 2500

cctaactggc cagtgtaaat gctattcttt ctaaataaga ttatttggaa 2550

acttccttca aactgcagga gggcgagctc tgagggcacg agaagctaaa 2600

actagctgct tttgatgaaa aagagtgcca gtctttggtc atctctaaac 2650

aaggcttatc accaatggag acagaaaact ctagttcaag agctgtacct 2700

cctttgaatc ccagccctac tcgaaataag tggtactatt tccatttagc 2750

ctttgagcaa atcacttaac tcaaaggcgt tgtggctcta agattaaacg 2800

acttt 2805

<210> 76
<211> 2637
<212> DNA
<213> Homo sapiens

1442
1435

<400> 76

gcggccgcgg agcaagaagg gcgccgcgtc gtgcggcccg cgcagccccc 50

ggagccatgg gcaagtgcag cgggcgctgc acgctggtcg ccttctgctg 100

cctgcagctg gtggctgcgc tggagcggca gatctttgac ttcctgggct 150

accagtgggc tcccatccta gccaacttcc tgcacatcat ggcagtcatc 200

ctgggcatct ttggcaccgt gcagtaccgc tcccggtacc tcatcctgta 250

tgcagcctgg ctggtgctct gggttggctg gaatgcattt atcatctgct 300

tctacttgga ggttggacag ctgtcccagg accgggactt catcatgacc 350

ttcaacacat ccctgcaccg ctcctggtgg atggagaatg ggccaggctg 400

cctggtgaca cctgttctga actcccgcct ggctctggag gaccaccatg 450

tcatctctgt cactggctgc ctgcttgact acccctacat tgaagccctc 500

agcagcgccc tgcagatctt cctggcactg ttcggcttcg tgttcgcctg 550

ctacgtgagc aaagtgttcc tggaggagga ggacagcttt gacttcatcg 600

gcggctttga ctcctacgga taccaggcgc cccagaagac gtcgcattta 650

cagctgcagc ctctgtacac gtcggggtag cctctgcccc gcgcccaccc 700

cggcgcctcg ccctgggctg accgcagctg ccgcgagctc gggccaaggc 750

gcaggcgtgt ccccctggtg gcccgcgctc tcactgcagc ctgtgcccaa 800

ccccgcgtct gcatctggag atgcggactt ggacgtggac ttggacttgg 850

acttggattt gagcttggct cttcgcagcc cggacttcgg aggagtgggg 900

cggggcgggg gaggggcacc acgggttttt tgtttttttgt ttgtttgttt 950

ttaatctcag ccttggcgtg agctggggcc ttcctctctt ctccagcctc 1000

tccctttcac tcttcgccca gcatcctgcc cccctgtcca aaaacagcag 1050

gacatcaggc ccatcccatc ccaccacact cactcaccag ctctggggaa 1100

agctactgtg aactaggagc aggattcctg ggttctaatc gcaggtccat 1150

cactgactgt gacgtctagc aaagcccttg ccctctctga gcctcggttt 1200

ccgcacctca agtaattaat cccttagcaa atggactctt ttagacttct 1250

catttaactc aattccctga gctagactgg gattaaaatt ctcattttgc 1300

agtacattaa aactgatgcc cagagatgtg atttgcttga ggccacacag 1350

ctagattttt ggtggaagtg ggccttgaac acagtgtact ttctgcagtt 1400

tctgactgta aaacccagtg tctgctctct gagttccatt tccaagcccc 1450

cctccatctt ggacctatgt ggtctccacc atattcacac accaccacca 1500

ccacttgcca atgcctctct taaagcaata tacccattcg ttctcttatt 1550

gggaactgga tggatgaagc cccaaattca gccccaccca cagagaagcc 1600

ttcctacact cagcctctgt ccacccttgg caaatctttc aagctctctc 1650

ctccaggaaa gtggggcccc aactcagtca ctccaccccc ttccaggtcc 1700

ctgaggctgg ttctactgta tccccatcac ctccacaact ccactcaccc 1750

ctgacggctc catccacctc accagttgga aggcttgtgg tttcagagag 1800

gagcaatgct ggtcagcgct gcccagactc cagtgtttac agatcaccag 1850

catttacaac caatccaatg gccagaagcc tcctctaaca agcccagaag 1900

gagttctgaa ggggcagatg ggggtgtgag tagtcgggga gtcgggattg 1950

ccagcaccct caccettcct tgggggcaag tagaggtgag aacactttcc 2000

ccacctccct ccacagacac tcctgaggac gctgcatccc acgcactgcc 2050

tggtgcgtcc atagagagag gatcaggtct cagcatttca tctgtgaaag 2100

aggcatggcc ctgggttaga aaggagggca ggagacatgg aggaactggg 2150

gggcacccag atggtgcaga tggtttgcac acctgagcct gtctgtggtg 2200

accattccgc tcctctccca ctaccctcca atctatcatt ccctactctc 2250

taaggccaaa atatcctgag caaggctggc aaccccaccc caccatccca 2300

aatgcaagca gccaggccca ggagttcctc tggcccccac aggcatggag 2350

ctcccagctg gtgggtacag cttgagaggg gggcagctcc ctcaggctaa 2400

gctactgccc ttcactgggc cagccctgcc tccagccctc acctctctca 2450

ccccaactct ccccccaagcc cctttctact caacgggtgt agccactggt 2500

gctttgaagc cttttgtttt tataagatgg tttttgcaag gggaccaggt 2550

tctcttttca ctgggacctt gcaaggaggg gagtgctctc ctggtttctg 2600

tgcaggcggg ttgattaaag atggtgtttt cttctct 2637

<210> 77
<211> 1829
<212> DNA
<213> Homo sapiens

<400> 77
ggcggggaga ggagaggaga gaagagccgc ggggggccca gcccggagcc 50

aggatgcccg cgccgcgcgc ccgggagcag ccccgcgtgc ccggggagcg 100

ccagccgctg ctgcctcgcg gtgcgcgggg ccctcgacgg tggcggcggg 150

cggcgggcgc ggccgtgctg ctggtggaga tgctggagcg cgccgccttc 200

ttcggcgtca ccgccaacct cgtgctgtac ctcaacagca ccaacttcaa 250

ctggaccggc gagcaggcga cgcgcgccgc gctggtattc ctgggcgcct 300

cctacctgct ggcgcccgtg ggcggctggc tggccgacgt gtacctgggc 350

cgctaccgcg cggtcgcgct cagcctgctg ctctacctgg ccgcctcggg 400

cctgctgccc gccaccgcct tccccgacgg ccgcagctcc ttctgcggag 450

agatgcccgc gtcgccgctg ggacctgcct gcccctcggc cggctgcccg 500

cgctcctcgc ccagccccta ctgcgcgccc gtcctctacg cgggcctgct 550

gctactcggc ctggccgcca gctccgtccg gagcaacctc acctccttcg 600

gtgccgacca ggtgatggat ctcggccgcg acgccacccg ccgcttcttc 650

aactggtttt actggagcat caacctgggt gctgtgctgt cgctgctggt 700

ggtggcgttt attcagcaga acatcagctt cctgctgggc tacagcatcc 750

ctgtgggctg tgtgggcctg gcatttttca tcttcctctt tgccaccccc 800

gtcttcatca ccaagccccc gatgggcagc caagtgtcct ctatgcttaa 850

gctcgctctc caaaactgct gcccccagct gtggcaacga cactcggcca 900

gagaccgtca atgtgcccgc gtgctggccg acgagaggtc tccccagcca 950

ggggcttccc cgcaagagga catcgccaac ttccaggtgc tggtgaagat 1000

cttgcccgtc atggtgaccc tggtgcccta ctggatggtc tacttccaga 1050

tgcagtccac ctatgtcctg cagggtcttc acctccacat cccaaacatt 1100

ttcccagcca acccggccaa catctctgtg gccctgagag cccagggcag 1150

cagctacacg atcccggaag cctggctcct cctggccaat gttgtggtgg 1200

tgctgattct ggtccctctg aaggaccgct tgatcgaccc tttactgctg 1250

cggtgcaagc tgcttccctc tgctctgcag aagatggcgc tggggatgtt 1300

ctttggtttt acctccgtca ttgtggcagg agtcctggag atggagcgct 1350

tacactacat ccaccacaac gagaccgtgt cccagcagat tggggaggtc 1400

ctgtacaacg cggcaccact gtccatctgg tggcagatcc ctcagtacct 1450

gctcattggg atcagtgaga tctttgccag catcccaggc ctggagtttg 1500

cctactcaga ggccccgcgc tccatgcagg gcgccatcat gggcatcttc 1550

ttctgcctgt cgggggtggg ctcactgttg ggctccagcc tagtggcact 1600

gctgtccttg cccgggggct ggctgcactg ccccaaggac tttgggaaca 1650

tcaacaattg ccggatggac ctctacttct tcctgctggc tggcattcag 1700

gccgtcacgg ctctcctatt tgtctggatc gctggacgct atgagagggc 1750

gtcccagggc ccagcctccc acagccgttt cagcagggac aggggctgaa 1800

caggccctat tccagccccc ttgcttcac 1829

<210> 78
<211> 5746
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 5621,, 5664
<223> unknown base

<400> 78
ggccgcgggc tctcgcgggg cggcgacgcc gcggggagga tgctgcttgc 50

cgcgcccgcg tcctcaccgt cctcccgggc cgcctgctgg ggctttgttg 100

tggcccggac gccgcgggcc acccccctgaa gtcgcctgcc gccgccgccg 150

ccgcacctag cggacgggcg ggcgggcgcg cgtgtgccca ggagtgcgcg 200

cctgtcgcgg tggtgggtgc aggactggac ccacgggccc attgtgcgcc 250

cgcccgcggc agccaggacc atgtgggtga acccggagga ggtgttgctg 300

gccaacgcgc tgtggatcac cgagagggcc aacccatact tcatcctgca 350

gcggaggaag ggccacgccg gcgatggagg cggcggcggc ggactggcgg 400

gcctgctggt gggtaccctt gatgttgtgt tggactccag cgcccgggtc 450

gctccttacc gaatcttgta ccagactcca gactccctgg tctactggac 500

catcgcctgt ggtggttcca ggaaagaaat cactgaacac tgggaatggc 550

ttgagcaaaa tctcttgcag acactctcca tctttgaaaa tgagaatgat 600

atcaccacat ttgtgagagg aaaaatacag ggcatcattg cagaatacaa 650

caaaatcaat gatgtaaagg aagatgatga cacggagaag tttaaagaag 700

ccattgtgaa atttcatagg ctgtttggga tgccagagga agagaaactc 750

1445
1438

gtcaactatt actcttgcag ctattggaag gggaaggtcc cccgtcaggg 800

ttggatgtac ctcagcatta accacctttg ctttattct tttcttatgg 850

gaagggaagc gaaactggtc atccggtggg tagacatcac tcagcttgag 900

aagaatgcca ccctgcttct gcctgatgtg atcaaagtga gcacacggtc 950

cagtgagcat ttcttctctg tattcctcaa catcaacgag accttcaagt 1000

taatggagca gcttgccaac atagccatga ggcaactctt agacaatgag 1050

ggatttgaac aagatcgatc cctgcccaaa ctcaaaagga aatctcctaa 1100

aaaagtgtct gctctaaaac gtgatcttga tgccagggca aagagtgaga 1150

gataccgtgc actttccgg ctgcccaaag atgaaaaatt agatggccac 1200

acagactgca ctctctggac tccatttaac aaaatgcaca ttttggggca 1250

gatgtttgtg tccacaaatt acatctgttt taccagcaag gaggagaact 1300

tatgtagcct cattatcccg ctccgtgagg tgacaattgt ggaaaaggca 1350

gacagctcca gtgtgctccc cagtccctta tccatcagca cccgaaacag 1400

gatgaccttc ctatttgcca acttgaaaga tagagacttt ctagtgcaga 1450

ggatctcaga tttcctgcaa cagactactt ccaaaatata ttctgacaag 1500

gagtttgcag gaagttacaa cagttcagat gatgaggtgt actctcgacc 1550

cagcagcctc gtctcctcca gcccccagag aagcacgagc tctgatgctg 1600

atggagagcg ccagtttaac ctaaatggca acagcgtccc cacagccaca 1650

cagaccctga tgaccatgta tcggcggcgg tctcccgagg agttcaaccc 1700

gaaattggcc aaagagtttc tgaaagagca agcctggaag attcactttg 1750

ctgagtatgg gcaagggatc tgcatgtacc gcacagagaa aacgcgggag 1800

ctggtgttga agggcatccc ggagagcatg cgtggggagc tctggctgct 1850

gctgtcaggt gccatcaatg agaaggccac acatcctggg tactatgaag 1900

acctagtgga gaagtccatg gggaagtata atctcgccac ggaggagatt 1950

gagagggatt tacaccgctc ccttccagaa cacccagctt ttcagaatga 2000

aatgggcatt gctgcactaa ggagagtctt aacagcttat gcttttcgaa 2050

atcccaacat agggtattgc caggccatga atattgtcac ttcagtgctg 2100

ctgctttatg ccaaagagga ggaagctttc tggctgcttg tggctttgtg 2150

tgagcgcatg ctcccagatt actacaacac cagagttgtg ggtgcactgg 2200

tggaccaagg tgtctttgag gagctagcac gagactacgt cccacagctg 2250

tacgactgca tgcaagacct gggcgtgatt tccaccatct ccctgtcttg 2300

gttcctcaca ctatttctca gtgtgatgcc ttttgagagt gcagttgtgg 2350

ttgttgactg tttcttctat gaaggaatta aagtgatatt ccagttggcc 2400

ctagctgtgc tggatgcaaa tgtggacaaa ctgttgaact gcaaggatga 2450

tggggaggcc atgaccgttt tgggaaggta tttagacagt gtgaccaata 2500

aagacagcac actgcctccc attcctcacc tccactcctt gctcagcgat 2550

gatgtggaac cttaccctga ggtagacatc tttagactca tcagaacttc 2600

ctacgagaaa ttcggaacta tccgggcaga tttgattgaa cagatgagat 2650

1446
1439

tcaaacagag actgaaagtg atccagacgc tggaggatac tacgaaacgc 2700

aacgtggtac gaaccattgt gacagaaact tcctttacca ttgatgagct 2750

ggaagaactt tatgctcttt tcaaggcaga acatctcacc agctgctact 2800

ggggcgggag cagcaacgcg ctggaccggc atgaccccag cctgccctac 2850

ctggaacagt atcgcattga cttcgagcag ttcaagggaa tgtttgctct 2900

tctctttcct tgggcatgtg gaactcactc tgacgttctg gcctcccgct 2950

tgttccagtt attagatgaa aatggagact ctttgattaa cttccgggag 3000

tttgtctctg ggctaagtgc tgcatgccat ggggacctca cagagaagct 3050

caaactcctg tacaaaatgc acgtcttgcc tgagccatcc tctgatcaag 3100

atgaaccaga ttctgctttt gaagcaactc agtacttctt tgaagatatt 3150

accccagaat gtacacatgt tgttggattg gatagcagaa gcaaacaggg 3200

tgcagatgat ggctttgtta cggtgagcct aaagccagac aaagggaaga 3250

gagcaaattc ccaagaaaat cgtaattatt tgagactgtg gactccagaa 3300

aataaatcta agtcaaagaa tgcaaaggat ttacccaaat taaatcaggg 3350

gcagttcatt gaactgtgta agacaatgta taacatgttc agcgaagacc 3400

ccaatgagca ggagctgtac catgccacgg cagcagtgac cagcctcctg 3450

ctggagattg gggaggtcgg caagttgttc gtggcccagc ctgcaaagga 3500

gggcgggagc ggaggcagtg ggccgtcctg ccaccagggc atcccaggcg 3550

tgctcttccc caagaaaggg ccaggccagc cttacgtggt ggagtctgtt 3600

gagcccctgc cggccagcct ggcccccgac agcgaggaac actcccttgg 3650

aggacaaatg gaggacatca agctggagga ctcctcgccc cgggacaacg 3700

gggcctgctc ctccatgctg atctctgacg acgacaccaa ggacgacagc 3750

tccatgtcct catactcggt gctgagtgcc ggctcccacg aggaggacaa 3800

gctgcactgc gaggacatcg gagaggacac ggtcctggtg cggagcggcc 3850

agggcacggc ggcactgccc cggagcacca gcctggaccg ggactgggcc 3900

atcaccttcg agcagttcct ggcctccctc ttaactgagc ctgccctggt 3950

caagtacttt gacaagcccg tgtgcatgat ggccaggatt accagtgcaa 4000

aaaacatccg gatgatgggc aagcccctca cctcggccag tgactatgaa 4050

atctcggcca tgtccggctg acacgggcgc cttcccgggg gagtgggagg 4100

agagggaggg gagggatttt ttatgttctt ctgtgttgag ttttttcttt 4150

ctttctttta aattaaatat ttattagtac ctggcttgaa gcctagtgtt 4200

ttcataatgt aattcaatga aaactgttgg agaaatattt aaacacctca 4250

atgtaggtac attacactct tgttgcgggg aggggattta ccagaataca 4300

gtttatttcg tgaattctaa aaaacaaaaa gatgaatctg tcagtgatat 4350

gtgtgtatta taacttatta atcttgctgt tgagctgtat acatggttta 4400

aaaaatagta ctgtttaatg ctaagtaagg cagcagtcat ttgtgtattc 4450

aggcttttta aataaaatta gagctgtaag gaaaatgaaa agccacaaat 4500

gcaagactgt tcttaaatgg aaggcatagt cagcgagggt aaatcctata 4550

ccactttagg aagtattaaa aatattttta agatttgaaa tatatttcat 4600

agaagtcctc tattcaaaat catattccac agatgttccc cttcaaaggg 4650

aaaacatttg gggttctaaa cagttatgaa agtaagtgat ttttacatga 4700

ttccagaata acacttgtat tgaccaattt agacagatac cagaccaatt 4750

ttgcatttaa gaaattgttc tgattattta cgtcaactca ttagaattca 4800

gtgaaaagta acagtctttt gtcacagaga atctgaaagt agcagcaaag 4850

acagagggct catgacaggt ttttgctttt gctttgcttt tgtttttgaa 4900

agagtaaaag tactgatgct tctgatactg gatgtttagc ttcttactgc 4950

aaaaacataa gtaaaacagt caactttacc atttccgtat tctccataga 5000

ttgaagaaat ttataccaca tatcgcatat gaccatcttt ccatcaaatc 5050

aatgtagaga taatgtaaac tgaaaaaaaa tctgcaagat aatgtaactg 5100

aatgttttaa aaacagaact tgtcacttta tataaaagaa tagtatgctc 5150

tatttcctga atggatgtgg aaatgaaagc tagcgcacct gcactttgaa 5200

ttcttgcttc ttttttatta ctgttatgat tttgcttttt acagatgttg 5250

gacgattttt tcttctgatt gttgaattca taatcatggt ctcatttcct 5300

ttgcttcttt ggaatatttc tttcaacaca ttcctttatt ttattataca 5350

ttgtgtcctt tttttagcta ttgctgctgt tgtttttttat tctatttaca 5400

ggatgatttt taaactgtca aatgaagtag tgttaacctc aaataggcta 5450

aatgtgaaca aataaaatac agcaaatact cagatacagc tttttatctt 5500

tgtgcttgag ttcctgccta aggaataaca ttattctttt gacaactttt 5550

gcagggaaaa ttatatcagg caaccatttt gattaagtaa ataaatttta 5600

taggcaaaca tatagagaga natacaattt gtagtatatc aatgactata 5650

tttaaaataa ggantataat tgttatcagt tatctaactt aaaatgctta 5700

tccataatga tcagtgatat tcagcttttt aaaatatgct tgttgg 5746

<210> 79
<211> 183
<212> PRT
<213> Homo sapiens

<400> 79

Met	Arg	Ala	Ala	Tyr	Leu	Phe	Leu	Leu	Phe	Leu	Pro	Ala	Gly	Leu
1				5					10					15
Leu	Ala	Gln	Gly	Gln	Tyr	Asp	Leu	Asp	Pro	Leu	Pro	Pro	Phe	Pro
				20					25					30
Asp	His	Val	Gln	Tyr	Thr	His	Tyr	Ser	Asp	Gln	Ile	Asp	Asn	Pro
				35					40					45
Asp	Tyr	Tyr	Asp	Tyr	Gln	Glu	Val	Thr	Pro	Arg	Pro	Ser	Glu	Glu
				50					55					60
Gln	Phe	Gln	Phe	Gln	Ser	Gln	Gln	Gln	Val	Gln	Gln	Glu	Val	Ile
				65					70					75
Pro	Ala	Pro	Thr	Pro	Glu	Pro	Gly	Asn	Ala	Glu	Leu	Glu	Pro	Thr
				80					85					90
Glu	Pro	Gly	Pro	Leu	Asp	Cys	Arg	Glu	Glu	Gln	Tyr	Pro	Cys	Thr
				95					100					105

1448
1471

Arg Leu Tyr Ser Ile His Arg Pro Cys Lys Gln Cys Leu Asn Glu
110 115 120

Val Cys Phe Tyr Ser Leu Arg Arg Val Tyr Val Ile Asn Lys Glu
125 130 135

Ile Cys Val Arg Thr Val Cys Ala His Glu Glu Leu Leu Arg Ala
140 145 150

Asp Leu Cys Arg Asp Lys Phe Ser Lys Cys Gly Val Met Ala Ser
155 160 165

Ser Gly Leu Cys Gln Ser Val Ala Ala Ser Cys Ala Arg Ser Cys
170 175 180

Gly Ser Cys

<210> 80
<211> 909
<212> PRT
<213> Homo sapiens

<400> 80
Met Ala Ala Arg Pro Leu Pro Val Ser Pro Ala Arg Ala Leu Leu
1 5 10 15

Leu Ala Leu Ala Gly Ala Leu Leu Ala Pro Cys Glu Ala Arg Gly
20 25 30

Val Ser Leu Trp Asn Glu Gly Arg Ala Asp Glu Val Val Ser Ala
35 40 45

Ser Val Arg Ser Gly Asp Leu Trp Ile Pro Val Lys Ser Phe Asp
50 55 60

Ser Lys Asn His Pro Glu Val Leu Asn Ile Arg Leu Gln Arg Glu
65 70 75

Ser Lys Glu Leu Ile Ile Asn Leu Glu Arg Asn Glu Gly Leu Ile
80 85 90

Ala Ser Ser Phe Thr Glu Thr His Tyr Leu Gln Asp Gly Thr Asp
95 100 105

Val Ser Leu Ala Arg Asn Tyr Thr Val Ile Leu Gly His Cys Tyr
110 115 120

Tyr His Gly His Val Arg Gly Tyr Ser Asp Ser Ala Val Ser Leu
125 130 135

Ser Thr Cys Ser Gly Leu Arg Gly Leu Ile Val Phe Glu Asn Glu
140 145 150

Ser Tyr Val Leu Glu Pro Met Lys Ser Ala Thr Asn Arg Tyr Lys
155 160 165

Leu Phe Pro Ala Lys Lys Leu Lys Ser Val Arg Gly Ser Cys Gly
170 175 180

Ser His His Asn Thr Pro Asn Leu Ala Ala Lys Asn Val Phe Pro
185 190 195

Pro Pro Ser Gln Thr Trp Ala Arg Arg His Lys Arg Glu Thr Leu
200 205 210

Lys Ala Thr Lys Tyr Val Glu Leu Val Ile Val Ala Asp Asn Arg
215 220 225

Glu Phe Gln Arg Gln Gly Lys Asp Leu Glu Lys Val Lys Gln Arg
230 235 240

Leu Ile Glu Ile Ala Asn His Val Asp Lys Phe Tyr Arg Pro Leu
245 250 255

Asn Ile Arg Ile Val Leu Val Gly Val Glu Val Trp Asn Asp Met

260 265 270

Asp Lys Cys Ser Val Ser Gln Asp Pro Phe Thr Ser Leu His Glu
275 280 285

Phe Leu Asp Trp Arg Lys Met Lys Leu Leu Pro Arg Lys Ser His
290 295 300

Asp Asn Ala Gln Leu Val Ser Gly Val Tyr Phe Gln Gly Thr Thr
305 310 315

Ile Gly Met Ala Pro Ile Met Ser Met Cys Thr Ala Asp Gln Ser
320 325 330

Gly Gly Ile Val Met Asp His Ser Asp Asn Pro Leu Gly Ala Ala
335 340 345

Val Thr Leu Ala His Glu Leu Gly His Asn Phe Gly Met Asn His
350 355 360

Asp Thr Leu Asp Arg Gly Cys Ser Cys Gln Met Ala Val Glu Lys
365 370 375

Gly Gly Cys Ile Met Asn Ala Ser Thr Gly Tyr Pro Phe Pro Met
380 385 390

Val Phe Ser Ser Cys Ser Arg Lys Asp Leu Glu Thr Ser Leu Glu
395 400 405

Lys Gly Met Gly Val Cys Leu Phe Asn Leu Pro Glu Val Arg Glu
410 415 420

Ser Phe Gly Gly Gln Lys Cys Gly Asn Arg Phe Val Glu Glu Gly
425 430 435

Glu Glu Cys Asp Cys Gly Glu Pro Glu Glu Cys Met Asn Arg Cys
440 445 450

Cys Asn Ala Thr Thr Cys Thr Leu Lys Pro Asp Ala Val Cys Ala
455 460 465

His Gly Leu Cys Cys Glu Asp Cys Gln Leu Lys Pro Ala Gly Thr
470 475 480

Ala Cys Arg Asp Ser Ser Asn Ser Cys Asp Leu Pro Glu Phe Cys
485 490 495

Thr Gly Ala Ser Pro His Cys Pro Ala Asn Val Tyr Leu His Asp
500 505 510

Gly His Ser Cys Gln Asp Val Asp Gly Tyr Cys Tyr Asn Gly Ile
515 520 525

Cys Gln Thr His Glu Gln Gln Cys Val Thr Leu Trp Gly Pro Gly
530 535 540

Ala Lys Pro Ala Pro Gly Ile Cys Phe Glu Arg Val Asn Ser Ala
545 550 555

Gly Asp Pro Tyr Gly Asn Cys Gly Lys Val Ser Lys Ser Ser Phe
560 565 570

Ala Lys Cys Glu Met Arg Asp Ala Lys Cys Gly Lys Ile Gln Cys
575 580 585

Gln Gly Gly Ala Ser Arg Pro Val Ile Gly Thr Asn Ala Val Ser
590 595 600

Ile Glu Thr Asn Ile Pro Leu Gln Gln Gly Gly Arg Ile Leu Cys
605 610 615

Arg Gly Thr His Val Tyr Leu Gly Asp Asp Met Pro Asp Pro Gly
620 625 630

Leu Val Leu Ala Gly Thr Lys Cys Ala Asp Gly Lys Ile Cys Leu
635 640 645

1450
1472

Asn Arg Gln Cys Gln Asn Ile Ser Val Phe Gly Val His Glu Cys
650 655 660

Ala Met Gln Cys His Gly Arg Gly Val Cys Asn Asn Arg Lys Asn
665 670 675

Cys His Cys Glu Ala His Trp Ala Pro Pro Phe Cys Asp Lys Phe
680 685 690

Gly Phe Gly Gly Ser Thr Asp Ser Gly Pro Ile Arg Gln Ala Asp
695 700 705

Asn Gln Gly Leu Thr Ile Gly Ile Leu Val Thr Ile Leu Cys Leu
710 715 720

Leu Ala Ala Gly Phe Val Val Tyr Leu Lys Arg Lys Thr Leu Ile
725 730 735

Arg Leu Leu Phe Thr Asn Lys Lys Thr Thr Ile Glu Lys Leu Arg
740 745 750

Cys Val Arg Pro Ser Arg Pro Pro Arg Gly Phe Gln Pro Cys Gln
755 760 765

Ala His Leu Gly His Leu Gly Lys Gly Leu Met Arg Lys Pro Pro
770 775 780

Asp Ser Tyr Pro Pro Lys Asp Asn Pro Arg Arg Leu Leu Gln Cys
785 790 795

Gln Asn Val Asp Ile Ser Arg Pro Leu Asn Gly Leu Asn Val Pro
800 805 810

Gln Pro Gln Ser Thr Gln Arg Val Leu Pro Pro Leu His Arg Ala
815 820 825

Pro Arg Ala Pro Ser Val Pro Ala Arg Pro Leu Pro Ala Lys Pro
830 835 840

Ala Leu Arg Gln Ala Gln Gly Thr Cys Lys Pro Asn Pro Pro Gln
845 850 855

Lys Pro Leu Pro Ala Asp Pro Leu Ala Arg Thr Thr Arg Leu Thr
860 865 870

His Ala Leu Ala Arg Thr Pro Gly Gln Trp Glu Thr Gly Leu Arg
875 880 885

Leu Ala Pro Leu Arg Pro Ala Pro Gln Tyr Pro His Gln Val Pro
890 895 900

Arg Ser Thr His Thr Ala Tyr Ile Lys
905

<210> 81
<211> 158
<212> PRT
<213> Homo sapiens

<400> 81
Met Ala Ser Arg Ser Met Arg Leu Leu Leu Leu Leu Ser Cys Leu
1 5 10 15

Ala Lys Thr Gly Val Leu Gly Asp Ile Ile Met Arg Pro Ser Cys
20 25 30

Ala Pro Gly Trp Phe Tyr His Lys Ser Asn Cys Tyr Gly Tyr Phe
35 40 45

Arg Lys Leu Arg Asn Trp Ser Asp Ala Glu Leu Glu Cys Gln Ser
50 55 60

Tyr Gly Asn Gly Ala His Leu Ala Ser Ile Leu Ser Leu Lys Glu
65 70 75

Ala Ser Thr Ile Ala Glu Tyr Ile Ser Gly Tyr Gln Arg Ser Gln
80 85 90

Pro Ile Trp Ile Gly Leu His Asp Pro Gln Lys Arg Gln Gln Trp
95 100 105

Gln Trp Ile Asp Gly Ala Met Tyr Leu Tyr Arg Ser Trp Ser Gly
110 115 120

Lys Ser Met Gly Gly Asn Lys His Cys Ala Glu Ile Ser Ser Asn
125 130 135

Asn Asn Phe Leu Thr Trp Ser Ser Asn Glu Cys Asn Lys Arg Gln
140 145 150

His Phe Leu Cys Lys Tyr Arg Pro
155

<210> 82
<211> 84
<212> PRT
<213> Homo sapiens

<400> 82
Met Ala Thr Met Glu Asn Lys Val Ile Cys Ala Leu Val Leu Val
1 5 10 15

Ser Met Leu Ala Leu Gly Thr Leu Ala Glu Ala Gln Thr Glu Thr
20 25 30

Cys Thr Val Ala Pro Arg Glu Arg Gln Asn Cys Gly Phe Pro Gly
35 40 45

Val Thr Pro Ser Gln Cys Ala Asn Lys Gly Cys Cys Phe Asp Asp
50 55 60

Thr Val Arg Gly Val Pro Trp Cys Phe Tyr Pro Asn Thr Ile Asp
65 70 75

Val Pro Pro Glu Glu Glu Cys Glu Phe
80

<210> 83
<211> 680
<212> PRT
<213> Homo sapiens

<400> 83
Met Leu Pro Gln Ile Pro Phe Leu Leu Leu Val Ser Leu Asn Leu
1 5 10 15

Val His Gly Val Phe Tyr Ala Glu Arg Tyr Gln Thr Pro Thr Gly
20 25 30

Ile Lys Gly Pro Leu Pro Asn Thr Lys Thr Gln Phe Phe Ile Pro
35 40 45

Tyr Thr Ile Lys Ser Lys Gly Ile Ala Val Arg Gly Glu Gln Gly
50 55 60

Thr Pro Gly Pro Pro Gly Pro Ala Gly Pro Arg Gly His Pro Gly
65 70 75

Pro Ser Gly Pro Pro Gly Lys Pro Gly Tyr Gly Ser Pro Gly Leu
80 85 90

Gln Gly Glu Pro Gly Leu Pro Gly Pro Pro Gly Pro Ser Ala Val
95 100 105

Gly Lys Pro Gly Val Pro Gly Leu Pro Gly Lys Pro Gly Glu Arg
110 115 120

Gly Pro Tyr Gly Pro Lys Gly Asp Val Gly Pro Ala Gly Leu Pro
125 130 135

Gly Pro Arg Gly Pro Pro Gly Pro Pro Gly Ile Pro Gly Pro Ala

1452
13424

```
                140                 145                 150

Gly Ile Ser Val Pro Gly Lys Pro Gly Gln Gln Gly Pro Thr Gly
                155                 160                 165

Ala Pro Gly Pro Arg Gly Phe Pro Gly Glu Lys Gly Ala Pro Gly
                170                 175                 180

Val Pro Gly Met Asn Gly Gln Lys Gly Glu Met Gly Tyr Gly Ala
                185                 190                 195

Pro Gly Arg Pro Gly Glu Arg Gly Leu Pro Gly Pro Gln Gly Pro
                200                 205                 210

Thr Gly Pro Ser Gly Pro Pro Gly Val Gly Lys Arg Gly Glu Asn
                215                 220                 225

Gly Val Pro Gly Gln Pro Gly Ile Lys Gly Asp Arg Gly Phe Pro
                230                 235                 240

Gly Glu Met Gly Pro Ile Gly Pro Pro Gly Pro Gln Gly Pro Pro
                245                 250                 255

Gly Glu Arg Gly Pro Glu Gly Ile Gly Lys Pro Gly Ala Ala Gly
                260                 265                 270

Ala Pro Gly Gln Pro Gly Ile Pro Gly Thr Lys Gly Leu Pro Gly
                275                 280                 285

Ala Pro Gly Ile Ala Gly Pro Pro Gly Pro Pro Gly Phe Gly Lys
                290                 295                 300

Pro Gly Leu Pro Gly Leu Lys Gly Glu Arg Gly Pro Ala Gly Leu
                305                 310                 315

Pro Gly Gly Pro Gly Ala Lys Gly Glu Gln Gly Pro Ala Gly Leu
                320                 325                 330

Pro Gly Lys Pro Gly Leu Thr Gly Pro Pro Gly Asn Met Gly Pro
                335                 340                 345

Gln Gly Pro Lys Gly Ile Pro Gly Ser His Gly Leu Pro Gly Pro
                350                 355                 360

Lys Gly Glu Thr Gly Pro Ala Gly Pro Ala Gly Tyr Pro Gly Ala
                365                 370                 375

Lys Gly Glu Arg Gly Ser Pro Gly Ser Asp Gly Lys Pro Gly Tyr
                380                 385                 390

Pro Gly Lys Pro Gly Leu Asp Gly Pro Lys Gly Asn Pro Gly Leu
                395                 400                 405

Pro Gly Pro Lys Gly Asp Pro Gly Val Gly Gly Pro Pro Gly Leu
                410                 415                 420

Pro Gly Pro Val Gly Pro Ala Gly Ala Lys Gly Met Pro Gly His
                425                 430                 435

Asn Gly Glu Ala Gly Pro Arg Gly Ala Pro Gly Ile Pro Gly Thr
                440                 445                 450

Arg Gly Pro Ile Gly Pro Pro Gly Ile Pro Gly Phe Pro Gly Ser
                455                 460                 465

Lys Gly Asp Pro Gly Ser Pro Gly Pro Pro Gly Pro Ala Gly Ile
                470                 475                 480

Ala Thr Lys Gly Leu Asn Gly Pro Thr Gly Pro Pro Gly Pro Pro
                485                 490                 495

Gly Pro Arg Gly Pro Ser Gly Glu Pro Gly Leu Pro Gly Pro Pro
                500                 505                 510

Gly Pro Pro Gly Pro Pro Gly Gln Ala Val Met Pro Glu Gly Phe
                515                 520                 525
```

1453
1445

```
Ile Lys Ala Gly Gln Arg Pro Ser Leu Ser Gly Thr Pro Leu Val
                530                 535                 540

Ser Ala Asn Gln Gly Val Thr Gly Met Pro Val Ser Ala Phe Thr
                545                 550                 555

Val Ile Leu Ser Lys Ala Tyr Pro Ala Ile Gly Thr Pro Ile Pro
                560                 565                 570

Phe Asp Lys Ile Leu Tyr Asn Arg Gln Gln His Tyr Asp Pro Arg
                575                 580                 585

Thr Gly Ile Phe Thr Cys Gln Ile Pro Gly Ile Tyr Tyr Phe Ser
                590                 595                 600

Tyr His Val His Val Lys Gly Thr His Val Trp Val Gly Leu Tyr
                605                 610                 615

Lys Asn Gly Thr Pro Val Met Tyr Thr Tyr Asp Glu Tyr Thr Lys
                620                 625                 630

Gly Tyr Leu Asp Gln Ala Ser Gly Ser Ala Ile Ile Asp Leu Thr
                635                 640                 645

Glu Asn Asp Gln Val Trp Leu Gln Leu Pro Asn Ala Glu Ser Asn
                650                 655                 660

Gly Leu Tyr Ser Ser Glu Tyr Val His Ser Ser Phe Ser Gly Phe
                665                 670                 675

Leu Val Ala Pro Met
                680
```

<210> 84
<211> 542
<212> PRT
<213> Homo sapiens

<400> 84

```
 Met Pro Leu Pro Pro Leu Leu Leu Leu Leu Leu Ala Ala Pro Trp
   1               5                  10                  15

 Gly Arg Ala Val Pro Cys Val Ser Gly Gly Leu Pro Lys Pro Ala
                20                  25                  30

 Asn Ile Thr Phe Leu Ser Ile Asn Met Lys Asn Val Leu Gln Trp
                35                  40                  45

 Thr Pro Pro Glu Gly Leu Gln Gly Val Lys Val Thr Tyr Thr Val
                50                  55                  60

 Gln Tyr Phe Ile Tyr Gly Gln Lys Lys Trp Leu Asn Lys Ser Glu
                65                  70                  75

 Cys Arg Asn Ile Asn Arg Thr Tyr Cys Asp Leu Ser Ala Glu Thr
                80                  85                  90

 Ser Asp Tyr Glu His Gln Tyr Tyr Ala Lys Val Lys Ala Ile Trp
                95                 100                 105

 Gly Thr Lys Cys Ser Lys Trp Ala Glu Ser Gly Arg Phe Tyr Pro
               110                 115                 120

 Phe Leu Glu Thr Gln Ile Gly Pro Pro Glu Val Ala Leu Thr Thr
               125                 130                 135

 Asp Glu Lys Ser Ile Ser Val Val Leu Thr Ala Pro Glu Lys Trp
               140                 145                 150

 Lys Arg Asn Pro Glu Asp Leu Pro Val Ser Met Gln Gln Ile Tyr
               155                 160                 165

 Ser Asn Leu Lys Tyr Asn Val Ser Val Leu Asn Thr Lys Ser Asn
               170                 175                 180
```

1446
1454

Arg Thr Trp Ser Gln Cys Val Thr Asn His Thr Leu Val Leu Thr
185 190 195

Trp Leu Glu Pro Asn Thr Leu Tyr Cys Val His Val Glu Ser Phe
200 205 210

Val Pro Gly Pro Pro Arg Arg Ala Gln Pro Ser Glu Lys Gln Cys
215 220 225

Ala Arg Thr Leu Lys Asp Gln Ser Ser Glu Phe Lys Ala Lys Ile
230 235 240

Ile Phe Trp Tyr Val Leu Pro Ile Ser Ile Thr Val Phe Leu Phe
245 250 255

Ser Val Met Gly Tyr Ser Ile Tyr Arg Tyr Ile His Val Gly Lys
260 265 270

Glu Lys His Pro Ala Asn Leu Ile Leu Ile Tyr Gly Asn Glu Phe
275 280 285

Asp Lys Arg Phe Phe Val Pro Ala Glu Lys Ile Val Ile Asn Phe
290 295 300

Ile Thr Leu Asn Ile Ser Asp Asp Ser Lys Ile Ser His Gln Asp
305 310 315

Met Ser Leu Leu Gly Lys Ser Ser Asp Val Ser Ser Leu Asn Asp
320 325 330

Pro Gln Pro Ser Gly Asn Leu Arg Pro Pro Gln Glu Glu Glu Glu
335 340 345

Val Lys His Leu Gly Tyr Ala Ser His Leu Met Glu Ile Phe Cys
350 355 360

Asp Ser Glu Glu Asn Thr Glu Gly Thr Ser Leu Thr Gln Gln Glu
365 370 375

Ser Leu Ser Arg Thr Ile Pro Pro Asp Lys Thr Val Ile Glu Tyr
380 385 390

Glu Tyr Asp Val Arg Thr Thr Asp Ile Cys Ala Gly Pro Glu Glu
395 400 405

Gln Glu Leu Ser Leu Gln Glu Glu Val Ser Thr Gln Gly Thr Leu
410 415 420

Leu Glu Ser Gln Ala Ala Leu Ala Val Leu Gly Pro Gln Thr Leu
425 430 435

Gln Tyr Ser Tyr Thr Pro Gln Leu Gln Asp Leu Asp Pro Leu Ala
440 445 450

Gln Glu His Thr Asp Ser Glu Glu Gly Pro Glu Glu Glu Pro Ser
455 460 465

Thr Thr Leu Val Asp Trp Asp Pro Gln Thr Gly Arg Leu Cys Ile
470 475 480

Pro Ser Leu Ser Ser Phe Asp Gln Asp Ser Glu Gly Cys Glu Pro
485 490 495

Ser Glu Gly Asp Gly Leu Gly Glu Glu Gly Leu Leu Ser Arg Leu
500 505 510

Tyr Glu Glu Pro Ala Pro Asp Arg Pro Pro Gly Glu Asn Glu Thr
515 520 525

Tyr Leu Met Gln Phe Met Glu Glu Trp Gly Leu Tyr Val Gln Met
530 535 540

Glu Asn

<210> 85

<211> 553
<212> PRT
<213> Homo sapiens

<400> 85

```
Met Arg Ala Pro Gly Arg Pro Ala Leu Arg Pro Leu Pro Leu Pro
  1               5                  10                  15

Pro Leu Leu Leu Leu Leu Leu Ala Ala Pro Trp Gly Arg Ala Val
             20                  25                  30

Pro Cys Val Ser Gly Gly Leu Pro Lys Pro Ala Asn Ile Thr Phe
             35                  40                  45

Leu Ser Ile Asn Met Lys Asn Val Leu Gln Trp Thr Pro Pro Glu
             50                  55                  60

Gly Leu Gln Gly Val Lys Val Thr Tyr Thr Val Gln Tyr Phe Ile
             65                  70                  75

Tyr Gly Gln Lys Lys Trp Leu Asn Lys Ser Glu Cys Arg Asn Ile
             80                  85                  90

Asn Arg Thr Tyr Cys Asp Leu Ser Ala Glu Thr Ser Asp Tyr Glu
             95                 100                 105

His Gln Tyr Tyr Ala Lys Val Lys Ala Ile Trp Gly Thr Lys Cys
            110                 115                 120

Ser Lys Trp Ala Glu Ser Gly Arg Phe Tyr Pro Phe Leu Glu Thr
            125                 130                 135

Gln Ile Gly Pro Pro Glu Val Ala Leu Thr Thr Asp Glu Lys Ser
            140                 145                 150

Ile Ser Val Val Leu Thr Ala Pro Glu Lys Trp Lys Arg Asn Pro
            155                 160                 165

Glu Asp Leu Pro Val Ser Met Gln Gln Ile Tyr Ser Asn Leu Lys
            170                 175                 180

Tyr Asn Val Ser Val Leu Asn Thr Lys Ser Asn Arg Thr Trp Ser
            185                 190                 195

Gln Cys Val Thr Asn His Thr Leu Val Leu Thr Trp Leu Glu Pro
            200                 205                 210

Asn Thr Leu Tyr Cys Val His Val Glu Ser Phe Val Pro Gly Pro
            215                 220                 225

Pro Arg Arg Ala Gln Pro Ser Glu Lys Gln Cys Ala Arg Thr Leu
            230                 235                 240

Lys Asp Gln Ser Ser Glu Phe Lys Ala Lys Ile Ile Phe Trp Tyr
            245                 250                 255

Val Leu Pro Ile Ser Ile Thr Val Phe Leu Phe Ser Val Met Gly
            260                 265                 270

Tyr Ser Ile Tyr Arg Tyr Ile His Val Gly Lys Glu Lys His Pro
            275                 280                 285

Ala Asn Leu Ile Leu Ile Tyr Gly Asn Glu Phe Asp Lys Arg Phe
            290                 295                 300

Phe Val Pro Ala Glu Lys Ile Val Ile Asn Phe Ile Thr Leu Asn
            305                 310                 315

Ile Ser Asp Asp Ser Lys Ile Ser His Gln Asp Met Ser Leu Leu
            320                 325                 330

Gly Lys Ser Ser Asp Val Ser Ser Leu Asn Asp Pro Gln Pro Ser
            335                 340                 345

Gly Asn Leu Arg Pro Pro Gln Glu Glu Glu Glu Val Lys His Leu
            350                 355                 360
```

Gly Tyr Ala Ser His Leu Met Glu Ile Phe Cys Asp Ser Glu Glu
365 370 375

Asn Thr Glu Gly Thr Ser Leu Thr Gln Gln Glu Ser Leu Ser Arg
380 385 390

Thr Ile Pro Pro Asp Lys Thr Val Ile Glu Tyr Glu Tyr Asp Val
395 400 405

Arg Thr Thr Asp Ile Cys Ala Gly Pro Glu Glu Gln Glu Leu Ser
410 415 420

Leu Gln Glu Glu Val Ser Thr Gln Gly Thr Leu Leu Glu Ser Gln
425 430 435

Ala Ala Leu Ala Val Leu Gly Pro Gln Thr Leu Gln Tyr Ser Tyr
440 445 450

Thr Pro Gln Leu Gln Asp Leu Asp Pro Leu Ala Gln Glu His Thr
455 460 465

Asp Ser Glu Glu Gly Pro Glu Glu Glu Pro Ser Thr Thr Leu Val
470 475 480

Asp Trp Asp Pro Gln Thr Gly Arg Leu Cys Ile Pro Ser Leu Ser
485 490 495

Ser Phe Asp Gln Asp Ser Glu Gly Cys Glu Pro Ser Glu Gly Asp
500 505 510

Gly Leu Gly Glu Glu Gly Leu Leu Ser Arg Leu Tyr Glu Glu Pro
515 520 525

Ala Pro Asp Arg Pro Pro Gly Glu Asn Glu Thr Tyr Leu Met Gln
530 535 540

Phe Met Glu Glu Trp Gly Leu Tyr Val Gln Met Glu Asn
545 550

<210> 86
<211> 95
<212> PRT
<213> Homo sapiens

<400> 86
Met Ala Ala Ala Arg Leu Cys Leu Ser Leu Leu Leu Leu Ser Thr
1 5 10 15

Cys Val Ala Leu Leu Leu Gln Pro Leu Leu Gly Ala Gln Gly Ala
20 25 30

Pro Leu Glu Pro Val Tyr Pro Gly Asp Asn Ala Thr Pro Glu Gln
35 40 45

Met Ala Gln Tyr Ala Ala Asp Leu Arg Arg Tyr Ile Asn Met Leu
50 55 60

Thr Arg Pro Arg Tyr Gly Lys Arg His Lys Glu Asp Thr Leu Ala
65 70 75

Phe Ser Glu Trp Gly Ser Pro His Ala Ala Val Pro Arg Glu Leu
80 85 90

Ser Pro Leu Asp Leu
95

<210> 87
<211> 344
<212> PRT
<213> Homo sapiens

<400> 87
Met Gly Pro Pro Ser Ala Pro Pro Cys Arg Leu His Val Pro Trp
1 5 10 15

Lys Glu Val Leu Leu Thr Ala Ser Leu Leu Thr Phe Trp Asn Pro
20 25 30

Pro Thr Thr Ala Lys Leu Thr Ile Glu Ser Thr Pro Phe Asn Val
35 40 45

Ala Glu Gly Lys Glu Val Leu Leu Leu Ala His Asn Leu Pro Gln
50 55 60

Asn Arg Ile Gly Tyr Ser Trp Tyr Lys Gly Glu Arg Val Asp Gly
65 70 75

Asn Ser Leu Ile Val Gly Tyr Val Ile Gly Thr Gln Gln Ala Thr
80 85 90

Pro Gly Pro Ala Tyr Ser Gly Arg Glu Thr Ile Tyr Pro Asn Ala
95 100 105

Ser Leu Leu Ile Gln Asn Val Thr Gln Asn Asp Thr Gly Phe Tyr
110 115 120

Thr Leu Gln Val Ile Lys Ser Asp Leu Val Asn Glu Glu Ala Thr
125 130 135

Gly Gln Phe His Val Tyr Pro Glu Leu Pro Lys Pro Ser Ile Ser
140 145 150

Ser Asn Asn Ser Asn Pro Val Glu Asp Lys Asp Ala Val Ala Phe
155 160 165

Thr Cys Glu Pro Glu Val Gln Asn Thr Thr Tyr Leu Trp Trp Val
170 175 180

Asn Gly Gln Ser Leu Pro Val Ser Pro Arg Leu Gln Leu Ser Asn
185 190 195

Gly Asn Met Thr Leu Thr Leu Leu Ser Val Lys Arg Asn Asp Ala
200 205 210

Gly Ser Tyr Glu Cys Glu Ile Gln Asn Pro Ala Ser Ala Asn Arg
215 220 225

Ser Asp Pro Val Thr Leu Asn Val Leu Tyr Gly Pro Asp Gly Pro
230 235 240

Thr Ile Ser Pro Ser Lys Ala Asn Tyr Arg Pro Gly Glu Asn Leu
245 250 255

Asn Leu Ser Cys His Ala Ala Ser Asn Pro Pro Ala Gln Tyr Ser
260 265 270

Trp Phe Ile Asn Gly Thr Phe Gln Gln Ser Thr Gln Glu Leu Phe
275 280 285

Ile Pro Asn Ile Thr Val Asn Asn Ser Gly Ser Tyr Met Cys Gln
290 295 300

Ala His Asn Ser Ala Thr Gly Leu Asn Arg Thr Thr Val Thr Met
305 310 315

Ile Thr Val Ser Gly Ser Ala Pro Val Leu Ser Ala Val Ala Thr
320 325 330

Val Gly Ile Thr Ile Gly Val Leu Ala Arg Val Ala Leu Ile
335 340

<210> 88
<211> 987
<212> PRT
<213> Homo sapiens

<400> 88
Met Glu Leu Arg Val Leu Leu Cys Trp Ala Ser Leu Ala Ala Ala
1 5 10 15

Leu Glu Glu Thr Leu Leu Asn Thr Lys Leu Glu Thr Ala Asp Leu

20 25 30

Lys Trp Val Thr Phe Pro Gln Val Asp Gly Gln Trp Glu Glu Leu
35 40 45

Ser Gly Leu Asp Glu Glu Gln His Ser Val Arg Thr Tyr Glu Val
50 55 60

Cys Asp Val Gln Arg Ala Pro Gly Gln Ala His Trp Leu Arg Thr
65 70 75

Gly Trp Val Pro Arg Arg Gly Ala Val His Val Tyr Ala Thr Leu
80 85 90

Arg Phe Thr Met Leu Glu Cys Leu Ser Leu Pro Arg Ala Gly Arg
95 100 105

Ser Cys Lys Glu Thr Phe Thr Val Phe Tyr Tyr Glu Ser Asp Ala
110 115 120

Asp Thr Ala Thr Ala Leu Thr Pro Ala Trp Met Glu Asn Pro Tyr
125 130 135

Ile Lys Val Asp Thr Val Ala Ala Glu His Leu Thr Arg Lys Arg
140 145 150

Pro Gly Ala Glu Ala Thr Gly Lys Val Asn Val Lys Thr Leu Arg
155 160 165

Leu Gly Pro Leu Ser Lys Ala Gly Phe Tyr Leu Ala Phe Gln Asp
170 175 180

Gln Gly Ala Cys Met Ala Leu Leu Ser Leu His Leu Phe Tyr Lys
185 190 195

Lys Cys Ala Gln Leu Thr Val Asn Leu Thr Arg Phe Pro Glu Thr
200 205 210

Val Pro Arg Glu Leu Val Val Pro Val Ala Gly Ser Cys Val Val
215 220 225

Asp Ala Val Pro Ala Pro Gly Pro Ser Pro Ser Leu Tyr Cys Arg
230 235 240

Glu Asp Gly Gln Trp Ala Glu Gln Pro Val Thr Gly Cys Ser Cys
245 250 255

Ala Pro Gly Phe Glu Ala Ala Glu Gly Asn Thr Lys Cys Arg Ala
260 265 270

Cys Ala Gln Gly Thr Phe Lys Pro Leu Ser Gly Glu Gly Ser Cys
275 280 285

Gln Pro Cys Pro Ala Asn Ser His Ser Asn Thr Ile Gly Ser Ala
290 295 300

Val Cys Gln Cys Arg Val Gly Tyr Phe Arg Ala Arg Thr Asp Pro
305 310 315

Arg Gly Ala Pro Cys Thr Thr Pro Pro Ser Ala Pro Arg Ser Val
320 325 330

Val Ser Arg Leu Asn Gly Ser Ser Leu His Leu Glu Trp Ser Ala
335 340 345

Pro Leu Glu Ser Gly Gly Arg Glu Asp Leu Thr Tyr Ala Leu Arg
350 355 360

Cys Arg Glu Cys Arg Pro Gly Gly Ser Cys Ala Pro Cys Gly Gly
365 370 375

Asp Leu Thr Phe Asp Pro Gly Pro Arg Asp Leu Val Glu Pro Trp
380 385 390

Val Val Val Arg Gly Leu Arg Pro Asp Phe Thr Tyr Thr Phe Glu
395 400 405

Val Thr Ala Leu Asn Gly Val Ser Ser Leu Ala Thr Gly Pro Val
410 415 420

Pro Phe Glu Pro Val Asn Val Thr Thr Asp Arg Glu Val Pro Pro
425 430 435

Ala Val Ser Asp Ile Arg Val Thr Arg Ser Ser Pro Ser Ser Leu
440 445 450

Ser Leu Ala Trp Ala Val Pro Arg Ala Pro Ser Gly Ala Val Leu
455 460 465

Asp Tyr Glu Val Lys Tyr His Glu Lys Gly Ala Glu Gly Pro Ser
470 475 480

Ser Val Arg Phe Leu Lys Thr Ser Glu Asn Arg Ala Glu Leu Arg
485 490 495

Gly Leu Lys Arg Gly Ala Ser Tyr Leu Val Gln Val Arg Ala Arg
500 505 510

Ser Glu Ala Gly Tyr Gly Pro Phe Gly Gln Glu His His Ser Gln
515 520 525

Thr Gln Leu Asp Glu Ser Glu Gly Trp Arg Glu Gln Leu Ala Leu
530 535 540

Ile Ala Gly Thr Ala Val Val Gly Val Val Leu Val Leu Val Val
545 550 555

Ile Val Val Ala Val Leu Cys Leu Arg Lys Gln Ser Asn Gly Arg
560 565 570

Glu Ala Glu Tyr Ser Asp Lys His Gly Gln Tyr Leu Ile Gly His
575 580 585

Gly Thr Lys Val Tyr Ile Asp Pro Phe Thr Tyr Glu Asp Pro Asn
590 595 600

Glu Ala Val Arg Glu Phe Ala Lys Glu Ile Asp Val Ser Tyr Val
605 610 615

Lys Ile Glu Glu Val Ile Gly Ala Gly Glu Phe Gly Glu Val Cys
620 625 630

Arg Gly Arg Leu Lys Ala Pro Gly Lys Lys Glu Ser Cys Val Ala
635 640 645

Ile Lys Thr Leu Lys Gly Gly Tyr Thr Glu Arg Gln Arg Arg Glu
650 655 660

Phe Leu Ser Glu Ala Ser Ile Met Gly Gln Phe Glu His Pro Asn
665 670 675

Ile Ile Arg Leu Glu Gly Val Val Thr Asn Ser Met Pro Val Met
680 685 690

Ile Leu Thr Glu Phe Met Glu Asn Gly Ala Leu Asp Ser Phe Leu
695 700 705

Arg Leu Asn Asp Gly Gln Phe Thr Val Ile Gln Leu Val Gly Met
710 715 720

Leu Arg Gly Ile Ala Ser Gly Met Arg Tyr Leu Ala Glu Met Ser
725 730 735

Tyr Val His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Asn Ser
740 745 750

Asn Leu Val Cys Lys Val Ser Asp Phe Gly Leu Ser Arg Phe Leu
755 760 765

Glu Glu Asn Ser Ser Asp Pro Thr Tyr Thr Ser Ser Leu Gly Gly
770 775 780

```
Lys Ile Pro Ile Arg Trp Thr Ala Pro Glu Ala Ile Ala Phe Arg
                785                 790                 795

Lys Phe Thr Ser Ala Ser Asp Ala Trp Ser Tyr Gly Ile Val Met
                800                 805                 810

Trp Glu Val Met Ser Phe Gly Glu Arg Pro Tyr Trp Asp Met Ser
                815                 820                 825

Asn Gln Asp Val Ile Asn Ala Ile Glu Gln Asp Tyr Arg Leu Pro
                830                 835                 840

Pro Pro Pro Asp Cys Pro Thr Ser Leu His Gln Leu Met Leu Asp
                845                 850                 855

Cys Trp Gln Lys Asp Arg Asn Ala Arg Pro Arg Phe Pro Gln Val
                860                 865                 870

Val Ser Ala Leu Asp Lys Met Ile Arg Asn Pro Ala Ser Leu Lys
                875                 880                 885

Ile Val Ala Arg Glu Asn Gly Gly Ala Ser His Pro Leu Leu Asp
                890                 895                 900

Gln Arg Gln Pro His Tyr Ser Ala Phe Gly Ser Val Gly Glu Trp
                905                 910                 915

Leu Arg Ala Ile Lys Met Gly Arg Tyr Glu Glu Ser Phe Ala Ala
                920                 925                 930

Ala Gly Phe Gly Ser Phe Glu Leu Val Ser Gln Ile Ser Ala Glu
                935                 940                 945

Asp Leu Leu Arg Ile Gly Val Thr Leu Ala Gly His Gln Lys Lys
                950                 955                 960

Ile Leu Ala Ser Val Gln His Met Lys Ser Gln Ala Lys Pro Gly
                965                 970                 975

Thr Pro Gly Gly Thr Gly Gly Pro Ala Pro Gln Tyr
                980                 985

<210> 89
<211> 911
<212> PRT
<213> Homo sapiens

<400> 89
Met Ala Gln Leu Phe Leu Pro Leu Leu Ala Ala Leu Val Leu Ala
  1               5                  10                  15

Gln Ala Pro Ala Ala Leu Ala Asp Val Leu Glu Gly Asp Ser Ser
                20                  25                  30

Glu Asp Arg Ala Phe Arg Val Arg Ile Ala Gly Asp Ala Pro Leu
                35                  40                  45

Gln Gly Val Leu Gly Gly Ala Leu Thr Ile Pro Cys His Val His
                50                  55                  60

Tyr Leu Arg Pro Pro Pro Ser Arg Arg Ala Val Leu Gly Ser Pro
                65                  70                  75

Arg Val Lys Trp Thr Phe Leu Ser Arg Gly Arg Glu Ala Glu Val
                80                  85                  90

Leu Val Ala Arg Gly Val Arg Val Lys Val Asn Glu Ala Tyr Arg
                95                 100                 105

Phe Arg Val Ala Leu Pro Ala Tyr Pro Ala Ser Leu Thr Asp Val
               110                 115                 120

Ser Leu Ala Leu Ser Glu Leu Arg Pro Asn Asp Ser Gly Ile Tyr
               125                 130                 135

Arg Cys Glu Val Gln His Gly Ile Asp Asp Ser Ser Asp Ala Val
```

```
                  140                 145                 150
Glu Val Lys Val Lys Gly Val Val Phe Leu Tyr Arg Glu Gly Ser
                  155                 160                 165

Ala Arg Tyr Ala Phe Ser Phe Ser Gly Ala Gln Glu Ala Cys Ala
                  170                 175                 180

Arg Ile Gly Ala His Ile Ala Thr Pro Glu Gln Leu Tyr Ala Ala
                  185                 190                 195

Tyr Leu Gly Gly Tyr Glu Gln Cys Asp Ala Gly Trp Leu Ser Asp
                  200                 205                 210

Gln Thr Val Arg Tyr Pro Ile Gln Thr Pro Arg Glu Ala Cys Tyr
                  215                 220                 225

Gly Asp Met Asp Gly Phe Pro Gly Val Arg Asn Tyr Gly Val Val
                  230                 235                 240

Asp Pro Asp Asp Leu Tyr Asp Val Tyr Cys Tyr Ala Glu Asp Leu
                  245                 250                 255

Asn Gly Glu Leu Phe Leu Gly Asp Pro Pro Glu Lys Leu Thr Leu
                  260                 265                 270

Glu Glu Ala Arg Ala Tyr Cys Gln Glu Arg Gly Ala Glu Ile Ala
                  275                 280                 285

Thr Thr Gly Gln Leu Tyr Ala Ala Trp Asp Gly Gly Leu Asp His
                  290                 295                 300

Cys Ser Pro Gly Trp Leu Ala Asp Gly Ser Val Arg Tyr Pro Ile
                  305                 310                 315

Val Thr Pro Ser Gln Arg Cys Gly Gly Gly Leu Pro Gly Val Lys
                  320                 325                 330

Thr Leu Phe Leu Phe Pro Asn Gln Thr Gly Phe Pro Asn Lys His
                  335                 340                 345

Ser Arg Phe Asn Val Tyr Cys Phe Arg Asp Ser Ala Gln Pro Ser
                  350                 355                 360

Ala Ile Pro Glu Ala Ser Asn Pro Ala Ser Asn Pro Ala Ser Asp
                  365                 370                 375

Gly Leu Glu Ala Ile Val Thr Val Thr Glu Thr Leu Glu Glu Leu
                  380                 385                 390

Gln Leu Pro Gln Glu Ala Thr Glu Ser Glu Ser Arg Gly Ala Ile
                  395                 400                 405

Tyr Ser Ile Pro Ile Met Glu Asp Gly Gly Gly Gly Ser Ser Thr
                  410                 415                 420

Pro Glu Asp Pro Ala Glu Ala Pro Arg Thr Leu Leu Glu Phe Glu
                  425                 430                 435

Thr Gln Ser Met Val Pro Pro Thr Gly Phe Ser Glu Glu Glu Gly
                  440                 445                 450

Lys Ala Leu Glu Glu Glu Glu Lys Tyr Glu Asp Glu Glu Glu Lys
                  455                 460                 465

Glu Glu Glu Glu Glu Glu Glu Glu Val Glu Asp Glu Ala Leu Trp
                  470                 475                 480

Ala Trp Pro Ser Glu Leu Ser Ser Pro Gly Pro Glu Ala Ser Leu
                  485                 490                 495

Pro Thr Glu Pro Ala Ala Gln Glu Lys Ser Leu Ser Gln Ala Pro
                  500                 505                 510

Ala Arg Ala Val Leu Gln Pro Gly Ala Ser Pro Leu Pro Asp Gly
                  515                 520                 525
```

1462

Glu Ser Glu Ala Ser Arg Pro Pro Arg Val His Gly Pro Pro Thr
530 535 540

Glu Thr Leu Pro Thr Pro Arg Glu Arg Asn Leu Ala Ser Pro Ser
545 550 555

Pro Ser Thr Leu Val Glu Ala Arg Glu Val Gly Glu Ala Thr Gly
560 565 570

Gly Pro Glu Leu Ser Gly Val Pro Arg Gly Glu Ser Glu Glu Thr
575 580 585

Gly Ser Ser Glu Gly Ala Pro Ser Leu Leu Pro Ala Thr Arg Ala
590 595 600

Pro Glu Gly Thr Arg Glu Leu Glu Ala Pro Ser Glu Asp Asn Ser
605 610 615

Gly Arg Thr Ala Pro Ala Gly Thr Ser Val Gln Ala Gln Pro Val
620 625 630

Leu Pro Thr Asp Ser Ala Ser Arg Gly Gly Val Ala Val Val Pro
635 640 645

Ala Ser Gly Asp Cys Val Pro Ser Pro Cys His Asn Gly Gly Thr
650 655 660

Cys Leu Glu Glu Glu Glu Gly Val Arg Cys Leu Cys Leu Pro Gly
665 670 675

Tyr Gly Gly Asp Leu Cys Asp Val Gly Leu Arg Phe Cys Asn Pro
680 685 690

Gly Trp Asp Ala Phe Gln Gly Ala Cys Tyr Lys His Phe Ser Thr
695 700 705

Arg Arg Ser Trp Glu Glu Ala Glu Thr Gln Cys Arg Met Tyr Gly
710 715 720

Ala His Leu Ala Ser Ile Ser Thr Pro Glu Glu Gln Asp Phe Ile
725 730 735

Asn Asn Arg Tyr Arg Glu Tyr Gln Trp Ile Gly Leu Asn Asp Arg
740 745 750

Thr Ile Glu Gly Asp Phe Leu Trp Ser Asp Gly Val Pro Leu Leu
755 760 765

Tyr Glu Asn Trp Asn Pro Gly Gln Pro Asp Ser Tyr Phe Leu Ser
770 775 780

Gly Glu Asn Cys Val Val Met Val Trp His Asp Gln Gly Gln Trp
785 790 795

Ser Asp Val Pro Cys Asn Tyr His Leu Ser Tyr Thr Cys Lys Met
800 805 810

Gly Leu Val Ser Cys Gly Pro Pro Pro Glu Leu Pro Leu Ala Gln
815 820 825

Val Phe Gly Arg Pro Arg Leu Arg Tyr Glu Val Asp Thr Val Leu
830 835 840

Arg Tyr Arg Cys Arg Glu Gly Leu Ala Gln Arg Asn Leu Pro Leu
845 850 855

Ile Arg Cys Gln Glu Asn Gly Arg Trp Glu Ala Pro Gln Ile Ser
860 865 870

Cys Val Pro Arg Arg Pro Ala Arg Ala Leu His Pro Glu Glu Asp
875 880 885

Pro Glu Gly Arg Gln Gly Arg Leu Leu Gly Arg Trp Lys Ala Leu
890 895 900

1463

Leu Ile Pro Pro Ser Ser Pro Met Pro Gly Pro
905 910

<210> 90
<211> 671
<212> PRT
<213> Homo sapiens

<400> 90
Met Ala Gln Leu Phe Leu Pro Leu Leu Ala Ala Leu Val Leu Ala
1 5 10 15

Gln Ala Pro Ala Ala Leu Ala Asp Val Leu Glu Gly Asp Ser Ser
20 25 30

Glu Asp Arg Ala Phe Arg Val Arg Ile Ala Gly Asp Ala Pro Leu
35 40 45

Gln Gly Val Leu Gly Gly Ala Leu Thr Ile Pro Cys His Val His
50 55 60

Tyr Leu Arg Pro Pro Pro Ser Arg Arg Ala Val Leu Gly Ser Pro
65 70 75

Arg Val Lys Trp Thr Phe Leu Ser Arg Gly Arg Glu Ala Glu Val
80 85 90

Leu Val Ala Arg Gly Val Arg Val Lys Val Asn Glu Ala Tyr Arg
95 100 105

Phe Arg Val Ala Leu Pro Ala Tyr Pro Ala Ser Leu Thr Asp Val
110 115 120

Ser Leu Ala Leu Ser Glu Leu Arg Pro Asn Asp Ser Gly Ile Tyr
125 130 135

Arg Cys Glu Val Gln His Gly Ile Asp Asp Ser Ser Asp Ala Val
140 145 150

Glu Val Lys Val Lys Gly Val Val Phe Leu Tyr Arg Glu Gly Ser
155 160 165

Ala Arg Tyr Ala Phe Ser Phe Ser Gly Ala Gln Glu Ala Cys Ala
170 175 180

Arg Ile Gly Ala His Ile Ala Thr Pro Glu Gln Leu Tyr Ala Ala
185 190 195

Tyr Leu Gly Gly Tyr Glu Gln Cys Asp Ala Gly Trp Leu Ser Asp
200 205 210

Gln Thr Val Arg Tyr Pro Ile Gln Thr Pro Arg Glu Ala Cys Tyr
215 220 225

Gly Asp Met Asp Gly Phe Pro Gly Val Arg Asn Tyr Gly Val Val
230 235 240

Asp Pro Asp Asp Leu Tyr Asp Val Tyr Cys Tyr Ala Glu Asp Leu
245 250 255

Asn Gly Glu Leu Phe Leu Gly Asp Pro Pro Glu Lys Leu Thr Leu
260 265 270

Glu Glu Ala Arg Ala Tyr Cys Gln Glu Arg Gly Ala Glu Ile Ala
275 280 285

Thr Thr Gly Gln Leu Tyr Ala Ala Trp Asp Gly Gly Leu Asp His
290 295 300

Cys Ser Pro Gly Trp Leu Ala Asp Gly Ser Val Arg Tyr Pro Ile
305 310 315

Val Thr Pro Ser Gln Arg Cys Gly Gly Gly Leu Pro Gly Val Lys
320 325 330

Thr Leu Phe Leu Phe Pro Asn Gln Thr Gly Phe Pro Asn Lys His

335 340 345

Ser Arg Phe Asn Val Tyr Cys Phe Arg Asp Ser Ala Gln Pro Ser
350 355 360

Ala Ile Pro Glu Ala Ser Asn Pro Ala Ser Asn Pro Ala Ser Asp
365 370 375

Gly Leu Glu Ala Ile Val Thr Val Thr Glu Thr Leu Glu Glu Leu
380 385 390

Gln Leu Pro Gln Glu Ala Thr Glu Ser Glu Ser Arg Gly Ala Ile
395 400 405

Tyr Ser Ile Pro Ile Met Glu Asp Gly Gly Gly Gly Ser Ser Thr
410 415 420

Pro Glu Asp Pro Ala Glu Ala Pro Arg Thr Leu Leu Glu Phe Glu
425 430 435

Thr Gln Ser Met Val Pro Pro Thr Gly Phe Ser Glu Glu Glu Gly
440 445 450

Lys Ala Leu Glu Glu Glu Glu Lys Tyr Glu Asp Glu Glu Glu Lys
455 460 465

Glu Glu Glu Glu Glu Glu Glu Glu Val Glu Asp Glu Ala Leu Trp
470 475 480

Ala Trp Pro Ser Glu Leu Ser Ser Pro Gly Pro Glu Ala Ser Leu
485 490 495

Pro Thr Glu Pro Ala Ala Gln Glu Glu Ser Leu Ser Gln Ala Pro
500 505 510

Ala Arg Ala Val Leu Gln Pro Gly Ala Ser Pro Leu Pro Asp Gly
515 520 525

Glu Ser Glu Ala Ser Arg Pro Pro Arg Val His Gly Pro Pro Thr
530 535 540

Glu Thr Leu Pro Thr Pro Arg Glu Arg Asn Leu Ala Ser Pro Ser
545 550 555

Pro Ser Thr Leu Val Glu Ala Arg Glu Val Gly Glu Ala Thr Gly
560 565 570

Gly Pro Glu Leu Ser Gly Val Pro Arg Gly Glu Ser Glu Glu Thr
575 580 585

Gly Ser Ser Glu Gly Ala Pro Ser Leu Leu Pro Ala Thr Arg Ala
590 595 600

Pro Glu Gly Thr Arg Glu Leu Glu Ala Pro Ser Glu Asp Asn Ser
605 610 615

Gly Arg Thr Ala Pro Ala Gly Thr Ser Val Gln Ala Gln Pro Val
620 625 630

Leu Pro Thr Asp Ser Ala Ser Arg Gly Gly Val Ala Val Val Pro
635 640 645

Ala Ser Gly Asn Ser Ala Gln Gly Ser Thr Ala Leu Ser Ile Leu
650 655 660

Leu Leu Phe Phe Pro Leu Gln Leu Trp Val Thr *
665 670

<210> 91
<211> 86
<212> PRT
<213> Homo sapiens

<400> 91
Met Ala Val Arg Gln Trp Val Ile Ala Leu Ala Leu Ala Ala Leu
1 5 10 15

1465 last

Leu Val Val Asp Arg Glu Val Pro Val Ala Ala Gly Lys Leu Pro
20 25 30

Phe Ser Arg Met Pro Ile Cys Glu His Met Val Glu Ser Pro Thr
35 40 45

Cys Ser Gln Met Ser Asn Leu Val Cys Gly Thr Asp Gly Leu Thr
50 55 60

Tyr Thr Asn Glu Cys Gln Leu Cys Leu Ala Arg Ile Lys Thr Lys
65 70 75

Gln Asp Ile Gln Ile Met Lys Asp Gly Lys Cys
80 85

<210> 92
<211> 459
<212> PRT
<213> Homo sapiens

<400> 92
Met Ala Pro Leu Cys Pro Ser Pro Trp Leu Pro Leu Leu Ile Pro
1 5 10 15

Ala Pro Ala Pro Gly Leu Thr Val Gln Leu Leu Leu Ser Leu Leu
20 25 30

Leu Leu Val Pro Val His Pro Gln Arg Leu Pro Arg Met Gln Glu
35 40 45

Asp Ser Pro Leu Gly Gly Gly Ser Ser Gly Glu Asp Asp Pro Leu
50 55 60

Gly Glu Glu Asp Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu
65 70 75

Asp Pro Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro
80 85 90

Gly Glu Glu Asp Leu Pro Glu Val Lys Pro Lys Ser Glu Glu Glu
95 100 105

Gly Ser Leu Lys Leu Glu Asp Leu Pro Thr Val Glu Ala Pro Gly
110 115 120

Asp Pro Gln Glu Pro Gln Asn Asn Ala His Arg Asp Lys Glu Gly
125 130 135

Asp Asp Gln Ser His Trp Arg Tyr Gly Gly Asp Pro Pro Trp Pro
140 145 150

Arg Val Ser Pro Ala Cys Ala Gly Arg Phe Gln Ser Pro Val Asp
155 160 165

Ile Arg Pro Gln Leu Ala Ala Phe Cys Pro Ala Leu Arg Pro Leu
170 175 180

Glu Leu Leu Gly Phe Gln Leu Pro Pro Leu Pro Glu Leu Arg Leu
185 190 195

Arg Asn Asn Gly His Ser Val Gln Leu Thr Leu Pro Pro Gly Leu
200 205 210

Glu Met Ala Leu Gly Pro Gly Arg Glu Tyr Arg Ala Leu Gln Leu
215 220 225

His Leu His Trp Gly Ala Ala Gly Arg Pro Gly Ser Glu His Thr
230 235 240

Val Glu Gly His Arg Phe Pro Ala Glu Ile His Val Val His Leu
245 250 255

Ser Thr Ala Phe Ala Arg Val Asp Glu Ala Leu Gly Arg Pro Gly
260 265 270

Gly Leu Ala Val Leu Ala Ala Phe Leu Glu Glu Gly Pro Glu Glu
275 280 285

Asn Ser Ala Tyr Glu Gln Leu Leu Ser Arg Leu Glu Glu Ile Ala
290 295 300

Glu Glu Gly Ser Glu Thr Gln Val Pro Gly Leu Asp Ile Ser Ala
305 310 315

Leu Leu Pro Ser Asp Phe Ser Arg Tyr Phe Arg Tyr Glu Gly Ser
320 325 330

Leu Thr Thr Pro Pro Cys Ala Gln Gly Val Ile Trp Thr Val Phe
335 340 345

Asn Gln Thr Val Met Leu Ser Ala Lys Gln Leu His Thr Leu Ser
350 355 360

Asp Thr Leu Trp Gly Pro Gly Asp Ser Arg Leu Gln Leu Asn Phe
365 370 375

Arg Ala Thr Gln Pro Leu Asn Gly Arg Val Ile Glu Ala Ser Phe
380 385 390

Pro Ala Gly Val Asp Ser Ser Pro Arg Ala Ala Glu Pro Val Gln
395 400 405

Leu Asn Ser Cys Leu Ala Ala Gly Asp Ile Leu Ala Leu Val Phe
410 415 420

Gly Leu Leu Phe Ala Val Thr Ser Val Ala Phe Leu Val Gln Met
425 430 435

Arg Arg Gln His Arg Arg Gly Thr Lys Gly Gly Val Ser Tyr Arg
440 445 450

Pro Ala Glu Val Ala Glu Thr Gly Ala
455

<210> 93
<211> 375
<212> PRT
<213> Homo sapiens

<400> 93
Met Asp Ala Leu Gln Leu Ala Asn Ser Ala Phe Ala Val Asp Leu
1 5 10 15

Phe Lys Gln Leu Cys Glu Lys Glu Pro Leu Gly Asn Val Leu Phe
20 25 30

Ser Pro Ile Cys Leu Ser Thr Ser Leu Ser Leu Ala Gln Val Gly
35 40 45

Ala Lys Gly Asp Thr Ala Asn Glu Ile Gly Gln Val Leu His Phe
50 55 60

Glu Asn Val Lys Asp Val Pro Phe Gly Phe Gln Thr Val Thr Ser
65 70 75

Asp Val Asn Lys Leu Ser Ser Phe Tyr Ser Leu Lys Leu Ile Lys
80 85 90

Arg Leu Tyr Val Asp Lys Ser Leu Asn Leu Ser Thr Glu Phe Ile
95 100 105

Ser Ser Thr Lys Arg Pro Tyr Ala Lys Glu Leu Glu Thr Val Asp
110 115 120

Phe Lys Asp Lys Leu Glu Glu Thr Lys Gly Gln Ile Asn Asn Ser
125 130 135

Ile Lys Asp Leu Thr Asp Gly His Phe Glu Asn Ile Leu Ala Asp
140 145 150

Asn Ser Val Asn Asp Gln Thr Lys Ile Leu Val Val Asn Ala Ala

155 160 165

Tyr Phe Val Gly Lys Trp Met Lys Lys Phe Ser Glu Ser Glu Thr
170 175 180

Lys Glu Cys Pro Phe Arg Val Asn Lys Thr Asp Thr Lys Pro Val
185 190 195

Gln Met Met Asn Met Glu Ala Thr Phe Cys Met Gly Asn Ile Asp
200 205 210

Ser Ile Asn Cys Lys Ile Ile Glu Leu Pro Phe Gln Asn Lys His
215 220 225

Leu Ser Met Phe Ile Leu Leu Pro Lys Asp Val Glu Asp Glu Ser
230 235 240

Thr Gly Leu Glu Lys Ile Glu Lys Gln Leu Asn Ser Glu Ser Leu
245 250 255

Ser Gln Trp Thr Asn Pro Ser Thr Met Ala Asn Ala Lys Val Lys
260 265 270

Leu Ser Ile Pro Lys Phe Lys Val Glu Lys Met Ile Asp Pro Lys
275 280 285

Ala Cys Leu Glu Asn Leu Gly Leu Lys His Ile Phe Ser Glu Asp
290 295 300

Thr Ser Asp Phe Ser Gly Met Ser Glu Thr Lys Gly Val Ala Leu
305 310 315

Ser Asn Val Ile His Lys Val Cys Leu Glu Ile Thr Glu Asp Gly
320 325 330

Gly Asp Ser Ile Glu Val Pro Gly Ala Arg Ile Leu Gln His Lys
335 340 345

Asp Glu Leu Asn Ala Asp His Pro Phe Ile Tyr Ile Ile Arg His
350 355 360

Asn Lys Thr Arg Asn Ile Ile Phe Phe Gly Lys Phe Cys Ser Pro
365 370 375

<210> 94
<211> 260
<212> PRT
<213> Homo sapiens

<400> 94
Met Asp Arg His Ser Ser Tyr Ile Phe Ile Trp Leu Gln Leu Glu
1 5 10 15

Leu Cys Ala Met Ala Val Leu Leu Thr Lys Gly Glu Ile Arg Cys
20 25 30

Tyr Cys Asp Ala Ala His Cys Val Ala Thr Gly Tyr Met Cys Lys
35 40 45

Ser Glu Leu Ser Ala Cys Phe Ser Arg Leu Leu Asp Pro Gln Asn
50 55 60

Ser Asn Ser Pro Leu Thr His Gly Cys Leu Asp Ser Leu Ala Ser
65 70 75

Thr Thr Asp Ile Cys Gln Ala Lys Gln Ala Arg Asn His Ser Gly
80 85 90

Thr Thr Ile Pro Thr Leu Glu Cys Cys His Glu Asp Met Cys Asn
95 100 105

Tyr Arg Gly Leu His Asp Val Leu Ser Pro Pro Arg Gly Glu Ala
110 115 120

Ser Gly Gln Gly Asn Arg Tyr Gln His Asp Gly Ser Arg Asn Leu
125 130 135

```
Ile Thr Lys Val Gln Glu Leu Thr Ser Ser Lys Glu Leu Trp Phe
                140                 145                 150

Arg Ala Ala Val Ile Ala Val Pro Ile Ala Gly Gly Leu Ile Leu
                155                 160                 165

Val Leu Leu Ile Met Leu Ala Leu Arg Met Leu Arg Ser Glu Asn
                170                 175                 180

Lys Arg Leu Gln Asp Gln Arg Gln Gln Met Leu Ser Arg Leu His
                185                 190                 195

Tyr Ser Phe His Gly His His Ser Lys Lys Gly Gln Val Ala Lys
                200                 205                 210

Leu Asp Leu Glu Cys Met Val Pro Val Ser Gly His Glu Asn Cys
                215                 220                 225

Cys Leu Thr Cys Asp Lys Met Arg Gln Ala Asp Leu Ser Asn Asp
                230                 235                 240

Lys Ile Leu Ser Leu Val His Trp Gly Met Tyr Ser Gly His Gly
                245                 250                 255

Lys Leu Glu Phe Val
                260
```

<210> 95
<211> 1095
<212> PRT
<213> Homo sapiens

<400> 95

```
Met Arg Asn Arg Arg Asn Asp Thr Leu Asp Ser Thr Arg Thr Leu
  1               5                  10                  15

Tyr Ser Ser Ala Ser Arg Ser Thr Asp Leu Ser Tyr Ser Glu Ser
                20                  25                  30

Asp Leu Val Asn Phe Ile Gln Ala Asn Phe Lys Lys Arg Glu Cys
                35                  40                  45

Val Phe Phe Thr Lys Asp Ser Lys Ala Thr Glu Asn Val Cys Lys
                50                  55                  60

Cys Gly Tyr Ala Gln Ser Gln His Met Glu Gly Thr Gln Ile Asn
                65                  70                  75

Gln Ser Glu Lys Trp Asn Tyr Lys Lys His Thr Lys Glu Phe Pro
                80                  85                  90

Thr Asp Ala Phe Gly Asp Ile Gln Phe Glu Thr Leu Gly Lys Lys
                95                  100                 105

Gly Lys Tyr Ile Arg Leu Ser Cys Asp Thr Asp Ala Glu Ile Leu
                110                 115                 120

Tyr Glu Leu Leu Thr Gln His Trp His Leu Lys Thr Pro Asn Leu
                125                 130                 135

Val Ile Ser Val Thr Gly Gly Ala Lys Asn Phe Ala Leu Lys Pro
                140                 145                 150

Arg Met Arg Lys Ile Phe Ser Arg Leu Ile Tyr Ile Ala Gln Ser
                155                 160                 165

Lys Gly Ala Trp Ile Leu Thr Gly Gly Thr His Tyr Gly Leu Thr
                170                 175                 180

Lys Tyr Ile Gly Glu Val Val Arg Asp Asn Thr Ile Ser Arg Ser
                185                 190                 195

Ser Glu Glu Asn Ile Val Ala Ile Gly Ile Ala Ala Trp Gly Met
                200                 205                 210
```

Val Ser Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu Gly
215 220 225

Tyr Phe Leu Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro
230 235 240

Leu Tyr Ile Leu Asp Asn Asn His Thr His Leu Leu Leu Val Asp
245 250 255

Asn Gly Cys His Gly His Pro Thr Val Glu Ala Lys Leu Arg Asn
260 265 270

Gln Leu Glu Lys His Ile Ser Glu Arg Thr Ile Gln Asp Ser Asn
275 280 285

Tyr Gly Gly Lys Ile Pro Ile Val Cys Phe Ala Gln Gly Gly Gly
290 295 300

Lys Glu Thr Leu Lys Ala Ile Asn Thr Ser Ile Lys Asn Lys Ile
305 310 315

Pro Cys Val Val Val Glu Gly Ser Gly Arg Ile Ala Asp Val Ile
320 325 330

Ala Ser Leu Val Glu Val Glu Asp Ala Pro Thr Ser Ser Ala Val
335 340 345

Lys Glu Lys Leu Val Arg Phe Leu Pro Arg Thr Val Ser Arg Leu
350 355 360

Ser Glu Glu Glu Thr Glu Ser Trp Ile Lys Trp Leu Lys Glu Ile
365 370 375

Leu Glu Cys Ser His Leu Leu Thr Val Ile Lys Met Glu Glu Ala
380 385 390

Gly Asp Glu Ile Val Ser Asn Ala Ile Ser Tyr Ala Leu Tyr Lys
395 400 405

Ala Phe Ser Thr Ser Glu Gln Asp Lys Asp Asn Trp Asn Gly Gln
410 415 420

Leu Lys Leu Leu Leu Glu Trp Asn Gln Leu Asp Leu Ala Asn Asp
425 430 435

Glu Ile Phe Thr Asn Asp Arg Arg Trp Glu Ser Ala Asp Leu Gln
440 445 450

Glu Val Met Phe Thr Ala Leu Ile Lys Asp Arg Pro Lys Phe Val
455 460 465

Arg Leu Phe Leu Glu Asn Gly Leu Asn Leu Arg Lys Phe Leu Thr
470 475 480

His Asp Val Leu Thr Glu Leu Phe Ser Asn His Phe Ser Thr Leu
485 490 495

Val Tyr Arg Asn Leu Gln Ile Ala Lys Asn Ser Tyr Asn Asp Ala
500 505 510

Leu Leu Thr Phe Val Trp Lys Leu Val Ala Asn Phe Arg Arg Gly
515 520 525

Phe Arg Lys Glu Asp Arg Asn Gly Arg Asp Glu Met Asp Ile Glu
530 535 540

Leu His Asp Val Ser Pro Ile Thr Arg His Pro Leu Gln Ala Leu
545 550 555

Phe Ile Trp Ala Ile Leu Gln Asn Lys Lys Glu Leu Ser Lys Val
560 565 570

Ile Trp Glu Gln Thr Arg Gly Cys Thr Leu Ala Ala Leu Gly Ala
575 580 585

Ser Lys Leu Leu Lys Thr Leu Ala Lys Val Lys Asn Asp Ile Asn

590 595 600

Ala Ala Gly Glu Ser Glu Glu Leu Ala Asn Glu Tyr Glu Thr Arg
605 610 615

Ala Val Glu Leu Phe Thr Glu Cys Tyr Ser Ser Asp Glu Asp Leu
620 625 630

Ala Glu Gln Leu Leu Val Tyr Ser Cys Glu Ala Trp Gly Gly Ser
635 640 645

Asn Cys Leu Glu Leu Ala Val Glu Ala Thr Asp Gln His Phe Thr
650 655 660

Ala Gln Pro Gly Val Gln Asn Phe Leu Ser Lys Gln Trp Tyr Gly
665 670 675

Glu Ile Ser Arg Asp Thr Lys Asn Trp Lys Ile Ile Leu Cys Leu
680 685 690

Phe Ile Ile Pro Leu Val Gly Cys Gly Phe Val Ser Phe Arg Lys
695 700 705

Lys Pro Val Asp Lys His Lys Lys Leu Leu Trp Tyr Tyr Val Ala
710 715 720

Phe Phe Thr Ser Pro Phe Val Val Phe Ser Trp Asn Val Val Phe
725 730 735

Tyr Ile Ala Phe Leu Leu Leu Phe Ala Tyr Val Leu Leu Met Asp
740 745 750

Phe His Ser Val Pro His Pro Pro Glu Leu Val Leu Tyr Ser Leu
755 760 765

Val Phe Val Leu Phe Cys Asp Glu Val Arg Gln Trp Tyr Val Asn
770 775 780

Gly Val Asn Tyr Phe Thr Asp Leu Trp Asn Val Met Asp Thr Leu
785 790 795

Gly Leu Phe Tyr Phe Ile Ala Gly Ile Val Phe Arg Leu His Ser
800 805 810

Ser Asn Lys Ser Ser Leu Tyr Ser Gly Arg Val Ile Phe Cys Leu
815 820 825

Asp Tyr Ile Ile Phe Thr Leu Arg Leu Ile His Ile Phe Thr Val
830 835 840

Ser Arg Asn Leu Gly Pro Lys Ile Ile Met Leu Gln Arg Met Leu
845 850 855

Ile Asp Val Phe Phe Phe Leu Phe Leu Phe Ala Val Trp Met Val
860 865 870

Ala Phe Gly Val Ala Arg Gln Gly Ile Leu Arg Gln Asn Glu Gln
875 880 885

Arg Trp Arg Trp Ile Phe Arg Ser Val Ile Tyr Glu Pro Tyr Leu
890 895 900

Ala Met Phe Gly Gln Val Pro Ser Asp Val Asp Gly Thr Thr Tyr
905 910 915

Asp Phe Ala His Cys Thr Phe Thr Gly Asn Glu Ser Lys Pro Leu
920 925 930

Cys Val Glu Leu Asp Glu His Asn Leu Pro Arg Phe Pro Glu Trp
935 940 945

Ile Thr Ile Pro Leu Val Cys Ile Tyr Met Leu Ser Thr Asn Ile
950 955 960

Leu Leu Val Asn Leu Leu Val Ala Met Phe Gly Tyr Thr Val Gly
965 970 975

Thr Val Gln Glu Asn Asn Asp Gln Val Trp Lys Phe Gln Arg Tyr
980 985 990

Phe Leu Val Gln Glu Tyr Cys Ser Arg Leu Asn Ile Pro Phe Pro
995 1000 1005

Phe Ile Val Phe Ala Tyr Phe Tyr Met Val Val Lys Lys Cys Phe
1010 1015 1020

Lys Cys Cys Cys Lys Glu Lys Asn Met Glu Ser Ser Val Cys Cys
1025 1030 1035

Phe Lys Asn Glu Asp Asn Glu Thr Leu Ala Trp Glu Gly Val Met
1040 1045 1050

Lys Glu Asn Tyr Leu Val Lys Ile Asn Thr Lys Ala Asn Asp Thr
1055 1060 1065

Ser Glu Glu Met Arg His Arg Phe Arg Gln Leu Asp Thr Lys Leu
1070 1075 1080

Asn Asp Leu Lys Gly Leu Leu Lys Glu Ile Ala Asn Lys Ile Lys
1085 1090 1095

<210> 96
<211> 1095
<212> PRT
<213> Homo sapiens

<400> 96
Met Arg Asn Arg Arg Asn Asp Thr Leu Asp Ser Thr Arg Thr Leu
1 5 10 15

Tyr Ser Ser Ala Ser Arg Ser Thr Asp Leu Ser Tyr Ser Glu Ser
20 25 30

Asp Leu Val Asn Phe Ile Gln Ala Asn Phe Lys Lys Arg Glu Cys
35 40 45

Val Phe Phe Thr Lys Asp Ser Lys Ala Thr Glu Asn Val Cys Lys
50 55 60

Cys Gly Tyr Ala Gln Ser Gln His Met Glu Gly Thr Gln Ile Asn
65 70 75

Gln Ser Glu Lys Trp Asn Tyr Lys Lys His Thr Lys Glu Phe Pro
80 85 90

Thr Asp Ala Phe Gly Asp Ile Gln Phe Glu Thr Leu Gly Lys Lys
95 100 105

Gly Lys Tyr Ile Arg Leu Ser Cys Asp Thr Asp Ala Glu Ile Leu
110 115 120

Tyr Glu Leu Leu Thr Gln His Trp His Leu Lys Thr Pro Asn Leu
125 130 135

Val Ile Ser Val Thr Gly Gly Ala Lys Asn Phe Ala Leu Lys Pro
140 145 150

Arg Met Arg Lys Ile Phe Ser Arg Leu Ile Tyr Ile Ala Gln Ser
155 160 165

Lys Gly Ala Trp Ile Leu Thr Gly Gly Thr His Tyr Gly Leu Met
170 175 180

Lys Tyr Ile Gly Glu Val Val Arg Asp Asn Thr Ile Ser Arg Ser
185 190 195

Ser Glu Glu Asn Ile Val Ala Ile Gly Ile Ala Ala Trp Gly Met
200 205 210

Val Ser Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu Gly
215 220 225

Tyr Phe Leu Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro
230 235 240

Leu Tyr Ile Leu Asp Asn Asn His Thr His Leu Leu Leu Val Asp
245 250 255

Asn Gly Cys His Gly His Pro Thr Val Glu Ala Lys Leu Arg Asn
260 265 270

Gln Leu Glu Lys Tyr Ile Ser Glu Arg Thr Ile Gln Asp Ser Asn
275 280 285

Tyr Gly Gly Lys Ile Pro Ile Val Cys Phe Ala Gln Gly Gly Gly
290 295 300

Lys Glu Thr Leu Lys Ala Ile Asn Thr Ser Ile Lys Asn Lys Ile
305 310 315

Pro Cys Val Val Val Glu Gly Ser Gly Gln Ile Ala Asp Val Ile
320 325 330

Ala Ser Leu Val Glu Val Glu Asp Ala Leu Thr Ser Ser Ala Val
335 340 345

Lys Glu Lys Leu Val Arg Phe Leu Pro Arg Thr Val Ser Arg Leu
350 355 360

Pro Glu Glu Glu Thr Glu Ser Trp Ile Lys Trp Leu Lys Glu Ile
365 370 375

Leu Glu Cys Ser His Leu Leu Thr Val Ile Lys Met Glu Glu Ala
380 385 390

Gly Asp Glu Ile Val Ser Asn Ala Ile Ser Tyr Ala Leu Tyr Lys
395 400 405

Ala Phe Ser Thr Ser Glu Gln Asp Lys Asp Asn Trp Asn Gly Gln
410 415 420

Leu Lys Leu Leu Leu Glu Trp Asn Gln Leu Asp Leu Ala Asn Asp
425 430 435

Glu Ile Phe Thr Asn Asp Arg Arg Trp Glu Ser Ala Asp Leu Gln
440 445 450

Glu Val Met Phe Thr Ala Leu Ile Lys Asp Arg Pro Lys Phe Val
455 460 465

Arg Leu Phe Leu Glu Asn Gly Leu Asn Leu Arg Lys Phe Leu Thr
470 475 480

His Asp Val Leu Thr Glu Leu Phe Ser Asn His Phe Ser Thr Leu
485 490 495

Val Tyr Arg Asn Leu Gln Ile Ala Lys Asn Ser Tyr Asn Asp Ala
500 505 510

Leu Leu Thr Phe Val Trp Lys Leu Val Ala Asn Phe Arg Arg Gly
515 520 525

Phe Arg Lys Glu Asp Arg Asn Gly Arg Asp Glu Met Asp Ile Glu
530 535 540

Leu His Asp Val Ser Pro Ile Thr Arg His Pro Leu Gln Ala Leu
545 550 555

Phe Ile Trp Ala Ile Leu Gln Asn Lys Lys Glu Leu Ser Lys Val
560 565 570

Ile Trp Glu Gln Thr Arg Gly Cys Thr Leu Ala Ala Leu Gly Ala
575 580 585

Ser Lys Leu Leu Lys Thr Leu Ala Lys Val Lys Asn Asp Ile Asn
590 595 600

Ala Ala Gly Glu Ser Glu Glu Leu Ala Asn Glu Tyr Glu Thr Arg

605 610 615

Ala Val Glu Leu Phe Thr Glu Cys Tyr Ser Ser Asp Glu Asp Leu
620 625 630

Ala Glu Gln Leu Leu Val Tyr Ser Cys Glu Ala Trp Gly Gly Ser
635 640 645

Asn Cys Leu Glu Leu Ala Val Glu Ala Thr Asp Gln His Phe Ile
650 655 660

Ala Gln Pro Gly Val Gln Asn Phe Leu Ser Lys Gln Trp Tyr Gly
665 670 675

Glu Ile Ser Arg Asp Thr Lys Asn Trp Lys Ile Ile Leu Cys Leu
680 685 690

Phe Ile Ile Pro Leu Val Gly Cys Gly Phe Val Ser Phe Arg Lys
695 700 705

Lys Pro Val Asp Lys His Lys Lys Leu Leu Trp Tyr Tyr Val Ala
710 715 720

Phe Phe Thr Ser Pro Phe Val Val Phe Ser Trp Asn Val Val Phe
725 730 735

Tyr Ile Ala Phe Leu Leu Leu Phe Ala Tyr Val Leu Leu Met Asp
740 745 750

Phe His Ser Val Pro His Pro Pro Glu Leu Val Leu Tyr Ser Leu
755 760 765

Val Phe Val Leu Phe Cys Asp Glu Val Arg Gln Trp Tyr Val Asn
770 775 780

Gly Val Asn Tyr Phe Thr Asp Leu Trp Asn Val Met Asp Thr Leu
785 790 795

Gly Leu Phe Tyr Phe Ile Ala Gly Ile Val Phe Arg Leu His Ser
800 805 810

Ser Asn Lys Ser Ser Leu Tyr Ser Gly Arg Val Ile Phe Cys Leu
815 820 825

Asp Tyr Ile Ile Phe Thr Leu Arg Leu Ile His Ile Phe Thr Val
830 835 840

Ser Arg Asn Leu Gly Pro Lys Ile Ile Met Leu Gln Arg Met Leu
845 850 855

Ile Asp Val Phe Phe Phe Leu Phe Leu Phe Ala Val Trp Met Val
860 865 870

Ala Phe Gly Val Ala Arg Gln Gly Ile Leu Arg Gln Asn Glu Gln
875 880 885

Arg Trp Arg Trp Ile Phe Arg Ser Val Ile Tyr Glu Pro Tyr Leu
890 895 900

Ala Met Phe Gly Gln Val Pro Ser Asp Val Asp Gly Thr Thr Tyr
905 910 915

Asp Phe Ala His Cys Thr Phe Thr Gly Asn Glu Ser Lys Pro Leu
920 925 930

Cys Val Glu Leu Asp Glu His Asn Leu Pro Arg Phe Pro Glu Trp
935 940 945

Ile Thr Ile Pro Leu Val Cys Ile Tyr Met Leu Ser Thr Asn Ile
950 955 960

Leu Leu Val Asn Leu Leu Val Ala Met Phe Gly Tyr Thr Val Gly
965 970 975

Thr Val Gln Glu Asn Asn Asp Gln Val Trp Lys Phe Gln Arg Tyr
980 985 990

Phe Leu Val Gln Glu Tyr Cys Ser Arg Leu Asn Ile Pro Phe Pro
995 1000 1005

Phe Ile Val Phe Ala Tyr Phe Tyr Met Val Val Lys Lys Cys Phe
1010 1015 1020

Lys Cys Cys Cys Lys Glu Lys Asn Met Glu Ser Ser Val Cys Cys
1025 1030 1035

Phe Lys Asn Glu Asp Asn Glu Thr Leu Ala Trp Glu Gly Val Met
1040 1045 1050

Lys Glu Asn Tyr Leu Val Lys Ile Asn Thr Lys Ala Asn Asp Thr
1055 1060 1065

Ser Glu Glu Met Arg His Arg Phe Arg Gln Leu Asp Thr Lys Leu
1070 1075 1080

Asn Asp Leu Lys Gly Leu Leu Lys Glu Ile Ala Asn Lys Ile Lys
1085 1090 1095

<210> 97
<211> 347
<212> PRT
<213> Homo sapiens

<400> 97
Arg Arg Gly Arg Gly Glu Val Leu Ala Val Glu Arg Gly Ser Gly
1 5 10 15

Ser Gly Gly Gly Gly Thr Arg Ser Gly Trp Pro Ala Pro Ala Ala
20 25 30

Gly Ala Asp Lys Lys Pro Ser Arg Cys Gly Ser Gly Arg Glu Gly
35 40 45

Glu Gly Val Ser Glu Gly His Lys Ser Met Thr Gly Leu Tyr Glu
50 55 60

Leu Val Trp Arg Val Leu His Ala Leu Leu Cys Leu His Arg Thr
65 70 75

Leu Thr Ser Trp Leu Arg Val Arg Phe Gly Thr Trp Asn Trp Ile
80 85 90

Trp Arg Arg Cys Cys Arg Ala Ala Ser Ala Ala Val Leu Ala Pro
95 100 105

Leu Gly Phe Thr Leu Arg Lys Pro Pro Ala Val Gly Arg Asn Arg
110 115 120

Arg His His Arg His Pro Arg Gly Gly Ser Cys Leu Ala Ala Ala
125 130 135

His His Arg Met Arg Trp Arg Ala Asp Gly Arg Ser Leu Glu Lys
140 145 150

Leu Pro Val His Met Gly Leu Val Ile Thr Glu Val Glu Gln Glu
155 160 165

Pro Ser Phe Ser Asp Ile Ala Ser Leu Val Val Trp Cys Met Ala
170 175 180

Val Gly Ile Ser Tyr Ile Ser Val Tyr Asp His Gln Gly Ile Phe
185 190 195

Lys Arg Asn Asn Ser Arg Leu Met Asp Glu Ile Leu Lys Gln Gln
200 205 210

Gln Glu Leu Leu Gly Leu Asp Cys Ser Lys Tyr Ser Pro Glu Phe
215 220 225

Ala Asn Ser Asn Asp Lys Asp Asp Gln Val Leu Asn Cys His Leu
230 235 240

Ala Val Lys Val Leu Ser Pro Glu Asp Gly Lys Ala Asp Ile Val
245 250 255

Arg Ala Ala Gln Asp Phe Cys Gln Leu Val Ala Gln Lys Gln Arg
260 265 270

Arg Pro Thr Asp Leu Asp Val Asp Thr Leu Ala Ser Leu Leu Ser
275 280 285

Ser Asn Gly Cys Pro Asp Pro Asp Leu Val Leu Lys Phe Gly Pro
290 295 300

Val Asp Ser Thr Leu Gly Phe Leu Pro Trp His Ile Arg Leu Thr
305 310 315

Glu Ile Val Ser Leu Pro Ser His Leu Asn Ile Ser Tyr Glu Asp
320 325 330

Phe Phe Ser Ala Leu Arg Gln Tyr Ala Ala Cys Glu Gln Arg Leu
335 340 345

Gly Lys

<210> 98
<211> 293
<212> PRT
<213> Homo sapiens

<400> 98
Met Thr Gly Leu Tyr Glu Leu Val Trp Arg Val Leu His Ala Leu
1 5 10 15

Leu Cys Leu His Arg Thr Leu Thr Ser Trp Leu Arg Val Arg Phe
20 25 30

Gly Thr Trp Asn Trp Ile Trp Arg Arg Cys Cys Arg Ala Ala Ser
35 40 45

Ala Ala Val Leu Ala Pro Leu Gly Phe Thr Leu Arg Lys Pro Pro
50 55 60

Ala Val Gly Arg Asn Arg Arg His His Arg His Pro Arg Gly Gly
65 70 75

Ser Cys Leu Ala Ala Ala His His Arg Met Arg Trp Arg Ala Asp
80 85 90

Gly Arg Ser Leu Glu Lys Leu Pro Val His Met Gly Leu Val Ile
95 100 105

Thr Glu Val Glu Gln Glu Pro Ser Phe Ser Asp Ile Ala Ser Leu
110 115 120

Val Val Trp Cys Met Ala Val Gly Ile Ser Tyr Ile Ser Val Tyr
125 130 135

Asp His Gln Gly Ile Phe Lys Arg Asn Asn Ser Arg Leu Met Asp
140 145 150

Glu Ile Leu Lys Gln Gln Gln Glu Leu Leu Gly Leu Asp Cys Ser
155 160 165

Lys Tyr Ser Pro Glu Phe Ala Asn Ser Asn Asp Lys Asp Asp Gln
170 175 180

Val Leu Asn Cys His Leu Ala Val Lys Val Leu Ser Pro Glu Asp
185 190 195

Gly Lys Ala Asp Ile Val Arg Ala Ala Gln Asp Phe Cys Gln Leu
200 205 210

Val Ala Gln Lys Gln Lys Arg Pro Thr Asp Leu Asp Val Asp Thr
215 220 225

Leu Ala Ser Leu Leu Ser Ser Asn Gly Cys Pro Asp Pro Asp Leu

230 235 240

Val Leu Lys Phe Gly Pro Val Asp Ser Thr Leu Gly Phe Leu Pro
245 250 255

Trp His Ile Arg Leu Thr Glu Ile Val Ser Leu Pro Ser His Leu
260 265 270

Asn Ile Ser Tyr Glu Asp Phe Phe Ser Ala Leu Arg Gln Tyr Ala
275 280 285

Ala Cys Glu Gln Arg Leu Gly Lys
290

<210> 99
<211> 553
<212> PRT
<213> Homo sapiens

<400> 99
Met Val Gln Arg Leu Trp Val Ser Arg Leu Leu Arg His Arg Lys
1 5 10 15

Ala Gln Leu Leu Leu Val Asn Leu Leu Thr Phe Gly Leu Glu Val
20 25 30

Cys Leu Ala Ala Gly Ile Thr Tyr Val Pro Pro Leu Leu Leu Glu
35 40 45

Val Gly Val Glu Glu Lys Phe Met Thr Met Val Leu Gly Ile Gly
50 55 60

Pro Val Leu Gly Leu Val Cys Val Pro Leu Leu Gly Ser Ala Ser
65 70 75

Asp His Trp Arg Gly Arg Tyr Gly Arg Arg Arg Pro Phe Ile Trp
80 85 90

Ala Leu Ser Leu Gly Ile Leu Leu Ser Leu Phe Leu Ile Pro Arg
95 100 105

Ala Gly Trp Leu Ala Gly Leu Leu Cys Pro Asp Pro Arg Pro Leu
110 115 120

Glu Leu Ala Leu Leu Ile Leu Gly Val Gly Leu Leu Asp Phe Cys
125 130 135

Gly Gln Val Cys Phe Thr Pro Leu Glu Ala Leu Leu Ser Asp Leu
140 145 150

Phe Arg Asp Pro Asp His Cys Arg Gln Ala Tyr Ser Val Tyr Ala
155 160 165

Phe Met Ile Ser Leu Gly Gly Cys Leu Gly Tyr Leu Leu Pro Ala
170 175 180

Ile Asp Trp Asp Thr Ser Ala Leu Ala Pro Tyr Leu Gly Thr Gln
185 190 195

Glu Glu Cys Leu Phe Gly Leu Leu Thr Leu Ile Phe Leu Thr Cys
200 205 210

Val Ala Ala Thr Leu Leu Val Ala Glu Glu Ala Ala Leu Gly Pro
215 220 225

Thr Glu Pro Ala Glu Gly Leu Ser Ala Pro Ser Leu Ser Pro His
230 235 240

Cys Cys Pro Cys Arg Ala Arg Leu Ala Phe Arg Asn Leu Gly Ala
245 250 255

Leu Leu Pro Arg Leu His Gln Leu Cys Cys Arg Met Pro Arg Thr
260 265 270

Leu Arg Arg Leu Phe Val Ala Glu Leu Cys Ser Trp Met Ala Leu
275 280 285

Met Thr Phe Thr Leu Phe Tyr Thr Asp Phe Val Gly Glu Gly Leu
290 295 300

Tyr Gln Gly Val Pro Arg Ala Glu Pro Gly Thr Glu Ala Arg Arg
305 310 315

His Tyr Asp Glu Gly Val Arg Met Gly Ser Leu Gly Leu Phe Leu
320 325 330

Gln Cys Ala Ile Ser Leu Val Phe Ser Leu Val Met Asp Arg Leu
335 340 345

Val Gln Arg Phe Gly Thr Arg Ala Val Tyr Leu Ala Ser Val Ala
350 355 360

Ala Phe Pro Val Ala Ala Gly Ala Thr Cys Leu Ser His Ser Val
365 370 375

Ala Val Val Thr Ala Ser Ala Ala Leu Thr Gly Phe Thr Phe Ser
380 385 390

Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala Ser Leu Tyr His Arg
395 400 405

Glu Lys Gln Val Phe Leu Pro Lys Tyr Arg Gly Asp Thr Gly Gly
410 415 420

Ala Ser Ser Glu Asp Ser Leu Met Thr Ser Phe Leu Pro Gly Pro
425 430 435

Lys Pro Gly Ala Pro Phe Pro Asn Gly His Val Gly Ala Gly Gly
440 445 450

Ser Gly Leu Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser Ala
455 460 465

Cys Asp Val Ser Val Arg Val Val Val Gly Glu Pro Thr Glu Ala
470 475 480

Arg Val Val Pro Gly Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu
485 490 495

Asp Ser Ala Phe Leu Leu Ser Gln Val Ala Pro Ser Leu Phe Met
500 505 510

Gly Ser Ile Val Gln Leu Ser Gln Ser Val Thr Ala Tyr Met Val
515 520 525

Ser Ala Ala Gly Leu Gly Leu Val Ala Ile Tyr Phe Ala Thr Gln
530 535 540

Val Val Phe Asp Lys Ser Asp Leu Ala Lys Tyr Ser Ala
545 550

<210> 100
<211> 987
<212> PRT
<213> Homo sapiens

<400> 100
Met Ala Leu Arg Arg Leu Gly Ala Ala Leu Leu Leu Leu Pro Leu
1 5 10 15

Leu Ala Ala Val Glu Glu Thr Leu Met Asp Ser Thr Thr Ala Thr
20 25 30

Ala Glu Leu Gly Trp Met Val His Pro Pro Ser Gly Trp Glu Glu
35 40 45

Val Ser Gly Tyr Asp Glu Asn Met Asn Thr Ile Arg Thr Tyr Gln
50 55 60

Val Cys Asn Val Phe Glu Ser Ser Gln Asn Asn Trp Leu Arg Thr
65 70 75

Lys Phe Ile Arg Arg Arg Gly Ala His Arg Ile His Val Glu Met
80 85 90

Lys Phe Ser Val Arg Asp Cys Ser Ser Ile Pro Ser Val Pro Gly
95 100 105

Ser Cys Lys Glu Thr Phe Asn Leu Tyr Tyr Tyr Glu Ala Asp Phe
110 115 120

Asp Ser Ala Thr Lys Thr Phe Pro Asn Trp Met Glu Asn Pro Trp
125 130 135

Val Lys Val Asp Thr Ile Ala Ala Asp Glu Ser Phe Ser Gln Val
140 145 150

Asp Leu Gly Gly Arg Val Met Lys Ile Asn Thr Glu Val Arg Ser
155 160 165

Phe Gly Pro Val Ser Arg Ser Gly Phe Tyr Leu Ala Phe Gln Asp
170 175 180

Tyr Gly Gly Cys Met Ser Leu Ile Ala Val Arg Val Phe Tyr Arg
185 190 195

Lys Cys Pro Arg Ile Ile Gln Asn Gly Ala Ile Phe Gln Glu Thr
200 205 210

Leu Ser Gly Ala Glu Ser Thr Ser Leu Val Ala Ala Arg Gly Ser
215 220 225

Cys Ile Ala Asn Ala Glu Glu Val Asp Val Pro Ile Lys Leu Tyr
230 235 240

Cys Asn Gly Asp Gly Glu Trp Leu Val Pro Ile Gly Arg Cys Met
245 250 255

Cys Lys Ala Gly Phe Glu Ala Val Glu Asn Gly Thr Val Cys Arg
260 265 270

Gly Cys Pro Ser Gly Thr Phe Lys Ala Asn Gln Gly Asp Glu Ala
275 280 285

Cys Thr His Cys Pro Ile Asn Ser Arg Thr Thr Ser Glu Gly Ala
290 295 300

Thr Asn Cys Val Cys Arg Asn Gly Tyr Tyr Arg Ala Asp Leu Asp
305 310 315

Pro Leu Asp Met Pro Cys Thr Thr Ile Pro Ser Ala Pro Gln Ala
320 325 330

Val Ile Ser Ser Val Asn Glu Thr Ser Leu Met Leu Glu Trp Thr
335 340 345

Pro Pro Arg Asp Ser Gly Gly Arg Glu Asp Leu Val Tyr Asn Ile
350 355 360

Ile Cys Lys Ser Cys Gly Ser Gly Arg Gly Ala Cys Thr Arg Cys
365 370 375

Gly Asp Asn Val Gln Tyr Ala Pro Arg Gln Leu Gly Leu Thr Glu
380 385 390

Pro Arg Ile Tyr Ile Ser Asp Leu Leu Ala His Thr Gln Tyr Thr
395 400 405

Phe Glu Ile Gln Ala Val Asn Gly Val Thr Asp Gln Ser Pro Phe
410 415 420

Ser Pro Gln Phe Ala Ser Val Asn Ile Thr Thr Asn Gln Ala Ala
425 430 435

Pro Ser Ala Val Ser Ile Met His Gln Val Ser Arg Thr Val Asp
440 445 450

Ser Ile Thr Leu Ser Trp Ser Gln Pro Asp Gln Pro Asn Gly Val

455 460 465

Ile Leu Asp Tyr Glu Leu Gln Tyr Tyr Glu Lys Glu Leu Ser Glu
470 475 480

Tyr Asn Ala Thr Ala Ile Lys Ser Pro Thr Asn Thr Val Thr Val
485 490 495

Gln Gly Leu Lys Ala Gly Ala Ile Tyr Val Phe Gln Val Arg Ala
500 505 510

Arg Thr Val Ala Gly Tyr Gly Arg Tyr Ser Gly Lys Met Tyr Phe
515 520 525

Gln Thr Met Thr Glu Ala Glu Tyr Gln Thr Ser Ile Gln Glu Lys
530 535 540

Leu Pro Leu Ile Ile Gly Ser Ser Ala Ala Gly Leu Val Phe Leu
545 550 555

Ile Ala Val Val Val Ile Ala Ile Val Cys Asn Arg Arg Arg Gly
560 565 570

Phe Glu Arg Ala Asp Ser Glu Tyr Thr Asp Lys Leu Gln His Tyr
575 580 585

Thr Ser Gly His Met Thr Pro Gly Met Lys Ile Tyr Ile Asp Pro
590 595 600

Phe Thr Tyr Glu Asp Pro Asn Glu Ala Val Arg Glu Phe Ala Lys
605 610 615

Glu Ile Asp Ile Ser Cys Val Lys Ile Glu Gln Val Ile Gly Ala
620 625 630

Gly Glu Phe Gly Glu Val Cys Ser Gly His Leu Lys Leu Pro Gly
635 640 645

Lys Arg Glu Ile Phe Val Ala Ile Lys Thr Leu Lys Ser Gly Tyr
650 655 660

Thr Glu Lys Gln Arg Arg Asp Phe Leu Ser Glu Ala Ser Ile Met
665 670 675

Gly Gln Phe Asp His Pro Asn Val Ile His Leu Glu Gly Val Val
680 685 690

Thr Lys Ser Thr Pro Val Met Ile Ile Thr Glu Phe Met Glu Asn
695 700 705

Gly Ser Leu Asp Ser Phe Leu Arg Gln Asn Asp Gly Gln Phe Thr
710 715 720

Val Ile Gln Leu Val Gly Met Leu Arg Gly Ile Ala Ala Gly Met
725 730 735

Lys Tyr Leu Ala Asp Met Asn Tyr Val His Arg Asp Leu Ala Ala
740 745 750

Arg Asn Ile Leu Val Asn Ser Asn Leu Val Cys Lys Val Ser Asp
755 760 765

Phe Gly Leu Ser Arg Phe Leu Glu Asp Asp Thr Ser Asp Pro Thr
770 775 780

Tyr Thr Ser Ala Leu Gly Gly Lys Ile Pro Ile Arg Trp Thr Ala
785 790 795

Pro Glu Ala Ile Gln Tyr Arg Lys Phe Thr Ser Ala Ser Asp Val
800 805 810

Trp Ser Tyr Gly Ile Val Met Trp Glu Val Met Ser Tyr Gly Glu
815 820 825

Arg Pro Tyr Trp Asp Met Thr Asn Gln Asp Val Ile Asn Ala Ile
830 835 840

Glu Gln Asp Tyr Arg Leu Pro Pro Pro Met Asp Cys Pro Ser Ala
845 850 855

Leu His Gln Leu Met Leu Asp Cys Trp Gln Lys Asp Arg Asn His
860 865 870

Arg Pro Lys Phe Gly Gln Ile Val Asn Thr Leu Asp Lys Met Ile
875 880 885

Arg Asn Pro Asn Ser Leu Lys Ala Met Ala Pro Leu Ser Ser Gly
890 895 900

Ile Asn Leu Pro Leu Leu Asp Arg Thr Ile Pro Asp Tyr Thr Ser
905 910 915

Phe Asn Thr Val Asp Glu Trp Leu Glu Ala Ile Lys Met Gly Gln
920 925 930

Tyr Lys Glu Ser Phe Ala Asn Ala Gly Phe Thr Ser Phe Asp Val
935 940 945

Val Ser Gln Met Met Met Glu Asp Ile Leu Arg Val Gly Leu Thr
950 955 960

Leu Ala Gly His Gln Lys Lys Ile Leu Asn Ser Ile Gln Val Met
965 970 975

Arg Ala Gln Met Asn Gln Ile Gln Ser Val Glu Val
980 985

<210> 101
<211> 986
<212> PRT
<213> Homo sapiens

<400> 101
Met Ala Leu Arg Arg Leu Gly Ala Ala Leu Leu Leu Leu Pro Leu
1 5 10 15

Leu Ala Ala Val Glu Glu Thr Leu Met Asp Ser Thr Thr Ala Thr
20 25 30

Ala Glu Leu Gly Trp Met Val His Pro Pro Ser Gly Trp Glu Glu
35 40 45

Val Ser Gly Tyr Asp Glu Asn Met Asn Thr Ile Arg Thr Tyr Gln
50 55 60

Val Cys Asn Val Phe Glu Ser Ser Gln Asn Asn Trp Leu Arg Thr
65 70 75

Lys Phe Ile Arg Arg Arg Gly Ala His Arg Ile His Val Glu Met
80 85 90

Lys Phe Ser Val Arg Asp Cys Ser Ser Ile Pro Ser Val Pro Gly
95 100 105

Ser Cys Lys Glu Thr Phe Asn Leu Tyr Tyr Tyr Glu Ala Asp Phe
110 115 120

Asp Ser Ala Thr Lys Thr Phe Pro Asn Trp Met Glu Asn Pro Trp
125 130 135

Val Lys Val Asp Thr Ile Ala Ala Asp Glu Ser Phe Ser Gln Val
140 145 150

Asp Leu Gly Gly Arg Val Met Lys Ile Asn Thr Glu Val Arg Ser
155 160 165

Phe Gly Pro Val Ser Arg Ser Gly Phe Tyr Leu Ala Phe Gln Asp
170 175 180

Tyr Gly Gly Cys Met Ser Leu Ile Ala Val Arg Val Phe Tyr Arg
185 190 195

Lys Cys Pro Arg Ile Ile Gln Asn Gly Ala Ile Phe Gln Glu Thr
200 205 210

Leu Ser Gly Ala Glu Ser Thr Ser Leu Val Ala Ala Arg Gly Ser
215 220 225

Cys Ile Ala Asn Ala Glu Glu Val Asp Val Pro Ile Lys Leu Tyr
230 235 240

Cys Asn Gly Asp Gly Glu Trp Leu Val Pro Ile Gly Arg Cys Met
245 250 255

Cys Lys Ala Gly Phe Glu Ala Val Glu Asn Gly Thr Val Cys Arg
260 265 270

Gly Cys Pro Ser Gly Thr Phe Lys Ala Asn Gln Gly Asp Glu Ala
275 280 285

Cys Thr His Cys Pro Ile Asn Ser Arg Thr Thr Ser Glu Gly Ala
290 295 300

Thr Asn Cys Val Cys Arg Asn Gly Tyr Tyr Arg Ala Asp Leu Asp
305 310 315

Pro Leu Asp Met Pro Cys Thr Thr Ile Pro Ser Ala Pro Gln Ala
320 325 330

Val Ile Ser Ser Val Asn Glu Thr Ser Leu Met Leu Glu Trp Thr
335 340 345

Pro Pro Arg Asp Ser Gly Gly Arg Glu Asp Leu Val Tyr Asn Ile
350 355 360

Ile Cys Lys Ser Cys Gly Ser Gly Arg Gly Ala Cys Thr Arg Cys
365 370 375

Gly Asp Asn Val Gln Tyr Ala Pro Arg Gln Leu Gly Leu Thr Glu
380 385 390

Pro Arg Ile Tyr Ile Ser Asp Leu Leu Ala His Thr Gln Tyr Thr
395 400 405

Phe Glu Ile Gln Ala Val Asn Gly Val Thr Asp Gln Ser Pro Phe
410 415 420

Ser Pro Gln Phe Ala Ser Val Asn Ile Thr Thr Asn Gln Ala Ala
425 430 435

Pro Ser Ala Val Ser Ile Met His Gln Val Ser Arg Thr Val Asp
440 445 450

Ser Ile Thr Leu Ser Trp Ser Gln Pro Asp Gln Pro Asn Gly Val
455 460 465

Ile Leu Asp Tyr Glu Leu Gln Tyr Tyr Glu Lys Glu Leu Ser Glu
470 475 480

Tyr Asn Ala Thr Ala Ile Lys Ser Pro Thr Asn Thr Val Thr Val
485 490 495

Gln Gly Leu Lys Ala Gly Ala Ile Tyr Val Phe Gln Val Arg Ala
500 505 510

Arg Thr Val Ala Gly Tyr Gly Arg Tyr Ser Gly Lys Met Tyr Phe
515 520 525

Gln Thr Met Thr Glu Ala Glu Tyr Gln Thr Ser Ile Gln Glu Lys
530 535 540

Leu Pro Leu Ile Ile Gly Ser Ser Ala Ala Gly Leu Val Phe Leu
545 550 555

Ile Ala Val Val Val Ile Ala Ile Val Cys Asn Arg Arg Gly Phe
560 565 570

Glu Arg Ala Asp Ser Glu Tyr Thr Asp Lys Leu Gln His Tyr Thr

575 580 585

Ser Gly His Met Thr Pro Gly Met Lys Ile Tyr Ile Asp Pro Phe
590 595 600

Thr Tyr Glu Asp Pro Asn Glu Ala Val Arg Glu Phe Ala Lys Glu
605 610 615

Ile Asp Ile Ser Cys Val Lys Ile Glu Gln Val Ile Gly Ala Gly
620 625 630

Glu Phe Gly Glu Val Cys Ser Gly His Leu Lys Leu Pro Gly Lys
635 640 645

Arg Glu Ile Phe Val Ala Ile Lys Thr Leu Lys Ser Gly Tyr Thr
650 655 660

Glu Lys Gln Arg Arg Asp Phe Leu Ser Glu Ala Ser Ile Met Gly
665 670 675

Gln Phe Asp His Pro Asn Val Ile His Leu Glu Gly Val Val Thr
680 685 690

Lys Ser Thr Pro Val Met Ile Ile Thr Glu Phe Met Glu Asn Gly
695 700 705

Ser Leu Asp Ser Phe Leu Arg Gln Asn Asp Gly Gln Phe Thr Val
710 715 720

Ile Gln Leu Val Gly Met Leu Arg Gly Ile Ala Ala Gly Met Lys
725 730 735

Tyr Leu Ala Asp Met Asn Tyr Val His Arg Asp Leu Ala Ala Arg
740 745 750

Asn Ile Leu Val Asn Ser Asn Leu Val Cys Lys Val Ser Asp Phe
755 760 765

Gly Leu Ser Arg Phe Leu Glu Asp Asp Thr Ser Asp Pro Thr Tyr
770 775 780

Thr Ser Ala Leu Gly Gly Lys Ile Pro Ile Arg Trp Thr Ala Pro
785 790 795

Glu Ala Ile Gln Tyr Arg Lys Phe Thr Ser Ala Ser Asp Val Trp
800 805 810

Ser Tyr Gly Ile Val Met Trp Glu Val Met Ser Tyr Gly Glu Arg
815 820 825

Pro Tyr Trp Asp Met Thr Asn Gln Asp Val Ile Asn Ala Ile Glu
830 835 840

Gln Asp Tyr Arg Leu Pro Pro Pro Met Asp Cys Pro Ser Ala Leu
845 850 855

His Gln Leu Met Leu Asp Cys Trp Gln Lys Asp Arg Asn His Arg
860 865 870

Pro Lys Phe Gly Gln Ile Val Asn Thr Leu Asp Lys Met Ile Arg
875 880 885

Asn Pro Asn Ser Leu Lys Ala Met Ala Pro Leu Ser Ser Gly Ile
890 895 900

Asn Leu Pro Leu Leu Asp Arg Thr Ile Pro Asp Tyr Thr Ser Phe
905 910 915

Asn Thr Val Asp Glu Trp Leu Glu Ala Ile Lys Met Gly Gln Tyr
920 925 930

Lys Glu Ser Phe Ala Asn Ala Gly Phe Thr Ser Phe Asp Val Val
935 940 945

Ser Gln Met Met Met Glu Asp Ile Leu Arg Val Gly Val Thr Leu
950 955 960

1483

Ala Gly His Gln Lys Lys Ile Leu Asn Ser Ile Gln Val Met Arg
965 970 975

Ala Gln Met Asn Gln Ile Gln Ser Val Glu Val
980 985

<210> 102
<211> 153
<212> PRT
<213> Homo sapiens

<400> 102
Met Asn Phe Ser Thr Ser Ser Ser Ser Phe Ala Tyr Asp Arg Glu
1 5 10 15

Phe Leu Arg Thr Leu Pro Gly Phe Leu Ile Val Ala Glu Ile Val
20 25 30

Leu Gly Leu Leu Val Trp Thr Leu Ile Ala Gly Thr Glu Tyr Phe
35 40 45

Arg Val Pro Ala Phe Gly Trp Val Met Phe Val Ala Val Phe Tyr
50 55 60

Trp Val Leu Thr Val Phe Phe Leu Ile Ile Tyr Ile Thr Met Thr
65 70 75

Tyr Thr Arg Ile Pro Gln Val Pro Trp Thr Thr Val Gly Leu Cys
80 85 90

Phe Asn Gly Ser Ala Phe Val Leu Tyr Leu Ser Ala Ala Val Val
95 100 105

Asp Ala Ser Ser Val Ser Pro Glu Arg Asp Ser His Asn Phe Asn
110 115 120

Ser Trp Ala Ala Ser Ser Phe Phe Ala Phe Leu Val Asn Ile Cys
125 130 135

Tyr Ala Gly Asn Thr Tyr Phe Ser Phe Ile Ala Trp Arg Ser Arg
140 145 150

Thr Ile Gln

<210> 103
<211> 153
<212> PRT
<213> Homo sapiens

<400> 103
Met Asn Phe Ser Thr Ser Ser Ser Ser Phe Ala Tyr Asp Arg Glu
1 5 10 15

Phe Leu Arg Thr Leu Pro Gly Phe Leu Ile Val Ala Glu Ile Val
20 25 30

Leu Gly Leu Leu Val Trp Thr Leu Ile Ala Gly Thr Glu Tyr Phe
35 40 45

Arg Val Pro Ala Phe Gly Trp Val Met Phe Val Ala Val Phe Tyr
50 55 60

Trp Val Leu Thr Val Phe Phe Leu Ile Ile Tyr Ile Thr Met Thr
65 70 75

Tyr Thr Arg Ile Pro Gln Val Pro Trp Thr Thr Val Gly Leu Cys
80 85 90

Phe Asn Gly Ser Ala Phe Val Leu Tyr Leu Ser Ala Ala Val Val
95 100 105

Asp Ala Ser Ser Val Ser Pro Glu Arg Asp Ser His Asn Phe Asn
110 115 120

Ser Trp Ala Ala Ser Ser Phe Phe Ala Phe Leu Val Thr Ile Cys
125 130 135

Tyr Ala Gly Asn Thr Tyr Phe Ser Phe Ile Ala Trp Arg Ser Arg
140 145 150

Thr Ile Gln

<210> 104
<211> 555
<212> PRT
<213> Homo sapiens

<400> 104
Met Pro Thr Thr Val Asp Asp Val Leu Glu His Gly Gly Glu Phe
1 5 10 15

His Phe Phe Gln Lys Gln Met Phe Phe Leu Leu Ala Leu Leu Ser
20 25 30

Ala Thr Phe Ala Pro Ile Tyr Val Gly Ile Val Phe Leu Gly Phe
35 40 45

Thr Pro Asp His Arg Cys Arg Ser Pro Gly Val Ala Glu Leu Ser
50 55 60

Leu Arg Cys Gly Trp Ser Pro Ala Glu Glu Leu Asn Tyr Thr Val
65 70 75

Pro Gly Pro Gly Pro Ala Gly Glu Ala Ser Pro Arg Gln Cys Arg
80 85 90

Arg Tyr Glu Val Asp Trp Asn Gln Ser Thr Phe Asp Cys Val Asp
95 100 105

Pro Leu Ala Ser Leu Asp Thr Asn Arg Ser Arg Leu Pro Leu Gly
110 115 120

Pro Cys Arg Asp Gly Trp Val Tyr Glu Thr Pro Gly Ser Ser Ile
125 130 135

Val Thr Glu Phe Asn Leu Val Cys Ala Asn Ser Trp Met Leu Asp
140 145 150

Leu Phe Gln Ser Ser Val Asn Val Gly Phe Phe Ile Gly Ser Met
155 160 165

Ser Ile Gly Tyr Ile Ala Asp Arg Phe Gly Arg Lys Leu Cys Leu
170 175 180

Leu Thr Thr Val Leu Ile Asn Ala Ala Ala Gly Val Leu Met Ala
185 190 195

Ile Ser Pro Thr Tyr Thr Trp Met Leu Ile Phe Arg Leu Ile Gln
200 205 210

Gly Leu Val Ser Lys Ala Gly Trp Leu Ile Gly Tyr Ile Leu Ile
215 220 225

Thr Glu Phe Val Gly Arg Arg Tyr Arg Arg Thr Val Gly Ile Phe
230 235 240

Tyr Gln Val Ala Tyr Thr Val Gly Leu Leu Val Leu Ala Gly Val
245 250 255

Ala Tyr Ala Leu Pro His Trp Arg Trp Leu Gln Phe Thr Val Ala
260 265 270

Leu Pro Asn Phe Phe Phe Leu Leu Tyr Tyr Trp Cys Ile Pro Glu
275 280 285

Ser Pro Arg Trp Leu Ile Ser Gln Asn Lys Asn Ala Glu Ala Met
290 295 300

Arg Ile Ile Lys His Ile Ala Lys Lys Asn Gly Lys Ser Leu Pro

305 310 315

Ala Ser Leu Gln Arg Leu Arg Leu Glu Glu Glu Thr Gly Lys Lys
320 325 330

Leu Asn Pro Ser Phe Leu Asp Leu Val Arg Thr Pro Gln Ile Arg
335 340 345

Lys His Thr Met Ile Leu Met Tyr Asn Trp Phe Thr Ser Ser Val
350 355 360

Leu Tyr Gln Gly Leu Ile Met His Met Gly Leu Ala Gly Asp Asn
365 370 375

Ile Tyr Leu Asp Phe Phe Tyr Ser Ala Leu Val Glu Phe Pro Ala
380 385 390

Ala Phe Met Ile Ile Leu Thr Ile Asp Arg Ile Gly Arg Arg Tyr
395 400 405

Pro Trp Ala Ala Ser Asn Met Val Ala Gly Ala Ala Cys Leu Ala
410 415 420

Ser Val Phe Ile Pro Gly Asp Leu Gln Trp Leu Lys Ile Ile Ile
425 430 435

Ser Cys Leu Gly Arg Met Gly Ile Thr Met Ala Tyr Glu Ile Val
440 445 450

Cys Leu Val Asn Ala Glu Leu Tyr Pro Thr Phe Ile Arg Asn Leu
455 460 465

Gly Val His Ile Cys Ser Ser Met Cys Asp Ile Gly Gly Ile Ile
470 475 480

Thr Pro Phe Leu Val Tyr Arg Leu Thr Asn Ile Trp Leu Glu Leu
485 490 495

Pro Leu Met Val Phe Gly Val Leu Gly Leu Val Ala Gly Gly Leu
500 505 510

Val Leu Leu Leu Pro Glu Thr Lys Gly Lys Ala Leu Pro Glu Thr
515 520 525

Ile Glu Glu Ala Glu Asn Met Gln Arg Pro Arg Lys Asn Lys Glu
530 535 540

Lys Met Ile Tyr Leu Gln Val Gln Lys Leu Asp Ile Pro Leu Asn
545 550 555

<210> 105
<211> 555
<212> PRT
<213> Homo sapiens

<400> 105
Met Pro Thr Thr Val Asp Asp Val Leu Glu His Gly Gly Glu Phe
1 5 10 15

His Phe Phe Gln Lys Gln Met Phe Phe Leu Leu Ala Leu Leu Ser
20 25 30

Ala Thr Phe Ala Pro Ile Tyr Val Gly Ile Val Phe Leu Gly Phe
35 40 45

Thr Pro Asp His Arg Cys Arg Ser Pro Gly Val Ala Glu Leu Ser
50 55 60

Leu Arg Cys Gly Trp Ser Pro Ala Glu Glu Leu Asn Tyr Thr Val
65 70 75

Pro Gly Pro Gly Pro Ala Gly Glu Ala Ser Pro Arg Gln Cys Arg
80 85 90

Arg Tyr Glu Val Asp Trp Asn Gln Ser Thr Phe Asp Cys Val Asp
95 100 105

Pro Leu Ala Ser Leu Asp Thr Asn Arg Ser Arg Leu Pro Leu Gly
110 115 120

Pro Cys Arg Asp Gly Trp Val Tyr Glu Thr Pro Gly Ser Ser Ile
125 130 135

Val Thr Glu Phe Asn Leu Val Cys Ala Asn Ser Trp Met Leu Asp
140 145 150

Leu Phe Gln Ser Ser Val Asn Val Gly Phe Phe Ile Gly Ser Met
155 160 165

Ser Ile Gly Tyr Ile Ala Asp Arg Phe Gly Arg Lys Leu Cys Leu
170 175 180

Leu Thr Thr Val Leu Ile Asn Ala Ala Ala Gly Val Leu Met Ala
185 190 195

Ile Ser Pro Thr Tyr Thr Trp Met Leu Ile Phe Arg Leu Ile Gln
200 205 210

Gly Leu Val Ser Lys Ala Gly Trp Leu Ile Gly Tyr Ile Leu Ile
215 220 225

Thr Glu Phe Val Gly Arg Arg Tyr Arg Arg Thr Val Gly Ile Phe
230 235 240

Tyr Gln Val Ala Tyr Thr Val Gly Leu Leu Val Leu Ala Gly Val
245 250 255

Ala Tyr Ala Leu Pro His Trp Arg Trp Leu Gln Phe Thr Val Ser
260 265 270

Leu Pro Asn Phe Phe Phe Leu Leu Tyr Tyr Trp Cys Ile Pro Glu
275 280 285

Ser Pro Arg Trp Leu Ile Ser Gln Asn Lys Asn Ala Glu Ala Met
290 295 300

Arg Ile Ile Lys His Ile Ala Lys Lys Asn Gly Lys Ser Leu Pro
305 310 315

Ala Ser Leu Gln Arg Leu Arg Leu Glu Glu Glu Thr Gly Lys Lys
320 325 330

Leu Asn Pro Ser Phe Leu Asp Leu Val Arg Thr Pro Gln Ile Arg
335 340 345

Lys His Thr Met Ile Leu Met Tyr Asn Trp Phe Thr Ser Ser Val
350 355 360

Leu Tyr Gln Gly Leu Ile Met His Met Gly Leu Ala Gly Asp Asn
365 370 375

Ile Tyr Leu Asp Phe Phe Tyr Ser Ala Leu Val Glu Phe Pro Ala
380 385 390

Ala Phe Met Ile Ile Leu Thr Ile Asp Arg Ile Gly Arg Arg Tyr
395 400 405

Pro Trp Ala Ala Ser Asn Met Val Ala Gly Ala Ala Cys Leu Ala
410 415 420

Ser Val Phe Ile Pro Gly Asp Leu Gln Trp Leu Lys Ile Ile Ile
425 430 435

Ser Cys Leu Gly Arg Met Gly Ile Thr Met Ala Tyr Glu Ile Val
440 445 450

Cys Leu Val Asn Ala Glu Leu Tyr Pro Thr Phe Ile Arg Asn Leu
455 460 465

Gly Val His Ile Cys Ser Ser Met Cys Asp Ile Gly Gly Ile Ile
470 475 480

Thr Pro Phe Leu Val Tyr Arg Leu Thr Asn Ile Trp Leu Glu Leu
485 490 495

Pro Leu Met Val Phe Gly Val Leu Gly Leu Val Ala Gly Gly Leu
500 505 510

Val Leu Leu Leu Pro Glu Thr Lys Gly Lys Ala Leu Pro Glu Thr
515 520 525

Ile Glu Glu Ala Glu Asn Met Gln Arg Pro Arg Lys Asn Lys Glu
530 535 540

Lys Met Ile Tyr Leu Gln Val Gln Lys Leu Asp Ile Pro Leu Asn
545 550 555

<210> 106
<211> 690
<212> PRT
<213> Homo sapiens

<400> 106
Met Ala Pro Trp Pro Glu Leu Gly Asp Ala Gln Pro Asn Pro Asp
1 5 10 15

Lys Tyr Leu Glu Gly Ala Ala Gly Gln Gln Pro Thr Ala Pro Asp
20 25 30

Lys Ser Lys Glu Thr Asn Lys Thr Asp Asn Thr Glu Ala Pro Val
35 40 45

Thr Lys Ile Glu Leu Leu Pro Ser Tyr Ser Thr Ala Thr Leu Ile
50 55 60

Asp Glu Pro Thr Glu Val Asp Asp Pro Trp Asn Leu Pro Thr Leu
65 70 75

Gln Asp Ser Gly Ile Lys Trp Ser Glu Arg Asp Thr Lys Gly Lys
80 85 90

Ile Leu Cys Phe Phe Gln Gly Ile Gly Arg Leu Ile Leu Leu Leu
95 100 105

Gly Phe Leu Tyr Phe Phe Val Cys Ser Leu Asp Ile Leu Ser Ser
110 115 120

Ala Phe Gln Leu Val Gly Gly Lys Met Ala Gly Gln Phe Phe Ser
125 130 135

Asn Ser Ser Ile Met Ser Asn Pro Leu Leu Gly Leu Val Ile Gly
140 145 150

Val Leu Val Thr Val Leu Val Gln Ser Ser Ser Thr Ser Thr Ser
155 160 165

Ile Val Val Ser Met Val Ser Ser Ser Leu Leu Thr Val Arg Ala
170 175 180

Ala Ile Pro Ile Ile Met Gly Ala Asn Ile Gly Thr Ser Ile Thr
185 190 195

Asn Thr Ile Val Ala Leu Met Gln Val Gly Asp Arg Ser Glu Phe
200 205 210

Arg Arg Ala Phe Ala Gly Ala Thr Val His Asp Phe Phe Asn Trp
215 220 225

Leu Ser Val Leu Val Leu Leu Pro Val Glu Val Ala Thr His Tyr
230 235 240

Leu Glu Ile Ile Thr Gln Leu Ile Val Glu Ser Phe His Phe Lys
245 250 255

Asn Gly Glu Asp Ala Pro Asp Leu Leu Lys Val Ile Thr Lys Pro
260 265 270

Phe Thr Lys Leu Ile Val Gln Leu Asp Lys Lys Val Ile Ser Gln

275 280 285

Ile Ala Met Asn Asp Glu Lys Ala Lys Asn Lys Ser Leu Val Lys
290 295 300

Ile Trp Cys Lys Thr Phe Thr Asn Lys Thr Gln Ile Asn Val Thr
305 310 315

Val Pro Ser Thr Ala Asn Cys Thr Ser Pro Ser Leu Cys Trp Thr
320 325 330

Asp Gly Ile Gln Asn Trp Thr Met Lys Asn Val Thr Tyr Lys Glu
335 340 345

Asn Ile Ala Lys Cys Gln His Ile Phe Val Asn Phe His Leu Pro
350 355 360

Asp Leu Ala Val Gly Thr Ile Leu Leu Ile Leu Ser Leu Leu Val
365 370 375

Leu Cys Gly Cys Leu Ile Met Ile Val Lys Ile Leu Gly Ser Val
380 385 390

Leu Lys Gly Gln Val Ala Thr Val Ile Lys Lys Thr Ile Asn Thr
395 400 405

Asp Phe Pro Phe Pro Phe Ala Trp Leu Thr Gly Tyr Leu Ala Ile
410 415 420

Leu Val Gly Ala Gly Met Thr Phe Ile Val Gln Ser Ser Ser Val
425 430 435

Phe Thr Ser Ala Leu Thr Pro Leu Ile Gly Ile Gly Val Ile Thr
440 445 450

Ile Glu Arg Ala Tyr Pro Leu Thr Leu Gly Ser Asn Ile Gly Thr
455 460 465

Thr Thr Thr Ala Ile Leu Ala Ala Leu Ala Ser Pro Gly Asn Ala
470 475 480

Leu Arg Ser Ser Leu Gln Ile Ala Leu Cys His Phe Phe Phe Asn
485 490 495

Ile Ser Gly Ile Leu Leu Trp Tyr Pro Ile Pro Phe Thr Arg Leu
500 505 510

Pro Ile Arg Met Ala Lys Gly Leu Gly Asn Ile Ser Ala Lys Tyr
515 520 525

Arg Trp Phe Ala Val Phe Tyr Leu Ile Ile Phe Phe Phe Leu Ile
530 535 540

Pro Leu Thr Val Phe Gly Leu Ser Leu Ala Gly Trp Arg Val Leu
545 550 555

Val Gly Val Gly Val Pro Val Val Phe Ile Ile Ile Leu Val Leu
560 565 570

Cys Leu Arg Leu Leu Gln Ser Arg Cys Pro Arg Val Leu Pro Lys
575 580 585

Lys Leu Gln Asn Trp Asn Phe Leu Pro Leu Trp Met Arg Ser Leu
590 595 600

Lys Pro Trp Asp Ala Val Val Ser Lys Phe Thr Gly Cys Phe Gln
605 610 615

Met Arg Cys Cys Tyr Cys Cys Arg Val Cys Cys Arg Ala Cys Cys
620 625 630

Leu Leu Cys Gly Cys Pro Lys Cys Cys Arg Cys Ser Lys Cys Cys
635 640 645

Glu Asp Leu Glu Glu Ala Gln Glu Gly Gln Asp Val Pro Val Lys
650 655 660

```
Ala Pro Glu Thr Phe Asp Asn Ile Thr Ile Ser Arg Glu Ala Gln
                665                 670                 675

Gly Glu Val Pro Ala Ser Asp Ser Lys Thr Glu Cys Thr Ala Leu
                680                 685                 690
```

<210> 107
<211> 690
<212> PRT
<213> Homo sapiens

<400> 107

```
Met Ala Pro Trp Pro Glu Leu Gly Asp Ala Gln Pro Asn Pro Asp
  1               5                  10                  15

Lys Tyr Leu Glu Gly Ala Ala Gly Gln Gln Pro Thr Ala Pro Asp
                20                  25                  30

Lys Ser Lys Glu Thr Asn Lys Thr Asp Asn Thr Glu Ala Pro Val
                35                  40                  45

Thr Lys Ile Glu Leu Leu Pro Ser Tyr Ser Thr Ala Thr Leu Ile
                50                  55                  60

Asp Glu Pro Thr Glu Val Asp Asp Pro Trp Asn Leu Pro Thr Leu
                65                  70                  75

Gln Asp Ser Gly Ile Lys Trp Ser Glu Arg Asp Thr Lys Gly Lys
                80                  85                  90

Ile Leu Cys Phe Phe Gln Gly Ile Gly Arg Leu Ile Leu Leu Leu
                95                 100                 105

Gly Phe Leu Tyr Phe Phe Val Cys Ser Leu Asp Ile Leu Ser Ser
               110                 115                 120

Ala Phe Gln Leu Val Gly Gly Lys Met Ala Gly Gln Phe Phe Ser
               125                 130                 135

Asn Ser Ser Ile Met Ser Asn Pro Leu Leu Gly Leu Val Ile Gly
               140                 145                 150

Val Leu Val Thr Val Leu Val Gln Ser Ser Ser Thr Ser Thr Ser
               155                 160                 165

Ile Val Val Ser Met Val Ser Ser Ser Leu Leu Thr Val Arg Ala
               170                 175                 180

Ala Ile Pro Ile Ile Met Gly Ala Asn Ile Gly Thr Ser Ile Thr
               185                 190                 195

Asn Thr Ile Val Ala Leu Met Gln Val Gly Asp Arg Ser Glu Phe
               200                 205                 210

Arg Arg Ala Phe Ala Gly Ala Thr Val His Asp Phe Phe Asn Trp
               215                 220                 225

Leu Ser Val Leu Val Leu Leu Pro Val Glu Val Ala Thr His Tyr
               230                 235                 240

Leu Glu Ile Ile Thr Gln Leu Ile Val Glu Ser Phe His Phe Lys
               245                 250                 255

Asn Gly Glu Asp Ala Pro Asp Leu Leu Lys Val Ile Thr Lys Pro
               260                 265                 270

Phe Thr Lys Leu Ile Val Gln Leu Asp Lys Lys Val Ile Ser Gln
               275                 280                 285

Ile Ala Met Asn Asp Glu Lys Ala Lys Asn Lys Ser Leu Val Lys
               290                 295                 300

Ile Trp Cys Lys Thr Phe Thr Asn Lys Thr Gln Ile Asn Val Thr
               305                 310                 315
```

1490 1400

Val Pro Ser Thr Ala Asn Cys Thr Ser Pro Ser Leu Cys Trp Thr
320 325 330

Asp Gly Ile Gln Asn Trp Thr Met Lys Asn Val Thr Tyr Lys Glu
335 340 345

Asn Ile Ala Lys Cys Gln His Ile Phe Val Asn Phe His Leu Pro
350 355 360

Asp Leu Ala Val Gly Thr Ile Leu Leu Ile Leu Ser Leu Leu Val
365 370 375

Leu Cys Gly Cys Leu Ile Met Ile Val Lys Ile Leu Gly Ser Val
380 385 390

Leu Lys Gly Gln Val Ala Thr Val Ile Lys Lys Thr Ile Asn Thr
395 400 405

Asp Phe Pro Phe Pro Phe Ala Trp Leu Thr Gly Tyr Leu Ala Ile
410 415 420

Leu Val Gly Ala Gly Met Thr Phe Ile Val Gln Ser Ser Ser Val
425 430 435

Phe Thr Ser Ala Leu Thr Pro Leu Ile Gly Ile Gly Val Ile Thr
440 445 450

Ile Glu Arg Ala Tyr Pro Leu Thr Leu Gly Ser Asn Ile Gly Thr
455 460 465

Thr Thr Thr Ala Ile Leu Ala Ala Leu Ala Ser Pro Gly Asn Ala
470 475 480

Leu Arg Ser Ser Leu Gln Ile Ala Leu Cys His Phe Phe Phe Asn
485 490 495

Ile Ser Gly Ile Leu Leu Trp Tyr Pro Ile Pro Phe Thr Arg Leu
500 505 510

Pro Ile Arg Met Ala Lys Gly Leu Gly Asn Ile Ser Ala Lys Tyr
515 520 525

Arg Trp Phe Ala Val Phe Tyr Leu Ile Ile Phe Phe Phe Leu Ile
530 535 540

Pro Leu Thr Val Phe Gly Leu Ser Leu Ala Gly Trp Arg Val Leu
545 550 555

Val Gly Val Gly Val Pro Val Val Phe Ile Ile Ile Leu Val Leu
560 565 570

Cys Leu Arg Leu Leu Gln Ser Arg Cys Pro Arg Val Leu Pro Lys
575 580 585

Lys Leu Gln Asn Trp Asn Phe Leu Pro Leu Trp Met Arg Ser Leu
590 595 600

Lys Pro Trp Asp Ala Val Val Ser Lys Phe Thr Gly Cys Phe Gln
605 610 615

Met Arg Cys Cys Cys Cys Cys Arg Val Cys Cys Arg Ala Cys Cys
620 625 630

Leu Leu Cys Gly Cys Pro Lys Cys Cys Arg Cys Ser Lys Cys Cys
635 640 645

Glu Asp Leu Glu Glu Ala Gln Glu Gly Gln Asp Val Pro Val Lys
650 655 660

Ala Pro Glu Thr Phe Asp Asn Ile Thr Ile Ser Arg Glu Ala Gln
665 670 675

Gly Glu Val Pro Ala Ser Asp Ser Lys Thr Glu Cys Thr Ala Leu
680 685 690

<210> 108

<211> 411
<212> PRT
<213> Homo sapiens

<400> 108

```
Met Trp Ser Gly Trp Trp Leu Trp Pro Leu Val Ala Val Cys Thr
  1               5                  10                  15

Ala Asp Phe Phe Arg Asp Glu Ala Glu Arg Ile Met Arg Asp Ser
            20                  25                  30

Pro Val Ile Asp Gly His Asn Asp Leu Pro Trp Gln Leu Leu Asp
            35                  40                  45

Met Phe Asn Asn Arg Leu Gln Asp Glu Arg Ala Asn Leu Thr Thr
            50                  55                  60

Leu Ala Gly Thr His Thr Asn Ile Pro Lys Leu Arg Ala Gly Phe
            65                  70                  75

Val Gly Gly Gln Phe Trp Ser Val Tyr Thr Pro Cys Asp Thr Gln
            80                  85                  90

Asn Lys Asp Ala Val Arg Arg Thr Leu Glu Gln Met Asp Val Val
            95                  100                 105

His Arg Met Cys Arg Met Tyr Pro Glu Thr Phe Leu Tyr Val Thr
           110                 115                 120

Ser Ser Ala Gly Ile Arg Gln Ala Phe Arg Glu Gly Lys Val Ala
           125                 130                 135

Ser Leu Ile Gly Val Glu Gly Gly His Ser Ile Asp Ser Ser Leu
           140                 145                 150

Gly Val Leu Arg Ala Leu Tyr Gln Leu Gly Met Arg Tyr Leu Thr
           155                 160                 165

Leu Thr His Ser Cys Asn Thr Pro Trp Ala Asp Asn Trp Leu Val
           170                 175                 180

Asp Thr Gly Asp Ser Glu Pro Gln Ser Gln Gly Leu Ser Pro Phe
           185                 190                 195

Gly Gln Arg Val Val Lys Glu Leu Asn Arg Leu Gly Val Leu Ile
           200                 205                 210

Asp Leu Ala His Val Ser Val Ala Thr Met Lys Ala Thr Leu Gln
           215                 220                 225

Leu Ser Arg Ala Pro Val Ile Phe Ser His Ser Ser Ala Tyr Ser
           230                 235                 240

Val Cys Ala Ser Arg Arg Asn Val Pro Asp Asp Val Leu Arg Leu
           245                 250                 255

Val Lys Gln Thr Asp Ser Leu Val Met Val Asn Phe Tyr Asn Asn
           260                 265                 270

Tyr Ile Ser Cys Thr Asn Lys Ala Asn Leu Ser Gln Val Ala Asp
           275                 280                 285

His Leu Asp His Ile Lys Glu Val Ala Gly Ala Arg Ala Val Gly
           290                 295                 300

Phe Gly Gly Asp Phe Asp Gly Val Pro Arg Val Pro Glu Gly Leu
           305                 310                 315

Glu Asp Val Ser Lys Tyr Pro Asp Leu Ile Ala Glu Leu Leu Arg
           320                 325                 330

Arg Asn Trp Thr Glu Ala Glu Val Lys Gly Ala Leu Ala Asp Asn
           335                 340                 345

Leu Leu Arg Val Phe Gln Ala Val Glu Gln Ala Ser Asn Leu Thr
           350                 355                 360
```

Gln Ala Pro Glu Glu Glu Pro Ile Pro Leu Asp Gln Leu Gly Gly
365 370 375

Ser Cys Arg Thr His Tyr Gly Tyr Ser Ser Gly Ala Ser Ser Leu
380 385 390

His Arg His Trp Gly Leu Leu Leu Ala Ser Leu Ala Pro Leu Val
395 400 405

Leu Cys Leu Ser Leu Leu
410

<210> 109
<211> 320
<212> PRT
<213> Homo sapiens

<400> 109
Met Ser Ser Cys Asn Phe Thr His Ala Thr Phe Val Leu Ile Gly
1 5 10 15

Ile Pro Gly Leu Glu Lys Ala His Phe Trp Val Gly Phe Pro Leu
20 25 30

Leu Ser Met Tyr Val Val Ala Met Phe Gly Asn Cys Ile Val Val
35 40 45

Phe Ile Val Arg Thr Glu Arg Ser Leu His Ala Pro Met Tyr Leu
50 55 60

Phe Leu Cys Met Leu Ala Ala Ile Asp Leu Ala Leu Ser Thr Ser
65 70 75

Thr Met Pro Lys Ile Leu Ala Leu Phe Trp Phe Asp Ser Arg Glu
80 85 90

Ile Ser Phe Glu Ala Cys Leu Thr Gln Met Phe Phe Ile His Ala
95 100 105

Leu Ser Ala Ile Glu Ser Thr Ile Leu Leu Ala Met Ala Phe Asp
110 115 120

Arg Tyr Val Ala Ile Cys His Pro Leu Arg His Ala Ala Val Leu
125 130 135

Asn Asn Thr Val Thr Ala Gln Ile Gly Ile Val Ala Val Val Arg
140 145 150

Gly Ser Leu Phe Phe Phe Pro Leu Pro Leu Leu Ile Lys Arg Leu
155 160 165

Ala Phe Cys His Ser Asn Val Leu Ser His Ser Tyr Cys Val His
170 175 180

Gln Asp Val Met Lys Leu Ala Tyr Ala Asp Thr Leu Pro Asn Val
185 190 195

Val Tyr Gly Leu Thr Ala Ile Leu Leu Val Met Gly Val Asp Val
200 205 210

Met Phe Ile Ser Leu Ser Tyr Phe Leu Ile Ile Arg Thr Val Leu
215 220 225

Gln Leu Pro Ser Lys Ser Glu Arg Ala Lys Ala Phe Gly Thr Cys
230 235 240

Val Ser His Ile Gly Val Val Leu Ala Phe Tyr Val Pro Leu Ile
245 250 255

Gly Leu Ser Val Val His Arg Phe Gly Asn Ser Leu His Pro Ile
260 265 270

Val Arg Val Val Met Gly Asp Ile Tyr Leu Leu Leu Pro Pro Val
275 280 285

Ile Asn Pro Ile Ile Tyr Gly Ala Lys Thr Lys Gln Ile Arg Thr
290 295 300

Arg Val Leu Ala Met Phe Lys Ile Ser Cys Asp Lys Asp Leu Gln
305 310 315

Ala Val Gly Gly Lys
320

<210> 110
<211> 2314
<212> PRT
<213> Homo sapiens

<400> 110
Met Arg Ile Leu Lys Arg Phe Leu Ala Cys Ile Gln Leu Leu Cys
1 5 10 15

Val Cys Arg Leu Asp Trp Ala Asn Gly Tyr Tyr Arg Gln Gln Arg
20 25 30

Lys Leu Val Glu Glu Ile Gly Trp Ser Tyr Thr Gly Ala Leu Asn
35 40 45

Gln Lys Asn Trp Gly Lys Lys Tyr Pro Thr Cys Asn Ser Pro Lys
50 55 60

Gln Ser Pro Ile Asn Ile Asp Glu Asp Leu Thr Gln Val Asn Val
65 70 75

Asn Leu Lys Lys Leu Lys Phe Gln Gly Trp Asp Lys Thr Ser Leu
80 85 90

Glu Asn Thr Phe Ile His Asn Thr Gly Lys Thr Val Glu Ile Asn
95 100 105

Leu Thr Asn Asp Tyr Arg Val Ser Gly Gly Val Ser Glu Met Val
110 115 120

Phe Lys Ala Ser Lys Ile Thr Phe His Trp Gly Lys Cys Asn Met
125 130 135

Ser Ser Asp Gly Ser Glu His Ser Leu Glu Gly Gln Lys Phe Pro
140 145 150

Leu Glu Met Gln Ile Tyr Cys Phe Asp Ala Asp Arg Phe Ser Ser
155 160 165

Phe Glu Glu Ala Val Lys Gly Lys Gly Lys Leu Arg Ala Leu Ser
170 175 180

Ile Leu Phe Glu Val Gly Thr Glu Glu Asn Leu Asp Phe Lys Ala
185 190 195

Ile Ile Asp Gly Val Glu Ser Val Ser Arg Phe Gly Lys Gln Ala
200 205 210

Ala Leu Asp Pro Phe Ile Leu Leu Asn Leu Leu Pro Asn Ser Thr
215 220 225

Asp Lys Tyr Tyr Ile Tyr Asn Gly Ser Leu Thr Ser Pro Pro Cys
230 235 240

Thr Asp Thr Val Asp Trp Ile Val Phe Lys Asp Thr Val Ser Ile
245 250 255

Ser Glu Ser Gln Leu Ala Val Phe Cys Glu Val Leu Thr Met Gln
260 265 270

Gln Ser Gly Tyr Val Met Leu Met Asp Tyr Leu Gln Asn Asn Phe
275 280 285

Arg Glu Gln Gln Tyr Lys Phe Ser Arg Gln Val Phe Ser Ser Tyr
290 295 300

Thr Gly Lys Glu Glu Ile His Glu Ala Val Cys Ser Ser Glu Pro

```
                305                 310                 315

Glu Asn Val Gln Ala Asp Pro Glu Asn Tyr Thr Ser Leu Leu Val
                320                 325                 330

Thr Trp Glu Arg Pro Arg Val Val Tyr Asp Thr Met Ile Glu Lys
                335                 340                 345

Phe Ala Val Leu Tyr Gln Gln Leu Asp Gly Glu Asp Gln Thr Lys
                350                 355                 360

His Glu Phe Leu Thr Asp Gly Tyr Gln Asp Leu Gly Ala Ile Leu
                365                 370                 375

Asn Asn Leu Leu Pro Asn Met Ser Tyr Val Leu Gln Ile Val Ala
                380                 385                 390

Ile Cys Thr Asn Gly Leu Tyr Gly Lys Tyr Ser Asp Gln Leu Ile
                395                 400                 405

Val Asp Met Pro Thr Asp Asn Pro Glu Leu Asp Leu Phe Pro Glu
                410                 415                 420

Leu Ile Gly Thr Glu Glu Ile Ile Lys Glu Glu Glu Glu Gly Lys
                425                 430                 435

Asp Ile Glu Glu Gly Ala Ile Val Asn Pro Gly Arg Asp Ser Ala
                440                 445                 450

Thr Asn Gln Ile Arg Lys Lys Glu Pro Gln Ile Ser Thr Thr Thr
                455                 460                 465

His Tyr Asn Arg Ile Gly Thr Lys Tyr Asn Glu Ala Lys Thr Asn
                470                 475                 480

Arg Ser Pro Thr Arg Gly Ser Glu Phe Ser Gly Lys Gly Asp Val
                485                 490                 495

Pro Asn Thr Ser Leu Asn Ser Thr Ser Gln Pro Val Thr Lys Leu
                500                 505                 510

Ala Thr Glu Lys Asp Ile Ser Leu Thr Ser Gln Thr Val Thr Glu
                515                 520                 525

Leu Pro Pro His Thr Val Glu Gly Thr Ser Ala Ser Leu Asn Asp
                530                 535                 540

Gly Ser Lys Thr Val Leu Arg Ser Pro His Met Asn Leu Ser Gly
                545                 550                 555

Thr Ala Glu Ser Leu Asn Thr Val Ser Ile Thr Glu Tyr Glu Glu
                560                 565                 570

Glu Ser Leu Leu Thr Ser Phe Lys Leu Asp Thr Gly Ala Glu Asp
                575                 580                 585

Ser Ser Gly Ser Ser Pro Ala Thr Ser Ala Ile Pro Phe Ile Ser
                590                 595                 600

Glu Asn Ile Ser Gln Gly Tyr Ile Phe Ser Ser Glu Asn Pro Glu
                605                 610                 615

Thr Ile Thr Tyr Asp Val Leu Ile Pro Glu Ser Ala Arg Asn Ala
                620                 625                 630

Ser Glu Asp Ser Thr Ser Ser Gly Ser Glu Glu Ser Leu Lys Asp
                635                 640                 645

Pro Ser Met Glu Gly Asn Val Trp Phe Pro Ser Ser Thr Asp Ile
                650                 655                 660

Thr Ala Gln Pro Asp Val Gly Ser Gly Arg Glu Ser Phe Leu Gln
                665                 670                 675

Thr Asn Tyr Thr Glu Ile Arg Val Asp Glu Ser Glu Lys Thr Thr
                680                 685                 690
```

Lys Ser Phe Ser Ala Gly Pro Val Met Ser Gln Gly Pro Ser Val
695 700 705

Thr Asp Leu Glu Met Pro His Tyr Ser Thr Phe Ala Tyr Phe Pro
710 715 720

Thr Glu Val Thr Pro His Ala Phe Thr Pro Ser Ser Arg Gln Gln
725 730 735

Asp Leu Val Ser Thr Val Asn Val Val Tyr Ser Gln Thr Thr Gln
740 745 750

Pro Val Tyr Asn Gly Glu Thr Pro Leu Gln Pro Ser Tyr Ser Ser
755 760 765

Glu Val Phe Pro Leu Val Thr Pro Leu Leu Leu Asp Asn Gln Ile
770 775 780

Leu Asn Thr Thr Pro Ala Ala Ser Ser Ser Asp Ser Ala Leu His
785 790 795

Ala Thr Pro Val Phe Pro Ser Val Asp Val Ser Phe Glu Ser Ile
800 805 810

Leu Ser Ser Tyr Asp Gly Ala Pro Leu Leu Pro Phe Ser Ser Ala
815 820 825

Ser Phe Ser Ser Glu Leu Phe Arg His Leu His Thr Val Ser Gln
830 835 840

Ile Leu Pro Gln Val Thr Ser Ala Thr Glu Ser Asp Lys Val Pro
845 850 855

Leu His Ala Ser Leu Pro Val Ala Gly Gly Asp Leu Leu Leu Glu
860 865 870

Pro Ser Leu Ala Gln Tyr Ser Asp Val Leu Ser Thr Thr His Ala
875 880 885

Ala Ser Glu Thr Leu Glu Phe Gly Ser Glu Ser Gly Val Leu Tyr
890 895 900

Lys Thr Leu Met Phe Ser Gln Val Glu Pro Pro Ser Ser Asp Ala
905 910 915

Met Met His Ala Arg Ser Ser Gly Pro Glu Pro Ser Tyr Ala Leu
920 925 930

Ser Asp Asn Glu Gly Ser Gln His Ile Phe Thr Val Ser Tyr Ser
935 940 945

Ser Ala Ile Pro Val His Asp Ser Val Gly Val Thr Tyr Gln Gly
950 955 960

Ser Leu Phe Ser Gly Pro Ser His Ile Pro Ile Pro Lys Ser Ser
965 970 975

Leu Ile Thr Pro Thr Ala Ser Leu Leu Gln Pro Thr His Ala Leu
980 985 990

Ser Gly Asp Gly Glu Trp Ser Gly Ala Ser Ser Asp Ser Glu Phe
995 1000 1005

Leu Leu Pro Asp Thr Asp Gly Leu Thr Ala Leu Asn Ile Ser Ser
1010 1015 1020

Pro Val Ser Val Ala Glu Phe Thr Tyr Thr Thr Ser Val Phe Gly
1025 1030 1035

Asp Asp Asn Lys Ala Leu Ser Lys Ser Glu Ile Ile Tyr Gly Asn
1040 1045 1050

Glu Thr Glu Leu Gln Ile Pro Ser Phe Asn Glu Met Val Tyr Pro
1055 1060 1065

Ser Glu Ser Thr Val Met Pro Asn Met Tyr Asp Asn Val Asn Lys
1070 1075 1080

Leu Asn Ala Ser Leu Gln Glu Thr Ser Val Ser Ile Ser Ser Thr
1085 1090 1095

Lys Gly Met Phe Pro Gly Ser Leu Ala His Thr Thr Thr Lys Val
1100 1105 1110

Phe Asp His Glu Ile Ser Gln Val Pro Glu Asn Asn Phe Ser Val
1115 1120 1125

Gln Pro Thr His Thr Val Ser Gln Ala Ser Gly Asp Thr Ser Leu
1130 1135 1140

Lys Pro Val Leu Ser Ala Asn Ser Glu Pro Ala Ser Ser Asp Pro
1145 1150 1155

Ala Ser Ser Glu Met Leu Ser Pro Ser Thr Gln Leu Leu Phe Tyr
1160 1165 1170

Glu Thr Ser Ala Ser Phe Ser Thr Glu Val Leu Leu Gln Pro Ser
1175 1180 1185

Phe Gln Ala Ser Asp Val Asp Thr Leu Leu Lys Thr Val Leu Pro
1190 1195 1200

Ala Val Pro Ser Asp Pro Ile Leu Val Glu Thr Pro Lys Val Asp
1205 1210 1216

Lys Ile Ser Ser Thr Met Leu His Leu Ile Val Ser Asn Ser Ala
1220 1225 1230

Ser Ser Glu Asn Met Leu His Ser Thr Ser Val Pro Val Phe Asp
1235 1240 1245

Val Ser Pro Thr Ser His Met His Ser Ala Ser Leu Gln Gly Leu
1250 1255 1260

Thr Ile Ser Tyr Ala Ser Glu Lys Tyr Glu Pro Val Leu Leu Lys
1265 1270 1275

Ser Glu Ser Ser His Gln Val Val Pro Ser Leu Tyr Ser Asn Asp
1280 1285 1290

Glu Leu Phe Gln Thr Ala Asn Leu Glu Ile Asn Gln Ala His Pro
1295 1300 1305

Pro Lys Gly Arg His Val Phe Ala Thr Pro Val Leu Ser Ile Asp
1310 1315 1320

Glu Pro Leu Asn Thr Leu Ile Asn Lys Leu Ile His Ser Asp Glu
1325 1330 1335

Ile Leu Thr Ser Thr Lys Ser Ser Val Thr Gly Lys Val Phe Ala
1340 1345 1350

Gly Ile Pro Thr Val Ala Ser Asp Thr Phe Val Ser Thr Asp His
1355 1360 1365

Ser Val Pro Ile Gly Asn Gly His Val Ala Ile Thr Ala Val Ser
1370 1375 1380

Pro His Arg Asp Gly Ser Val Thr Ser Thr Lys Leu Leu Phe Pro
1385 1390 1395

Ser Lys Ala Thr Ser Glu Leu Ser His Ser Ala Lys Ser Asp Ala
1400 1405 1410

Gly Leu Val Gly Gly Gly Glu Asp Gly Asp Thr Asp Asp Asp Gly
1415 1420 1425

Asp Asp Asp Asp Asp Arg Asp Ser Asp Gly Leu Ser Ile His Lys
1430 1435 1440

Cys Met Ser Cys Ser Ser Tyr Arg Glu Ser Gln Glu Lys Val Met

1445 1450 1455

Asn Asp Ser Asp Thr His Glu Asn Ser Leu Met Asp Gln Asn Asn
1460 1465 1470

Pro Ile Ser Tyr Ser Leu Ser Glu Asn Ser Glu Glu Asp Asn Arg
1475 1480 1485

Val Thr Ser Val Ser Ser Asp Ser Gln Thr Gly Met Asp Arg Ser
1490 1495 1500

Pro Gly Lys Ser Pro Ser Ala Asn Gly Leu Ser Gln Lys His Asn
1505 1510 1515

Asp Gly Lys Glu Glu Asn Asp Ile Gln Thr Gly Ser Ala Leu Leu
1520 1525 1530

Pro Leu Ser Pro Glu Ser Lys Ala Trp Ala Val Leu Thr Ser Asp
1535 1540 1545

Glu Glu Ser Gly Ser Gly Gln Gly Thr Ser Asp Ser Leu Asn Glu
1550 1555 1560

Asn Glu Thr Ser Thr Asp Phe Ser Phe Ala Asp Thr Asn Glu Lys
1565 1570 1575

Asp Ala Asp Gly Ile Leu Ala Ala Gly Asp Ser Glu Ile Thr Pro
1580 1585 1590

Gly Phe Pro Gln Ser Pro Thr Ser Ser Val Thr Ser Glu Asn Ser
1595 1600 1605

Glu Val Phe His Val Ser Glu Ala Glu Ala Ser Asn Ser Ser His
1610 1615 1620

Glu Ser Arg Ile Gly Leu Ala Glu Gly Leu Glu Ser Glu Lys Lys
1625 1630 1635

Ala Val Ile Pro Leu Val Ile Val Ser Ala Leu Thr Phe Ile Cys
1640 1645 1650

Leu Val Val Leu Val Gly Ile Leu Ile Tyr Trp Arg Lys Cys Phe
1655 1660 1665

Gln Thr Ala His Phe Tyr Leu Glu Asp Ser Thr Ser Pro Arg Val
1670 1675 1680

Ile Ser Thr Pro Pro Thr Pro Ile Phe Pro Ile Ser Asp Asp Val
1685 1690 1695

Gly Ala Ile Pro Ile Lys His Phe Pro Lys His Val Ala Asp Leu
1700 1705 1710

His Ala Ser Ser Gly Phe Thr Glu Glu Phe Glu Thr Leu Lys Glu
1715 1720 1725

Phe Tyr Gln Glu Val Gln Ser Cys Thr Val Asp Leu Gly Ile Thr
1730 1735 1740

Ala Asp Ser Ser Asn His Pro Asp Asn Lys His Lys Asn Arg Tyr
1745 1750 1755

Ile Asn Ile Val Ala Tyr Asp His Ser Arg Val Lys Leu Ala Gln
1760 1765 1770

Leu Ala Glu Lys Asp Gly Lys Leu Thr Asp Tyr Ile Asn Ala Asn
1775 1780 1785

Tyr Val Asp Gly Tyr Asn Arg Pro Lys Ala Tyr Ile Ala Ala Gln
1790 1795 1800

Gly Pro Leu Lys Ser Thr Ala Glu Asp Phe Trp Arg Met Ile Trp
1805 1810 1815

Glu His Asn Val Glu Val Ile Val Met Ile Thr Asn Leu Val Glu
1820 1825 1830

Lys Gly Arg Arg Lys Cys Asp Gln Tyr Trp Pro Ala Asp Gly Ser
1835 1840 1845

Glu Glu Tyr Gly Asn Phe Leu Val Thr Gln Lys Ser Val Gln Val
1850 1855 1860

Leu Ala Tyr Tyr Thr Val Arg Asn Phe Thr Leu Arg Asn Thr Lys
1865 1870 1875

Ile Lys Lys Gly Ser Gln Lys Gly Arg Pro Ser Gly Arg Val Val
1880 1885 1890

Thr Gln Tyr His Tyr Thr Gln Trp Pro Asp Met Gly Val Pro Glu
1895 1900 1905

Tyr Ser Leu Pro Val Leu Thr Phe Val Arg Lys Ala Ala Tyr Ala
1910 1915 1920

Lys Arg His Ala Val Gly Pro Val Val Val His Cys Ser Ala Gly
1925 1930 1935

Val Gly Arg Thr Gly Thr Tyr Ile Val Leu Asp Ser Met Leu Gln
1940 1945 1950

Gln Ile Gln His Glu Gly Thr Val Asn Ile Phe Gly Phe Leu Lys
1955 1960 1965

His Ile Arg Ser Gln Arg Asn Tyr Leu Val Gln Thr Glu Glu Gln
1970 1975 1980

Tyr Val Phe Ile His Asp Thr Leu Val Glu Ala Ile Leu Ser Lys
1985 1990 1995

Glu Thr Glu Val Leu Asp Ser His Ile His Ala Tyr Val Asn Ala
2000 2005 2010

Leu Leu Ile Pro Gly Pro Ala Gly Lys Thr Lys Leu Glu Lys Gln
2015 2020 2025

Phe Gln Leu Leu Ser Gln Ser Asn Ile Gln Gln Ser Asp Tyr Ser
2030 2035 2040

Ala Ala Leu Lys Gln Cys Asn Arg Glu Lys Asn Arg Thr Ser Ser
2045 2050 2055

Ile Ile Pro Val Glu Arg Ser Arg Val Gly Ile Ser Ser Leu Ser
2060 2065 2070

Gly Glu Gly Thr Asp Tyr Ile Asn Ala Ser Tyr Ile Met Gly Tyr
2075 2080 2085

Tyr Gln Ser Asn Glu Phe Ile Ile Thr Gln His Pro Leu Leu His
2090 2095 2100

Thr Ile Lys Asp Phe Trp Arg Met Ile Trp Asp His Asn Ala Gln
2105 2110 2115

Leu Val Val Met Ile Pro Asp Gly Gln Asn Met Ala Glu Asp Glu
2120 2125 2130

Phe Val Tyr Trp Pro Asn Lys Asp Glu Pro Ile Asn Cys Glu Ser
2135 2140 2145

Phe Lys Val Thr Leu Met Ala Glu Glu His Lys Cys Leu Ser Asn
2150 2155 2160

Glu Glu Lys Leu Ile Ile Gln Asp Phe Ile Leu Glu Ala Thr Gln
2165 2170 2175

Asp Asp Tyr Val Leu Glu Val Arg His Phe Gln Cys Pro Lys Trp
2180 2185 2190

Pro Asn Pro Asp Ser Pro Ile Ser Lys Thr Phe Glu Leu Ile Ser
2195 2200 2205

Val Ile Lys Glu Glu Ala Ala Asn Arg Asp Gly Pro Met Ile Val
2210 2215 2220

His Asp Glu His Gly Gly Val Thr Ala Gly Thr Phe Cys Ala Leu
2225 2230 2235

Thr Thr Leu Met His Gln Leu Glu Lys Glu Asn Ser Val Asp Val
2240 2245 2250

Tyr Gln Val Ala Lys Met Ile Asn Leu Met Arg Pro Gly Val Phe
2255 2260 2265

Ala Asp Ile Glu Gln Tyr Gln Phe Leu Tyr Lys Val Ile Leu Ser
2270 2275 2280

Leu Val Ser Thr Arg Gln Glu Glu Asn Pro Ser Thr Ser Leu Asp
2285 2290 2295

Ser Asn Gly Ala Ala Leu Pro Asp Gly Asn Ile Ala Glu Ser Leu
2300 2305 2310

Glu Ser Leu Val

<210> 111
<211> 774
<212> PRT
<213> Homo sapiens

<400> 111
Met Cys Ala Arg Met Ala Gly Arg Thr Arg Ala Ala Pro Arg Gly
1 5 10 15

Pro Tyr Gly Pro Trp Leu Cys Leu Leu Val Ala Leu Ala Leu Asp
20 25 30

Val Val Arg Val Asp Cys Gly Gln Ala Pro Leu Asp Pro Val Tyr
35 40 45

Leu His Val Thr Ala Ala Arg Pro Ala Gln Pro Thr Leu Trp Thr
50 55 60

Ala Lys Leu Asp Arg Phe Lys Gly Ser Arg His His Thr Thr Leu
65 70 75

Ile Thr Cys His Arg Ala Gly Leu Thr Glu Pro Asp Ser Ser Ser
80 85 90

Pro Leu Glu Leu Ser Glu Phe Leu Trp Val Asp Phe Val Val Glu
95 100 105

Asn Ser Thr Gly Gly Gly Val Ala Val Thr Arg Pro Val Thr Trp
110 115 120

Gln Leu Glu Tyr Pro Gly Gln Ala Pro Glu Ala Glu Lys Asp Lys
125 130 135

Met Val Trp Glu Ile Leu Val Ser Glu Arg Asp Ile Arg Ala Leu
140 145 150

Ile Pro Leu Ala Lys Ala Glu Glu Leu Val Asn Thr Ala Pro Leu
155 160 165

Thr Gly Val Pro Gln His Val Pro Val Arg Leu Val Thr Val Asp
170 175 180

Gly Gly Gly Ala Leu Val Glu Val Thr Glu His Val Gly Cys Glu
185 190 195

Ser Ala Asn Thr Gln Val Leu Gln Val Ser Glu Ala Cys Asp Ala
200 205 210

Val Phe Val Ala Gly Lys Glu Ser Arg Gly Ala Arg Gly Val Arg
215 220 225

Val Asp Phe Trp Trp Arg Arg Leu Arg Ala Ser Leu Arg Leu Thr

230 235 240

Val Trp Ala Pro Leu Leu Pro Leu Arg Ile Glu Leu Thr Asp Thr
245 250 255

Thr Leu Glu Gln Val Arg Gly Trp Arg Val Pro Gly Pro Ala Glu
260 265 270

Gly Pro Ala Glu Pro Ala Ala Glu Ala Ser Asp Glu Ala Glu Arg
275 280 285

Arg Ala Arg Gly Cys His Leu Gln Tyr Gln Arg Ala Gly Val Arg
290 295 300

Phe Leu Ala Pro Phe Ala Ala His Pro Leu Asp Gly Gly Arg Arg
305 310 315

Leu Thr His Leu Leu Gly Pro Asp Trp Leu Leu Asp Val Ser His
320 325 330

Leu Val Ala Pro His Ala Arg Val Leu Asp Ser Arg Val Ala Ser
335 340 345

Leu Glu Gly Gly Arg Val Val Val Gly Arg Glu Pro Gly Val Thr
350 355 360

Ser Ile Glu Val Arg Ser Pro Leu Ser Asp Ser Ile Leu Gly Glu
365 370 375

Gln Ala Leu Ala Val Thr Asp Asp Lys Val Ser Val Leu Glu Leu
380 385 390

Arg Val Gln Pro Val Met Gly Ile Ser Leu Thr Leu Ser Arg Gly
395 400 405

Thr Ala His Pro Gly Glu Val Thr Ala Thr Cys Trp Ala Gln Ser
410 415 420

Ala Leu Pro Ala Pro Lys Gln Glu Val Ala Leu Ser Leu Trp Leu
425 430 435

Ser Phe Ser Asp His Thr Val Ala Pro Ala Glu Leu Tyr Asp Arg
440 445 450

Arg Asp Leu Gly Leu Ser Val Ser Ala Glu Glu Pro Gly Ala Ile
455 460 465

Leu Pro Ala Glu Glu Gln Gly Ala Gln Leu Gly Val Val Val Ser
470 475 480

Gly Ala Gly Ala Glu Gly Leu Pro Leu His Val Ala Leu His Pro
485 490 495

Pro Glu Pro Cys Arg Arg Gly Arg His Arg Val Pro Leu Ala Ser
500 505 510

Gly Thr Ala Trp Leu Gly Leu Pro Pro Ala Ser Thr Pro Ala Pro
515 520 525

Ala Leu Pro Ser Ser Pro Ala Trp Ser Pro Pro Ala Thr Glu Ala
530 535 540

Thr Met Gly Gly Lys Arg Gln Val Ala Gly Ser Val Gly Gly Asn
545 550 555

Thr Gly Val Arg Gly Lys Phe Glu Arg Ala Glu Glu Glu Ala Arg
560 565 570

Lys Glu Glu Thr Glu Ala Arg Glu Glu Glu Glu Glu Glu Glu Glu
575 580 585

Glu Met Val Pro Ala Pro Gln His Val Thr Glu Leu Glu Leu Gly
590 595 600

Met Tyr Ala Leu Leu Gly Val Phe Cys Val Ala Ile Phe Ile Phe
605 610 615

1501

Leu Val Asn Gly Val Val Phe Val Leu Arg Tyr Gln Arg Lys Glu
620 625 630

Pro Pro Asp Ser Ala Thr Asp Pro Thr Ser Pro Gln Pro His Asn
635 640 645

Trp Val Trp Leu Gly Thr Asp Gln Glu Glu Leu Ser Arg Gln Leu
650 655 660

Asp Arg Gln Ser Pro Gly Pro Pro Lys Gly Glu Gly Ser Cys Pro
665 670 675

Cys Glu Ser Gly Gly Gly Gly Glu Ala Pro Thr Leu Ala Pro Gly
680 685 690

Pro Pro Gly Gly Thr Thr Ser Ser Ser Ser Thr Leu Ala Arg Lys
695 700 705

Glu Ala Gly Gly Arg Arg Lys Arg Val Glu Phe Val Thr Phe Val
710 715 720

Pro Ala Pro Pro Ala Gln Ser Pro Glu Glu Pro Val Gly Ala Pro
725 730 735

Ala Val Gln Ser Ile Leu Val Ala Gly Glu Glu Asp Ile Arg Trp
740 745 750

Val Cys Glu Asp Met Gly Leu Lys Asp Pro Glu Glu Leu Arg Asn
755 760 765

Tyr Met Glu Arg Ile Arg Gly Ser Ser
770

<210> 112
<211> 420
<212> PRT
<213> Homo sapiens

<400> 112
Met Ala Thr Lys Thr Glu Leu Ser Pro Thr Ala Arg Glu Ser Lys
1 5 10 15

Asn Ala Gln Asp Met Gln Val Asp Glu Thr Leu Ile Pro Arg Lys
20 25 30

Gly Pro Ser Leu Cys Ser Ala Arg Tyr Gly Ile Ala Leu Val Leu
35 40 45

His Phe Cys Asn Phe Thr Thr Ile Ala Gln Asn Val Ile Met Asn
50 55 60

Ile Thr Met Val Ala Met Val Asn Ser Thr Ser Pro Gln Ser Gln
65 70 75

Leu Asn Asp Ser Ser Glu Val Leu Pro Val Asp Ser Phe Gly Gly
80 85 90

Leu Ser Lys Ala Pro Lys Ser Leu Pro Ala Lys Ser Ser Ile Leu
95 100 105

Gly Gly Gln Phe Ala Ile Trp Glu Lys Trp Gly Pro Pro Gln Glu
110 115 120

Arg Ser Arg Leu Cys Ser Ile Ala Leu Ser Gly Met Leu Leu Gly
125 130 135

Cys Phe Thr Ala Ile Leu Ile Gly Gly Phe Ile Ser Glu Thr Leu
140 145 150

Gly Trp Pro Phe Val Phe Tyr Ile Phe Gly Gly Val Gly Cys Val
155 160 165

Cys Cys Leu Leu Trp Phe Val Val Ile Tyr Asp Asp Pro Phe Ser
170 175 180

Tyr Pro Trp Ile Ser Thr Ser Glu Lys Glu Tyr Ile Ile Ser Ser
185 190 195

Leu Lys Gln Gln Val Gly Ser Ser Lys Gln Pro Leu Pro Ile Lys
200 205 210

Ala Met Leu Arg Ser Leu Pro Ile Trp Ser Ile Cys Leu Gly Cys
215 220 225

Phe Ser His Gln Trp Leu Val Ser Thr Met Val Val Tyr Ile Pro
230 235 240

Thr Tyr Ile Ser Ser Val Tyr His Val Asn Ile Arg Asp Asn Gly
245 250 255

Leu Leu Ser Ala Leu Pro Phe Ile Val Ala Trp Val Ile Gly Met
260 265 270

Val Gly Gly Tyr Leu Ala Asp Phe Leu Leu Thr Lys Lys Phe Arg
275 280 285

Leu Ile Thr Val Arg Lys Ile Ala Thr Ile Leu Gly Ser Leu Pro
290 295 300

Ser Ser Ala Leu Ile Val Ser Leu Pro Tyr Leu Asn Ser Gly Tyr
305 310 315

Ile Thr Ala Thr Ala Leu Leu Thr Leu Ser Cys Gly Leu Ser Thr
320 325 330

Leu Cys Gln Ser Gly Ile Tyr Ile Asn Val Leu Asp Ile Ala Pro
335 340 345

Arg Tyr Ser Ser Phe Leu Met Gly Ala Ser Arg Gly Phe Ser Ser
350 355 360

Ile Ala Pro Val Ile Val Pro Thr Val Ser Gly Phe Leu Leu Ser
365 370 375

Gln Asp Pro Glu Phe Gly Trp Arg Asn Val Phe Phe Leu Leu Phe
380 385 390

Ala Val Asn Leu Leu Gly Leu Leu Phe Tyr Leu Ile Phe Gly Glu
395 400 405

Ala Asp Val Gln Glu Trp Ala Lys Glu Arg Lys Leu Thr Arg Leu
410 415 420

<210> 113
<211> 401
<212> PRT
<213> Homo sapiens

<400> 113
Met Gln Val Asp Glu Thr Leu Ile Pro Arg Lys Val Pro Ser Leu
1 5 10 15

Cys Ser Ala Arg Tyr Gly Ile Ala Leu Val Leu His Phe Cys Asn
20 25 30

Phe Thr Thr Ile Ala Gln Asn Val Ile Met Asn Ile Thr Met Val
35 40 45

Ala Met Val Asn Ser Thr Ser Pro Gln Ser Gln Leu Asn Asp Ser
50 55 60

Ser Glu Val Leu Pro Val Asp Ser Phe Gly Gly Leu Ser Lys Ala
65 70 75

Pro Lys Ser Leu Pro Thr Lys Ser Ser Ile Leu Gly Gly Gln Phe
80 85 90

Ala Ile Trp Glu Lys Trp Gly Pro Pro Gln Glu Arg Ser Arg Leu
95 100 105

Cys Ser Ile Ala Leu Ser Gly Met Leu Leu Gly Cys Phe Thr Ala

110 115 120

Ile Leu Ile Gly Gly Phe Ile Ser Glu Thr Leu Gly Trp Pro Phe
125 130 135

Val Phe Tyr Ile Phe Gly Gly Val Gly Cys Val Cys Cys Leu Leu
140 145 150

Trp Phe Val Val Ile Tyr Asp Asp Pro Val Ser Tyr Pro Trp Ile
155 160 165

Ser Thr Ser Glu Lys Glu Tyr Ile Ile Ser Ser Leu Lys Gln Gln
170 175 180

Val Arg Ser Ser Lys Gln Pro Leu Pro Ile Lys Ala Met Leu Arg
185 190 195

Ser Leu Pro Ile Trp Ser Ile Cys Leu Gly Cys Phe Ser His Gln
200 205 210

Trp Leu Val Ser Thr Met Val Val Tyr Ile Pro Thr Tyr Ile Ser
215 220 225

Ser Val Tyr His Val Asn Ile Arg Asp Asn Gly Leu Leu Ser Ala
230 235 240

Leu Pro Phe Ile Val Ala Trp Val Ile Gly Met Val Gly Gly Tyr
245 250 255

Leu Ala Asp Phe Leu Leu Thr Lys Lys Phe Arg Leu Ile Thr Val
260 265 270

Arg Lys Ile Ala Thr Ile Leu Gly Ser Leu Pro Ser Ser Ala Leu
275 280 285

Ile Val Ser Leu Pro Tyr Leu Asn Ser Gly Tyr Ile Thr Ala Thr
290 295 300

Ala Leu Leu Thr Leu Ser Cys Gly Leu Ser Thr Leu Cys Gln Ser
305 310 315

Gly Ile Tyr Ile Asn Val Leu Asp Ile Ala Pro Arg Tyr Ser Ser
320 325 330

Phe Leu Met Gly Ala Ser Arg Gly Phe Ser Ser Ile Ala Pro Val
335 340 345

Ile Val Pro Thr Val Ser Gly Phe Leu Leu Ser Gln Asp Pro Glu
350 355 360

Phe Gly Trp Arg Asn Val Phe Phe Leu Leu Phe Ala Val Asn Leu
365 370 375

Leu Gly Leu Leu Phe Tyr Leu Ile Phe Gly Glu Ala Asp Val Gln
380 385 390

Glu Trp Ala Lys Glu Arg Lys Leu Thr Arg Leu
395 400

<210> 114
<211> 507
<212> PRT
<213> Homo sapiens

<400> 114
Met Ala Gly Ala Gly Pro Lys Arg Arg Ala Leu Ala Ala Pro Ala
1 5 10 15

Ala Glu Glu Lys Glu Glu Ala Arg Glu Lys Met Leu Ala Ala Lys
20 25 30

Ser Ala Asp Gly Ser Ala Pro Ala Gly Glu Gly Glu Gly Val Thr
35 40 45

Leu Gln Arg Asn Ile Thr Leu Leu Asn Gly Val Ala Ile Ile Val
50 55 60

1504

Gly Thr Ile Ile Gly Ser Gly Ile Phe Val Thr Pro Thr Gly Val
65 70 75

Leu Lys Glu Ala Gly Ser Pro Gly Leu Ala Leu Val Val Trp Ala
80 85 90

Ala Cys Gly Val Phe Ser Ile Val Gly Ala Leu Cys Tyr Ala Glu
95 100 105

Leu Gly Thr Thr Ile Ser Lys Ser Gly Gly Asp Tyr Ala Tyr Met
110 115 120

Leu Glu Val Tyr Gly Ser Leu Pro Ala Phe Leu Lys Leu Trp Ile
125 130 135

Glu Leu Leu Ile Ile Arg Pro Ser Ser Gln Tyr Ile Val Ala Leu
140 145 150

Val Phe Ala Thr Tyr Leu Leu Lys Pro Leu Phe Pro Thr Cys Pro
155 160 165

Val Pro Glu Glu Ala Ala Lys Leu Val Ala Cys Leu Cys Val Leu
170 175 180

Leu Leu Thr Ala Val Asn Cys Tyr Ser Val Lys Ala Ala Thr Arg
185 190 195

Val Gln Asp Ala Phe Ala Ala Ala Lys Leu Leu Ala Leu Ala Leu
200 205 210

Ile Ile Leu Leu Gly Phe Val Gln Ile Gly Lys Gly Asp Val Ser
215 220 225

Asn Leu Asp Pro Lys Phe Ser Phe Glu Gly Thr Lys Leu Asp Val
230 235 240

Gly Asn Ile Val Leu Ala Leu Tyr Ser Gly Leu Phe Ala Tyr Gly
245 250 255

Gly Trp Asn Tyr Leu Asn Phe Val Thr Glu Glu Met Ile Asn Pro
260 265 270

Tyr Arg Asn Leu Pro Leu Ala Ile Ile Ile Ser Leu Pro Ile Val
275 280 285

Thr Leu Val Tyr Val Leu Thr Asn Leu Ala Tyr Phe Thr Thr Leu
290 295 300

Ser Thr Glu Gln Met Leu Ser Ser Glu Ala Val Ala Val Asp Phe
305 310 315

Gly Asn Tyr His Leu Gly Val Met Ser Trp Ile Ile Pro Val Phe
320 325 330

Val Gly Leu Ser Cys Phe Gly Ser Val Asn Gly Ser Leu Phe Thr
335 340 345

Ser Ser Arg Leu Phe Phe Val Gly Ser Arg Glu Gly His Leu Pro
350 355 360

Ser Ile Leu Ser Met Ile His Pro Gln Leu Leu Thr Pro Val Pro
365 370 375

Ser Leu Val Phe Thr Cys Val Met Thr Leu Leu Tyr Ala Phe Ser
380 385 390

Lys Asp Ile Phe Ser Val Ile Asn Phe Phe Ser Phe Phe Asn Trp
395 400 405

Leu Cys Val Ala Leu Ala Ile Ile Gly Met Ile Trp Leu Arg His
410 415 420

Arg Lys Pro Glu Leu Glu Arg Pro Ile Lys Val Asn Leu Ala Leu
425 430 435

1505 Tank

```
Pro Val Phe Phe Ile Leu Ala Cys Leu Phe Leu Ile Ala Val Ser
                440                 445                 450

Phe Trp Lys Thr Pro Val Glu Cys Gly Ile Gly Phe Thr Ile Ile
                455                 460                 465

Leu Ser Gly Leu Pro Val Tyr Phe Phe Gly Val Trp Trp Lys Asn
                470                 475                 480

Lys Pro Lys Trp Leu Leu Gln Gly Ile Phe Ser Thr Thr Val Leu
                485                 490                 495

Cys Gln Lys Leu Met Gln Val Val Pro Gln Glu Thr
                500                 505
```

<210> 115
<211> 507
<212> PRT
<213> Homo sapiens

<400> 115

```
Met Ala Gly Ala Gly Pro Lys Arg Arg Ala Leu Ala Ala Pro Ala
  1               5                  10                  15

Ala Glu Glu Lys Glu Glu Ala Arg Glu Lys Met Leu Ala Ala Lys
                20                  25                  30

Ser Ala Asp Gly Ser Ala Pro Ala Gly Glu Gly Glu Gly Val Thr
                35                  40                  45

Leu Gln Arg Asn Ile Thr Leu Leu Asn Gly Val Ala Ile Ile Val
                50                  55                  60

Gly Thr Ile Ile Gly Ser Gly Ile Phe Val Thr Pro Thr Gly Val
                65                  70                  75

Leu Lys Glu Ala Gly Ser Pro Gly Leu Ala Leu Val Val Trp Ala
                80                  85                  90

Ala Cys Gly Val Phe Ser Ile Val Gly Ala Leu Cys Tyr Ala Glu
                95                 100                 105

Leu Gly Thr Thr Ile Ser Lys Ser Gly Gly Asp Tyr Ala Tyr Met
               110                 115                 120

Leu Glu Val Tyr Gly Ser Leu Pro Ala Phe Leu Lys Leu Trp Ile
               125                 130                 135

Glu Leu Leu Ile Ile Arg Pro Ser Ser Gln Tyr Ile Val Ala Leu
               140                 145                 150

Val Phe Ala Thr Tyr Leu Leu Lys Pro Leu Phe Pro Thr Cys Pro
               155                 160                 165

Val Pro Glu Glu Ala Ala Lys Leu Val Ala Cys Leu Cys Val Leu
               170                 175                 180

Leu Leu Thr Ala Val Asn Cys Tyr Ser Val Lys Ala Ala Thr Arg
               185                 190                 195

Val Gln Asp Ala Phe Ala Ala Ala Lys Leu Leu Ala Leu Ala Leu
               200                 205                 210

Ile Ile Leu Leu Gly Phe Val Gln Ile Gly Lys Gly Asp Val Ser
               215                 220                 225

Asn Leu Asp Pro Asn Phe Ser Phe Glu Gly Thr Lys Leu Asp Val
               230                 235                 240

Gly Asn Ile Val Leu Ala Leu Tyr Ser Gly Leu Phe Ala Tyr Gly
               245                 250                 255

Gly Trp Asn Tyr Leu Asn Phe Val Thr Glu Glu Met Ile Asn Pro
               260                 265                 270

Tyr Arg Asn Leu Pro Leu Ala Ile Ile Ile Ser Leu Pro Ile Val
```

```
                275                 280                 285

Thr Leu Val Tyr Val Leu Thr Asn Leu Ala Tyr Phe Thr Thr Leu
                290                 295                 300

Ser Thr Glu Gln Met Leu Ser Ser Glu Ala Val Ala Val Asp Phe
                305                 310                 315

Gly Asn Tyr His Leu Gly Val Met Ser Trp Ile Ile Pro Val Phe
                320                 325                 330

Val Gly Leu Ser Cys Phe Gly Ser Val Asn Gly Ser Leu Phe Thr
                335                 340                 345

Ser Ser Arg Leu Phe Phe Val Gly Ser Arg Glu Gly His Leu Pro
                350                 355                 360

Ser Ile Leu Ser Met Ile His Pro Gln Leu Leu Thr Pro Val Pro
                365                 370                 375

Ser Leu Val Phe Thr Cys Val Met Thr Leu Leu Tyr Ala Phe Ser
                380                 385                 390

Lys Asp Ile Phe Ser Val Ile Asn Phe Phe Ser Phe Phe Asn Trp
                395                 400                 405

Leu Cys Val Ala Leu Ala Ile Ile Gly Met Ile Trp Leu Arg His
                410                 415                 420

Arg Lys Pro Glu Leu Glu Arg Pro Ile Lys Val Asn Leu Ala Leu
                425                 430                 435

Pro Val Phe Phe Ile Leu Ala Cys Leu Phe Leu Ile Ala Val Ser
                440                 445                 450

Phe Trp Lys Thr Pro Val Glu Cys Gly Ile Gly Phe Thr Ile Ile
                455                 460                 465

Leu Ser Gly Leu Pro Val Tyr Phe Phe Gly Val Trp Trp Lys Asn
                470                 475                 480

Lys Pro Lys Trp Leu Leu Gln Gly Ile Phe Ser Thr Thr Val Leu
                485                 490                 495

Cys Gln Lys Leu Met Gln Val Val Pro Gln Glu Thr
                500                 505

<210> 116
<211> 622
<212> PRT
<213> Homo sapiens

<400> 116
 Met Ala Leu Pro Thr Ala Arg Pro Leu Leu Gly Ser Cys Gly Thr
   1               5                  10                  15

 Pro Ala Leu Gly Ser Leu Leu Phe Leu Leu Phe Ser Leu Gly Trp
                  20                  25                  30

 Val Gln Pro Ser Arg Thr Leu Ala Gly Glu Thr Gly Gln Glu Ala
                  35                  40                  45

 Ala Pro Leu Asp Gly Val Leu Ala Asn Pro Pro Asn Ile Ser Ser
                  50                  55                  60

 Leu Ser Pro Arg Gln Leu Leu Gly Phe Pro Cys Ala Glu Val Ser
                  65                  70                  75

 Gly Leu Ser Thr Glu Arg Val Arg Glu Leu Ala Val Ala Leu Ala
                  80                  85                  90

 Gln Lys Asn Val Lys Leu Ser Thr Glu Gln Leu Arg Cys Leu Ala
                  95                 100                 105

 His Arg Leu Ser Glu Pro Pro Glu Asp Leu Asp Ala Leu Pro Leu
                 110                 115                 120
```

Asp Leu Leu Leu Phe Leu Asn Pro Asp Ala Phe Ser Gly Pro Gln
125 130 135

Ala Cys Thr Arg Phe Phe Ser Arg Ile Thr Lys Ala Asn Val Asp
140 145 150

Leu Leu Pro Arg Gly Ala Pro Glu Arg Gln Arg Leu Leu Pro Ala
155 160 165

Ala Leu Ala Cys Trp Gly Val Arg Gly Ser Leu Leu Ser Glu Ala
170 175 180

Asp Val Arg Ala Leu Gly Gly Leu Ala Cys Asp Leu Pro Gly Arg
185 190 195

Phe Val Ala Glu Ser Ala Glu Val Leu Leu Pro Arg Leu Val Ser
200 205 210

Cys Pro Gly Pro Leu Asp Gln Asp Gln Gln Glu Ala Ala Arg Ala
215 220 225

Ala Leu Gln Gly Gly Gly Pro Pro Tyr Gly Pro Pro Ser Thr Trp
230 235 240

Ser Val Ser Thr Met Asp Ala Leu Arg Gly Leu Leu Pro Val Leu
245 250 255

Gly Gln Pro Ile Ile Arg Ser Ile Pro Gln Gly Ile Val Ala Ala
260 265 270

Trp Arg Gln Arg Ser Ser Arg Asp Pro Ser Trp Arg Gln Pro Glu
275 280 285

Arg Thr Ile Leu Arg Pro Arg Phe Arg Arg Glu Val Glu Lys Thr
290 295 300

Ala Cys Pro Ser Gly Lys Lys Ala Arg Glu Ile Asp Glu Ser Leu
305 310 315

Ile Phe Tyr Lys Lys Trp Glu Leu Glu Ala Cys Val Asp Ala Ala
320 325 330

Leu Leu Ala Thr Gln Met Asp Arg Val Asn Ala Ile Pro Phe Thr
335 340 345

Tyr Glu Gln Leu Asp Val Leu Lys His Lys Leu Asp Glu Leu Tyr
350 355 360

Pro Gln Gly Tyr Pro Glu Ser Val Ile Gln His Leu Gly Tyr Leu
365 370 375

Phe Leu Lys Met Ser Pro Glu Asp Ile Arg Lys Trp Asn Val Thr
380 385 390

Ser Leu Glu Thr Leu Lys Ala Leu Leu Glu Val Asn Lys Gly His
395 400 405

Glu Met Ser Pro Gln Val Ala Thr Leu Ile Asp Arg Phe Val Lys
410 415 420

Gly Arg Gly Gln Leu Asp Lys Asp Thr Leu Asp Thr Leu Thr Ala
425 430 435

Phe Tyr Pro Gly Tyr Leu Cys Ser Leu Ser Pro Glu Glu Leu Ser
440 445 450

Ser Val Pro Pro Ser Ser Ile Trp Ala Val Arg Pro Gln Asp Leu
455 460 465

Asp Thr Cys Asp Pro Arg Gln Leu Asp Val Leu Tyr Pro Lys Ala
470 475 480

Arg Leu Ala Phe Gln Asn Met Asn Gly Ser Glu Tyr Phe Val Lys
485 490 495

Ile Gln Ser Phe Leu Gly Gly Ala Pro Thr Glu Asp Leu Lys Ala
500 505 510

Leu Ser Gln Gln Asn Val Ser Met Asp Leu Ala Thr Phe Met Lys
515 520 525

Leu Arg Thr Asp Ala Val Leu Pro Leu Thr Val Ala Glu Val Gln
530 535 540

Lys Leu Leu Gly Pro His Val Glu Gly Leu Lys Ala Glu Glu Arg
545 550 555

His Arg Pro Val Arg Asp Trp Ile Leu Arg Gln Arg Gln Asp Asp
560 565 570

Leu Asp Thr Leu Gly Leu Gly Leu Gln Gly Gly Ile Pro Asn Gly
575 580 585

Tyr Leu Val Leu Asp Leu Ser Val Gln Glu Ala Leu Ser Gly Thr
590 595 600

Pro Cys Leu Leu Gly Pro Gly Pro Val Leu Thr Val Leu Ala Leu
605 610 615

Leu Leu Ala Ser Thr Leu Ala
620

<210> 117
<211> 621
<212> PRT
<213> Homo sapiens

<400> 117
Ala Leu Pro Thr Ala Arg Pro Leu Leu Gly Ser Cys Gly Thr Pro
1 5 10 15

Ala Leu Gly Ser Leu Leu Phe Leu Leu Phe Ser Leu Gly Trp Val
20 25 30

Gln Pro Ser Arg Thr Leu Ala Gly Glu Thr Gly Gln Glu Ala Ala
35 40 45

Pro Leu Asp Gly Val Leu Ala Asn Pro Pro Asn Ile Ser Ser Leu
50 55 60

Ser Pro Arg Gln Leu Leu Gly Phe Pro Cys Ala Glu Val Ser Gly
65 70 75

Leu Ser Thr Glu Arg Val Arg Glu Leu Ala Val Ala Leu Ala Gln
80 85 90

Lys Asn Val Lys Leu Ser Thr Glu Gln Leu Arg Cys Leu Ala His
95 100 105

Arg Leu Ser Glu Pro Pro Glu Asp Leu Asp Ala Leu Pro Leu Asp
110 115 120

Leu Leu Leu Phe Leu Asn Pro Asp Ala Phe Ser Gly Pro Gln Ala
125 130 135

Cys Thr Arg Phe Phe Ser Arg Ile Thr Lys Ala Asn Val Asp Leu
140 145 150

Leu Pro Arg Gly Ala Pro Glu Arg Gln Arg Leu Leu Pro Ala Ala
155 160 165

Leu Ala Cys Trp Gly Val Arg Gly Ser Leu Leu Ser Glu Ala Asp
170 175 180

Val Arg Ala Leu Gly Gly Leu Ala Cys Asp Leu Pro Gly Arg Phe
185 190 195

Val Ala Glu Ser Ala Glu Val Leu Leu Pro Arg Leu Val Ser Cys
200 205 210

Pro Gly Pro Leu Asp Gln Asp Gln Gln Glu Ala Ala Arg Ala Ala

1509 +9000

215 220 225

Leu Gln Gly Gly Gly Pro Pro Tyr Gly Pro Pro Ser Thr Trp Ser
230 235 240

Val Ser Thr Met Asp Ala Leu Arg Gly Leu Leu Pro Val Leu Gly
245 250 255

Gln Pro Ile Ile Arg Ser Ile Pro Gln Gly Ile Val Ala Ala Trp
260 265 270

Arg Gln Arg Ser Ser Arg Asp Pro Ser Trp Arg Gln Pro Glu Arg
275 280 285

Thr Ile Leu Arg Pro Arg Phe Arg Arg Glu Val Glu Lys Thr Ala
290 295 300

Cys Pro Ser Gly Lys Lys Ala Arg Glu Ile Asp Glu Ser Leu Ile
305 310 315

Phe Tyr Lys Lys Trp Glu Leu Glu Ala Cys Val Asp Ala Ala Leu
320 325 330

Leu Ala Thr Gln Met Asp Arg Val Asn Ala Ile Pro Phe Thr Tyr
335 340 345

Glu Gln Leu Asp Val Leu Lys His Lys Leu Asp Glu Leu Tyr Pro
350 355 360

Gln Gly Tyr Pro Glu Ser Val Ile Gln His Leu Gly Tyr Leu Phe
365 370 375

Leu Lys Met Ser Pro Glu Asp Ile Arg Lys Trp Asn Val Thr Ser
380 385 390

Leu Glu Thr Leu Lys Ala Leu Leu Glu Val Asn Lys Gly His Glu
395 400 405

Met Ser Pro Gln Val Ala Thr Leu Ile Asp Arg Phe Val Lys Gly
410 415 420

Arg Gly Gln Leu Asp Lys Asp Thr Leu Asp Thr Leu Thr Ala Phe
425 430 435

Tyr Pro Gly Tyr Leu Cys Ser Leu Ser Pro Glu Glu Leu Ser Ser
440 445 450

Val Pro Pro Ser Ser Ile Trp Ala Val Arg Pro Gln Asp Leu Asp
455 460 465

Thr Cys Asp Pro Arg Gln Leu Asp Val Leu Tyr Pro Lys Ala Arg
470 475 480

Leu Ala Phe Gln Asn Met Asn Gly Ser Glu Tyr Phe Val Lys Ile
485 490 495

Gln Ser Phe Leu Gly Gly Ala Pro Thr Glu Asp Leu Lys Ala Leu
500 505 510

Ser Gln Gln Asn Val Ser Met Asp Leu Ala Thr Phe Met Lys Leu
515 520 525

Arg Thr Asp Ala Val Leu Pro Leu Thr Val Ala Glu Val Gln Lys
530 535 540

Leu Leu Gly Pro His Val Glu Gly Leu Lys Ala Glu Glu Arg His
545 550 555

Arg Pro Val Arg Asp Trp Ile Leu Arg Gln Arg Gln Asp Asp Leu
560 565 570

Asp Thr Leu Gly Leu Gly Leu Gln Gly Gly Ile Pro Asn Gly Tyr
575 580 585

Leu Val Leu Asp Leu Ser Val Gln Glu Ala Leu Ser Gly Thr Pro
590 595 600

1510 1560

Cys Leu Leu Gly Pro Gly Pro Val Leu Thr Val Leu Ala Leu Leu
605 610 615

Leu Ala Ser Thr Leu Ala
620

<210> 118
<211> 357
<212> PRT
<213> Homo sapiens

<400> 118
Met Ala Thr Thr Val Pro Asp Gly Cys Arg Asn Gly Leu Lys Ser
1 5 10 15

Lys Tyr Tyr Arg Leu Cys Asp Lys Ala Glu Ala Trp Gly Ile Val
20 25 30

Leu Glu Thr Val Ala Thr Ala Gly Val Val Thr Ser Val Ala Phe
35 40 45

Met Leu Thr Leu Pro Ile Leu Val Cys Lys Val Gln Asp Ser Asn
50 55 60

Arg Arg Lys Met Leu Pro Thr Gln Phe Leu Phe Leu Leu Gly Val
65 70 75

Leu Gly Ile Phe Gly Leu Thr Phe Ala Phe Ile Ile Gly Leu Asp
80 85 90

Gly Ser Thr Gly Pro Thr Arg Phe Phe Leu Phe Gly Ile Leu Phe
95 100 105

Ser Ile Cys Phe Ser Cys Leu Leu Ala His Ala Val Ser Leu Thr
110 115 120

Lys Leu Val Arg Gly Arg Lys Pro Leu Ser Leu Leu Val Ile Leu
125 130 135

Gly Leu Ala Val Gly Phe Ser Leu Val Gln Asp Val Ile Ala Ile
140 145 150

Glu Tyr Ile Val Leu Thr Met Asn Arg Thr Asn Val Asn Val Phe
155 160 165

Ser Glu Leu Ser Ala Pro Arg Arg Asn Glu Asp Phe Val Leu Leu
170 175 180

Leu Thr Tyr Val Leu Phe Leu Met Ala Leu Thr Phe Leu Met Ser
185 190 195

Ser Phe Thr Phe Cys Gly Ser Phe Thr Gly Trp Lys Arg His Gly
200 205 210

Ala His Ile Tyr Leu Thr Met Leu Leu Ser Ile Ala Ile Trp Val
215 220 225

Ala Trp Ile Thr Leu Leu Met Leu Pro Asp Phe Asp Arg Arg Trp
230 235 240

Asp Asp Thr Ile Leu Ser Ser Ala Leu Ala Ala Asn Gly Trp Val
245 250 255

Phe Leu Leu Ala Tyr Val Ser Pro Glu Phe Trp Leu Leu Thr Lys
260 265 270

Gln Arg Asn Pro Met Asp Tyr Pro Val Glu Asp Ala Phe Cys Lys
275 280 285

Pro Gln Leu Val Lys Lys Ser Tyr Gly Val Glu Asn Arg Ala Tyr
290 295 300

Ser Gln Glu Glu Ile Thr Gln Gly Phe Glu Glu Thr Gly Asp Thr
305 310 315

1511 +50+

Leu Tyr Ala Pro Tyr Ser Thr His Phe Gln Leu Gln Asn Gln Pro
320 325 330

Pro Gln Lys Glu Phe Ser Ile Pro Arg Ala His Ala Trp Pro Ser
335 340 345

Pro Tyr Lys Asp Tyr Glu Val Lys Lys Glu Gly Ser
350 355

<210> 119
<211> 836
<212> PRT
<213> Homo sapiens

<400> 119
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
125 130 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
140 145 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
155 160 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
170 175 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
185 190 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
275 280 285

Val Leu Glu Arg Phe Met Gly Asp Lys Val Ala Ser Glu Ala Leu
290 295 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile

305 310 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
320 325 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
335 340 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
350 355 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
365 370 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
380 385 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
395 400 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
410 415 420

Val Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
425 430 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
440 445 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
455 460 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
470 475 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
485 490 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
500 505 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
515 520 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
530 535 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
545 550 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
560 565 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
575 580 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
590 595 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
605 610 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
620 625 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
635 640 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
650 655 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Thr Thr Lys Ile Ile Thr
665 670 675

Lys Val Val Glu Pro Lys Ile Lys Val Ile Glu Gly Ser Leu Gln
680 685 690

1513

```
Pro Ile Ile Lys Thr Glu Gly Pro Thr Leu Thr Lys Val Lys Ile
                695                 700                 705

Glu Gly Glu Pro Glu Phe Arg Leu Ile Lys Glu Gly Glu Thr Ile
                710                 715                 720

Thr Glu Val Ile His Gly Glu Pro Ile Ile Lys Lys Tyr Thr Lys
                725                 730                 735

Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys Glu Thr Arg
                740                 745                 750

Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr Thr Arg Ile
                755                 760                 765

Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys Lys Leu Leu
                770                 775                 780

Gln Glu Glu Val Thr Lys Val Thr Lys Phe Ile Glu Gly Gly Asp
                785                 790                 795

Gly His Leu Phe Glu Asp Glu Glu Ile Lys Arg Leu Leu Gln Gly
                800                 805                 810

Asp Thr Pro Val Arg Lys Leu Gln Ala Asn Lys Lys Val Gln Gly
                815                 820                 825

Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser Gln
                830                 835

<210> 120
<211> 809
<212> PRT
<213> Homo sapiens

<400> 120
 Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
   1               5                  10                  15

 Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
                20                  25                  30

 His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
                35                  40                  45

 Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
                50                  55                  60

 Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
                65                  70                  75

 Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
                80                  85                  90

 Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
                95                 100                 105

 Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
               110                 115                 120

 Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
               125                 130                 135

 Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
               140                 145                 150

 Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
               155                 160                 165

 His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
               170                 175                 180

 Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
               185                 190                 195
```

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
275 280 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
290 295 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
305 310 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
320 325 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
335 340 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
350 355 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
365 370 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
380 385 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
395 400 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
410 415 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
425 430 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
440 445 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
455 460 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
470 475 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
485 490 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
500 505 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
515 520 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
530 535 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
545 550 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
560 565 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu

575 580 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
590 595 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
605 610 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
620 625 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
635 640 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
650 655 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Arg Pro Thr Leu Thr Lys
665 670 675

Val Lys Ile Glu Gly Glu Pro Glu Phe Arg Leu Ile Lys Glu Gly
680 685 690

Glu Thr Ile Thr Glu Val Ile His Gly Glu Pro Ile Ile Lys Lys
695 700 705

Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
710 715 720

Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
725 730 735

Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
740 745 750

Lys Leu Leu Gln Glu Glu Val Thr Lys Val Thr Lys Phe Ile Glu
755 760 765

Gly Gly Asp Gly His Leu Phe Glu Asp Glu Glu Ile Lys Arg Leu
770 775 780

Leu Gln Gly Asp Thr Pro Val Arg Lys Leu Gln Ala Asn Lys Lys
785 790 795

Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser Gln
800 805

<210> 121
<211> 751
<212> PRT
<213> Homo sapiens

<400> 121
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

1516 1500

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
125 130 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
140 145 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
155 160 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
170 175 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
185 190 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
275 280 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
290 295 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
305 310 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
320 325 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
335 340 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
350 355 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
365 370 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
380 385 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
395 400 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
410 415 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
425 430 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
440 445 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
455 460 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
470 475 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
485 490 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
500 505 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
515 520 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
530 535 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
545 550 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
560 565 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
575 580 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
590 595 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
605 610 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
620 625 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
635 640 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
650 655 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Lys Pro Ile Ile Lys Lys
665 670 675

Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
680 685 690

Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
695 700 705

Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
710 715 720

Lys Leu Leu Gln Glu Asp Thr Pro Val Arg Lys Leu Gln Ala Asn
725 730 735

Lys Lys Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser
740 745 750

Gln

<210> 122
<211> 781
<212> PRT
<213> Homo sapiens

<400> 122
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys

110 115 120

Ile Leu Ile Gly Gly Phe Ile Ser Glu Thr Leu Gly Trp Pro Phe
125 130 135

Val Phe Tyr Ile Phe Gly Gly Val Gly Cys Val Cys Cys Leu Leu
140 145 150

Trp Phe Val Val Ile Tyr Asp Asp Pro Val Ser Tyr Pro Trp Ile
155 160 165

Ser Thr Ser Glu Lys Glu Tyr Ile Ile Ser Ser Leu Lys Gln Gln
170 175 180

Val Arg Ser Ser Lys Gln Pro Leu Pro Ile Lys Ala Met Leu Arg
185 190 195

Ser Leu Pro Ile Trp Ser Ile Cys Leu Gly Cys Phe Ser His Gln
200 205 210

Trp Leu Val Ser Thr Met Val Val Tyr Ile Pro Thr Tyr Ile Ser
215 220 225

Ser Val Tyr His Val Asn Ile Arg Asp Asn Gly Leu Leu Ser Ala
230 235 240

Leu Pro Phe Ile Val Ala Trp Val Ile Gly Met Val Gly Gly Tyr
245 250 255

Leu Ala Asp Phe Leu Leu Thr Lys Lys Phe Arg Leu Ile Thr Val
260 265 270

Arg Lys Ile Ala Thr Ile Leu Gly Ser Leu Pro Ser Ser Ala Leu
275 280 285

Ile Val Ser Leu Pro Tyr Leu Asn Ser Gly Tyr Ile Thr Ala Thr
290 295 300

Ala Leu Leu Thr Leu Ser Cys Gly Leu Ser Thr Leu Cys Gln Ser
305 310 315

Gly Ile Tyr Ile Asn Val Leu Asp Ile Ala Pro Arg Tyr Ser Ser
320 325 330

Phe Leu Met Gly Ala Ser Arg Gly Phe Ser Ser Ile Ala Pro Val
335 340 345

Ile Val Pro Thr Val Ser Gly Phe Leu Leu Ser Gln Asp Pro Glu
350 355 360

Phe Gly Trp Arg Asn Val Phe Phe Leu Leu Phe Ala Val Asn Leu
365 370 375

Leu Gly Leu Leu Phe Tyr Leu Ile Phe Gly Glu Ala Asp Val Gln
380 385 390

Glu Trp Ala Lys Glu Arg Lys Leu Thr Arg Leu
395 400

<210> 114
<211> 507
<212> PRT
<213> Homo sapiens

<400> 114
Met Ala Gly Ala Gly Pro Lys Arg Arg Ala Leu Ala Ala Pro Ala
1 5 10 15

Ala Glu Glu Lys Glu Glu Ala Arg Glu Lys Met Leu Ala Ala Lys
20 25 30

Ser Ala Asp Gly Ser Ala Pro Ala Gly Glu Gly Glu Gly Val Thr
35 40 45

Leu Gln Arg Asn Ile Thr Leu Leu Asn Gly Val Ala Ile Ile Val
50 55 60

Gly Thr Ile Ile Gly Ser Gly Ile Phe Val Thr Pro Thr Gly Val
65 70 75

Leu Lys Glu Ala Gly Ser Pro Gly Leu Ala Leu Val Val Trp Ala
80 85 90

Ala Cys Gly Val Phe Ser Ile Val Gly Ala Leu Cys Tyr Ala Glu
95 100 105

Leu Gly Thr Thr Ile Ser Lys Ser Gly Gly Asp Tyr Ala Tyr Met
110 115 120

Leu Glu Val Tyr Gly Ser Leu Pro Ala Phe Leu Lys Leu Trp Ile
125 130 135

Glu Leu Leu Ile Ile Arg Pro Ser Ser Gln Tyr Ile Val Ala Leu
140 145 150

Val Phe Ala Thr Tyr Leu Leu Lys Pro Leu Phe Pro Thr Cys Pro
155 160 165

Val Pro Glu Glu Ala Ala Lys Leu Val Ala Cys Leu Cys Val Leu
170 175 180

Leu Leu Thr Ala Val Asn Cys Tyr Ser Val Lys Ala Ala Thr Arg
185 190 195

Val Gln Asp Ala Phe Ala Ala Ala Lys Leu Leu Ala Leu Ala Leu
200 205 210

Ile Ile Leu Leu Gly Phe Val Gln Ile Gly Lys Gly Asp Val Ser
215 220 225

Asn Leu Asp Pro Lys Phe Ser Phe Glu Gly Thr Lys Leu Asp Val
230 235 240

Gly Asn Ile Val Leu Ala Leu Tyr Ser Gly Leu Phe Ala Tyr Gly
245 250 255

Gly Trp Asn Tyr Leu Asn Phe Val Thr Glu Glu Met Ile Asn Pro
260 265 270

Tyr Arg Asn Leu Pro Leu Ala Ile Ile Ile Ser Leu Pro Ile Val
275 280 285

Thr Leu Val Tyr Val Leu Thr Asn Leu Ala Tyr Phe Thr Thr Leu
290 295 300

Ser Thr Glu Gln Met Leu Ser Ser Glu Ala Val Ala Val Asp Phe
305 310 315

Gly Asn Tyr His Leu Gly Val Met Ser Trp Ile Ile Pro Val Phe
320 325 330

Val Gly Leu Ser Cys Phe Gly Ser Val Asn Gly Ser Leu Phe Thr
335 340 345

Ser Ser Arg Leu Phe Phe Val Gly Ser Arg Glu Gly His Leu Pro
350 355 360

Ser Ile Leu Ser Met Ile His Pro Gln Leu Leu Thr Pro Val Pro
365 370 375

Ser Leu Val Phe Thr Cys Val Met Thr Leu Leu Tyr Ala Phe Ser
380 385 390

Lys Asp Ile Phe Ser Val Ile Asn Phe Phe Ser Phe Phe Asn Trp
395 400 405

Leu Cys Val Ala Leu Ala Ile Ile Gly Met Ile Trp Leu Arg His
410 415 420

Arg Lys Pro Glu Leu Glu Arg Pro Ile Lys Val Asn Leu Ala Leu
425 430 435

Pro Val Phe Phe Ile Leu Ala Cys Leu Phe Leu Ile Ala Val Ser
440 445 450

Phe Trp Lys Thr Pro Val Glu Cys Gly Ile Gly Phe Thr Ile Ile
455 460 465

Leu Ser Gly Leu Pro Val Tyr Phe Phe Gly Val Trp Trp Lys Asn
470 475 480

Lys Pro Lys Trp Leu Leu Gln Gly Ile Phe Ser Thr Thr Val Leu
485 490 495

Cys Gln Lys Leu Met Gln Val Val Pro Gln Glu Thr
500 505

<210> 115
<211> 507
<212> PRT
<213> Homo sapiens

<400> 115
Met Ala Gly Ala Gly Pro Lys Arg Arg Ala Leu Ala Ala Pro Ala
1 5 10 15

Ala Glu Glu Lys Glu Glu Ala Arg Glu Lys Met Leu Ala Ala Lys
20 25 30

Ser Ala Asp Gly Ser Ala Pro Ala Gly Glu Gly Glu Gly Val Thr
35 40 45

Leu Gln Arg Asn Ile Thr Leu Leu Asn Gly Val Ala Ile Ile Val
50 55 60

Gly Thr Ile Ile Gly Ser Gly Ile Phe Val Thr Pro Thr Gly Val
65 70 75

Leu Lys Glu Ala Gly Ser Pro Gly Leu Ala Leu Val Val Trp Ala
80 85 90

Ala Cys Gly Val Phe Ser Ile Val Gly Ala Leu Cys Tyr Ala Glu
95 100 105

Leu Gly Thr Thr Ile Ser Lys Ser Gly Gly Asp Tyr Ala Tyr Met
110 115 120

Leu Glu Val Tyr Gly Ser Leu Pro Ala Phe Leu Lys Leu Trp Ile
125 130 135

Glu Leu Leu Ile Ile Arg Pro Ser Ser Gln Tyr Ile Val Ala Leu
140 145 150

Val Phe Ala Thr Tyr Leu Leu Lys Pro Leu Phe Pro Thr Cys Pro
155 160 165

Val Pro Glu Glu Ala Ala Lys Leu Val Ala Cys Leu Cys Val Leu
170 175 180

Leu Leu Thr Ala Val Asn Cys Tyr Ser Val Lys Ala Ala Thr Arg
185 190 195

Val Gln Asp Ala Phe Ala Ala Ala Lys Leu Leu Ala Leu Ala Leu
200 205 210

Ile Ile Leu Leu Gly Phe Val Gln Ile Gly Lys Gly Asp Val Ser
215 220 225

Asn Leu Asp Pro Asn Phe Ser Phe Glu Gly Thr Lys Leu Asp Val
230 235 240

Gly Asn Ile Val Leu Ala Leu Tyr Ser Gly Leu Phe Ala Tyr Gly
245 250 255

Gly Trp Asn Tyr Leu Asn Phe Val Thr Glu Glu Met Ile Asn Pro
260 265 270

Tyr Arg Asn Leu Pro Leu Ala Ile Ile Ile Ser Leu Pro Ile Val

275 280 285

Thr Leu Val Tyr Val Leu Thr Asn Leu Ala Tyr Phe Thr Thr Leu
290 295 300

Ser Thr Glu Gln Met Leu Ser Ser Glu Ala Val Ala Val Asp Phe
305 310 315

Gly Asn Tyr His Leu Gly Val Met Ser Trp Ile Ile Pro Val Phe
320 325 330

Val Gly Leu Ser Cys Phe Gly Ser Val Asn Gly Ser Leu Phe Thr
335 340 345

Ser Ser Arg Leu Phe Phe Val Gly Ser Arg Glu Gly His Leu Pro
350 355 360

Ser Ile Leu Ser Met Ile His Pro Gln Leu Leu Thr Pro Val Pro
365 370 375

Ser Leu Val Phe Thr Cys Val Met Thr Leu Leu Tyr Ala Phe Ser
380 385 390

Lys Asp Ile Phe Ser Val Ile Asn Phe Phe Ser Phe Phe Asn Trp
395 400 405

Leu Cys Val Ala Leu Ala Ile Ile Gly Met Ile Trp Leu Arg His
410 415 420

Arg Lys Pro Glu Leu Glu Arg Pro Ile Lys Val Asn Leu Ala Leu
425 430 435

Pro Val Phe Phe Ile Leu Ala Cys Leu Phe Leu Ile Ala Val Ser
440 445 450

Phe Trp Lys Thr Pro Val Glu Cys Gly Ile Gly Phe Thr Ile Ile
455 460 465

Leu Ser Gly Leu Pro Val Tyr Phe Phe Gly Val Trp Trp Lys Asn
470 475 480

Lys Pro Lys Trp Leu Leu Gln Gly Ile Phe Ser Thr Thr Val Leu
485 490 495

Cys Gln Lys Leu Met Glh Val Val Pro Gln Glu Thr
500 505

<210> 116
<211> 622
<212> PRT
<213> Homo sapiens

<400> 116
Met Ala Leu Pro Thr Ala Arg Pro Leu Leu Gly Ser Cys Gly Thr
1 5 10 15

Pro Ala Leu Gly Ser Leu Leu Phe Leu Leu Phe Ser Leu Gly Trp
20 25 30

Val Gln Pro Ser Arg Thr Leu Ala Gly Glu Thr Gly Gln Glu Ala
35 40 45

Ala Pro Leu Asp Gly Val Leu Ala Asn Pro Pro Asn Ile Ser Ser
50 55 60

Leu Ser Pro Arg Gln Leu Leu Gly Phe Pro Cys Ala Glu Val Ser
65 70 75

Gly Leu Ser Thr Glu Arg Val Arg Glu Leu Ala Val Ala Leu Ala
80 85 90

Gln Lys Asn Val Lys Leu Ser Thr Glu Gln Leu Arg Cys Leu Ala
95 100 105

His Arg Leu Ser Glu Pro Pro Glu Asp Leu Asp Ala Leu Pro Leu
110 115 120

Asp Leu Leu Leu Phe Leu Asn Pro Asp Ala Phe Ser Gly Pro Gln
125 130 135

Ala Cys Thr Arg Phe Phe Ser Arg Ile Thr Lys Ala Asn Val Asp
140 145 150

Leu Leu Pro Arg Gly Ala Pro Glu Arg Gln Arg Leu Leu Pro Ala
155 160 165

Ala Leu Ala Cys Trp Gly Val Arg Gly Ser Leu Leu Ser Glu Ala
170 175 180

Asp Val Arg Ala Leu Gly Gly Leu Ala Cys Asp Leu Pro Gly Arg
185 190 195

Phe Val Ala Glu Ser Ala Glu Val Leu Leu Pro Arg Leu Val Ser
200 205 210

Cys Pro Gly Pro Leu Asp Gln Asp Gln Gln Glu Ala Ala Arg Ala
215 220 225

Ala Leu Gln Gly Gly Gly Pro Pro Tyr Gly Pro Pro Ser Thr Trp
230 235 240

Ser Val Ser Thr Met Asp Ala Leu Arg Gly Leu Leu Pro Val Leu
245 250 255

Gly Gln Pro Ile Ile Arg Ser Ile Pro Gln Gly Ile Val Ala Ala
260 265 270

Trp Arg Gln Arg Ser Ser Arg Asp Pro Ser Trp Arg Gln Pro Glu
275 280 285

Arg Thr Ile Leu Arg Pro Arg Phe Arg Arg Glu Val Glu Lys Thr
290 295 300

Ala Cys Pro Ser Gly Lys Lys Ala Arg Glu Ile Asp Glu Ser Leu
305 310 315

Ile Phe Tyr Lys Lys Trp Glu Leu Glu Ala Cys Val Asp Ala Ala
320 325 330

Leu Leu Ala Thr Gln Met Asp Arg Val Asn Ala Ile Pro Phe Thr
335 340 345

Tyr Glu Gln Leu Asp Val Leu Lys His Lys Leu Asp Glu Leu Tyr
350 355 360

Pro Gln Gly Tyr Pro Glu Ser Val Ile Gln His Leu Gly Tyr Leu
365 370 375

Phe Leu Lys Met Ser Pro Glu Asp Ile Arg Lys Trp Asn Val Thr
380 385 390

Ser Leu Glu Thr Leu Lys Ala Leu Leu Glu Val Asn Lys Gly His
395 400 405

Glu Met Ser Pro Gln Val Ala Thr Leu Ile Asp Arg Phe Val Lys
410 415 420

Gly Arg Gly Gln Leu Asp Lys Asp Thr Leu Asp Thr Leu Thr Ala
425 430 435

Phe Tyr Pro Gly Tyr Leu Cys Ser Leu Ser Pro Glu Glu Leu Ser
440 445 450

Ser Val Pro Pro Ser Ser Ile Trp Ala Val Arg Pro Gln Asp Leu
455 460 465

Asp Thr Cys Asp Pro Arg Gln Leu Asp Val Leu Tyr Pro Lys Ala
470 475 480

Arg Leu Ala Phe Gln Asn Met Asn Gly Ser Glu Tyr Phe Val Lys
485 490 495

Ile Gln Ser Phe Leu Gly Gly Ala Pro Thr Glu Asp Leu Lys Ala
500 505 510

Leu Ser Gln Gln Asn Val Ser Met Asp Leu Ala Thr Phe Met Lys
515 520 525

Leu Arg Thr Asp Ala Val Leu Pro Leu Thr Val Ala Glu Val Gln
530 535 540

Lys Leu Leu Gly Pro His Val Glu Gly Leu Lys Ala Glu Glu Arg
545 550 555

His Arg Pro Val Arg Asp Trp Ile Leu Arg Gln Arg Gln Asp Asp
560 565 570

Leu Asp Thr Leu Gly Leu Gly Leu Gln Gly Gly Ile Pro Asn Gly
575 580 585

Tyr Leu Val Leu Asp Leu Ser Val Gln Glu Ala Leu Ser Gly Thr
590 595 600

Pro Cys Leu Leu Gly Pro Gly Pro Val Leu Thr Val Leu Ala Leu
605 610 615

Leu Leu Ala Ser Thr Leu Ala
620

<210> 117
<211> 621
<212> PRT
<213> Homo sapiens

<400> 117
Ala Leu Pro Thr Ala Arg Pro Leu Leu Gly Ser Cys Gly Thr Pro
1 5 10 15

Ala Leu Gly Ser Leu Leu Phe Leu Leu Phe Ser Leu Gly Trp Val
20 25 30

Gln Pro Ser Arg Thr Leu Ala Gly Glu Thr Gly Gln Glu Ala Ala
35 40 45

Pro Leu Asp Gly Val Leu Ala Asn Pro Pro Asn Ile Ser Ser Leu
50 55 60

Ser Pro Arg Gln Leu Leu Gly Phe Pro Cys Ala Glu Val Ser Gly
65 70 75

Leu Ser Thr Glu Arg Val Arg Glu Leu Ala Val Ala Leu Ala Gln
80 85 90

Lys Asn Val Lys Leu Ser Thr Glu Gln Leu Arg Cys Leu Ala His
95 100 105

Arg Leu Ser Glu Pro Pro Glu Asp Leu Asp Ala Leu Pro Leu Asp
110 115 120

Leu Leu Leu Phe Leu Asn Pro Asp Ala Phe Ser Gly Pro Gln Ala
125 130 135

Cys Thr Arg Phe Phe Ser Arg Ile Thr Lys Ala Asn Val Asp Leu
140 145 150

Leu Pro Arg Gly Ala Pro Glu Arg Gln Arg Leu Leu Pro Ala Ala
155 160 165

Leu Ala Cys Trp Gly Val Arg Gly Ser Leu Leu Ser Glu Ala Asp
170 175 180

Val Arg Ala Leu Gly Gly Leu Ala Cys Asp Leu Pro Gly Arg Phe
185 190 195

Val Ala Glu Ser Ala Glu Val Leu Leu Pro Arg Leu Val Ser Cys
200 205 210

Pro Gly Pro Leu Asp Gln Asp Gln Gln Glu Ala Ala Arg Ala Ala

215 220 225

Leu Gln Gly Gly Gly Pro Pro Tyr Gly Pro Pro Ser Thr Trp Ser
230 235 240

Val Ser Thr Met Asp Ala Leu Arg Gly Leu Leu Pro Val Leu Gly
245 250 255

Gln Pro Ile Ile Arg Ser Ile Pro Gln Gly Ile Val Ala Ala Trp
260 265 270

Arg Gln Arg Ser Ser Arg Asp Pro Ser Trp Arg Gln Pro Glu Arg
275 280 285

Thr Ile Leu Arg Pro Arg Phe Arg Arg Glu Val Glu Lys Thr Ala
290 295 300

Cys Pro Ser Gly Lys Lys Ala Arg Glu Ile Asp Glu Ser Leu Ile
305 310 315

Phe Tyr Lys Lys Trp Glu Leu Glu Ala Cys Val Asp Ala Ala Leu
320 325 330

Leu Ala Thr Gln Met Asp Arg Val Asn Ala Ile Pro Phe Thr Tyr
335 340 345

Glu Gln Leu Asp Val Leu Lys His Lys Leu Asp Glu Leu Tyr Pro
350 355 360

Gln Gly Tyr Pro Glu Ser Val Ile Gln His Leu Gly Tyr Leu Phe
365 370 375

Leu Lys Met Ser Pro Glu Asp Ile Arg Lys Trp Asn Val Thr Ser
380 385 390

Leu Glu Thr Leu Lys Ala Leu Leu Glu Val Asn Lys Gly His Glu
395 400 405

Met Ser Pro Gln Val Ala Thr Leu Ile Asp Arg Phe Val Lys Gly
410 415 420

Arg Gly Gln Leu Asp Lys Asp Thr Leu Asp Thr Leu Thr Ala Phe
425 430 435

Tyr Pro Gly Tyr Leu Cys Ser Leu Ser Pro Glu Glu Leu Ser Ser
440 445 450

Val Pro Pro Ser Ser Ile Trp Ala Val Arg Pro Gln Asp Leu Asp
455 460 465

Thr Cys Asp Pro Arg Gln Leu Asp Val Leu Tyr Pro Lys Ala Arg
470 475 480

Leu Ala Phe Gln Asn Met Asn Gly Ser Glu Tyr Phe Val Lys Ile
485 490 495

Gln Ser Phe Leu Gly Gly Ala Pro Thr Glu Asp Leu Lys Ala Leu
500 505 510

Ser Gln Gln Asn Val Ser Met Asp Leu Ala Thr Phe Met Lys Leu
515 520 525

Arg Thr Asp Ala Val Leu Pro Leu Thr Val Ala Glu Val Gln Lys
530 535 540

Leu Leu Gly Pro His Val Glu Gly Leu Lys Ala Glu Glu Arg His
545 550 555

Arg Pro Val Arg Asp Trp Ile Leu Arg Gln Arg Gln Asp Asp Leu
560 565 570

Asp Thr Leu Gly Leu Gly Leu Gln Gly Gly Ile Pro Asn Gly Tyr
575 580 585

Leu Val Leu Asp Leu Ser Val Gln Glu Ala Leu Ser Gly Thr Pro
590 595 600

Cys Leu Leu Gly Pro Gly Pro Val Leu Thr Val Leu Ala Leu Leu
605 610 615

Leu Ala Ser Thr Leu Ala
620

<210> 118
<211> 357
<212> PRT
<213> Homo sapiens

<400> 118
Met Ala Thr Thr Val Pro Asp Gly Cys Arg Asn Gly Leu Lys Ser
1 5 10 15

Lys Tyr Tyr Arg Leu Cys Asp Lys Ala Glu Ala Trp Gly Ile Val
20 25 30

Leu Glu Thr Val Ala Thr Ala Gly Val Val Thr Ser Val Ala Phe
35 40 45

Met Leu Thr Leu Pro Ile Leu Val Cys Lys Val Gln Asp Ser Asn
50 55 60

Arg Arg Lys Met Leu Pro Thr Gln Phe Leu Phe Leu Leu Gly Val
65 70 75

Leu Gly Ile Phe Gly Leu Thr Phe Ala Phe Ile Ile Gly Leu Asp
80 85 90

Gly Ser Thr Gly Pro Thr Arg Phe Phe Leu Phe Gly Ile Leu Phe
95 100 105

Ser Ile Cys Phe Ser Cys Leu Leu Ala His Ala Val Ser Leu Thr
110 115 120

Lys Leu Val Arg Gly Arg Lys Pro Leu Ser Leu Leu Val Ile Leu
125 130 135

Gly Leu Ala Val Gly Phe Ser Leu Val Gln Asp Val Ile Ala Ile
140 145 150

Glu Tyr Ile Val Leu Thr Met Asn Arg Thr Asn Val Asn Val Phe
155 160 165

Ser Glu Leu Ser Ala Pro Arg Arg Asn Glu Asp Phe Val Leu Leu
170 175 180

Leu Thr Tyr Val Leu Phe Leu Met Ala Leu Thr Phe Leu Met Ser
185 190 195

Ser Phe Thr Phe Cys Gly Ser Phe Thr Gly Trp Lys Arg His Gly
200 205 210

Ala His Ile Tyr Leu Thr Met Leu Leu Ser Ile Ala Ile Trp Val
215 220 225

Ala Trp Ile Thr Leu Leu Met Leu Pro Asp Phe Asp Arg Arg Trp
230 235 240

Asp Asp Thr Ile Leu Ser Ser Ala Leu Ala Ala Asn Gly Trp Val
245 250 255

Phe Leu Leu Ala Tyr Val Ser Pro Glu Phe Trp Leu Leu Thr Lys
260 265 270

Gln Arg Asn Pro Met Asp Tyr Pro Val Glu Asp Ala Phe Cys Lys
275 280 285

Pro Gln Leu Val Lys Lys Ser Tyr Gly Val Glu Asn Arg Ala Tyr
290 295 300

Ser Gln Glu Glu Ile Thr Gln Gly Phe Glu Glu Thr Gly Asp Thr
305 310 315

Leu Tyr Ala Pro Tyr Ser Thr His Phe Gln Leu Gln Asn Gln Pro
320 325 330

Pro Gln Lys Glu Phe Ser Ile Pro Arg Ala His Ala Trp Pro Ser
335 340 345

Pro Tyr Lys Asp Tyr Glu Val Lys Lys Glu Gly Ser
350 355

<210> 119
<211> 836
<212> PRT
<213> Homo sapiens

<400> 119
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
125 130 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
140 145 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
155 160 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
170 175 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
185 190 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
275 280 285

Val Leu Glu Arg Phe Met Gly Asp Lys Val Ala Ser Glu Ala Leu
290 295 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile

305 310 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
320 325 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
335 340 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
350 355 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
365 370 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
380 385 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
395 400 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
410 415 420

Val Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
425 430 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
440 445 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
455 460 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
470 475 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
485 490 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
500 505 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
515 520 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
530 535 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
545 550 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
560 565 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
575 580 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
590 595 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
605 610 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
620 625 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
635 640 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
650 655 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Thr Thr Lys Ile Ile Thr
665 670 675

Lys Val Val Glu Pro Lys Ile Lys Val Ile Glu Gly Ser Leu Gln
680 685 690

1528 1548.

```
Pro Ile Ile Lys Thr Glu Gly Pro Thr Leu Thr Lys Val Lys Ile
                695                 700                 705

Glu Gly Glu Pro Glu Phe Arg Leu Ile Lys Glu Gly Glu Thr Ile
                710                 715                 720

Thr Glu Val Ile His Gly Glu Pro Ile Ile Lys Lys Tyr Thr Lys
                725                 730                 735

Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys Glu Thr Arg
                740                 745                 750

Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr Thr Arg Ile
                755                 760                 765

Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys Lys Leu Leu
                770                 775                 780

Gln Glu Glu Val Thr Lys Val Thr Lys Phe Ile Glu Gly Gly Asp
                785                 790                 795

Gly His Leu Phe Glu Asp Glu Glu Ile Lys Arg Leu Leu Gln Gly
                800                 805                 810

Asp Thr Pro Val Arg Lys Leu Gln Ala Asn Lys Lys Val Gln Gly
                815                 820                 825

Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser Gln
                830                 835

<210> 120
<211> 809
<212> PRT
<213> Homo sapiens

<400> 120
 Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
   1               5                  10                  15

 Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
                20                  25                  30

 His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
                35                  40                  45

 Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
                50                  55                  60

 Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
                65                  70                  75

 Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
                80                  85                  90

 Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
                95                 100                 105

 Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
               110                 115                 120

 Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
               125                 130                 135

 Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
               140                 145                 150

 Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
               155                 160                 165

 His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
               170                 175                 180

 Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
               185                 190                 195
```

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
275 280 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
290 295 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
305 310 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
320 325 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
335 340 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
350 355 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
365 370 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
380 385 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
395 400 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
410 415 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
425 430 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
440 445 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
455 460 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
470 475 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
485 490 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
500 505 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
515 520 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
530 535 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
545 550 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
560 565 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu

575 580 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
590 595 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
605 610 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
620 625 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
635 640 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
650 655 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Arg Pro Thr Leu Thr Lys
665 670 675

Val Lys Ile Glu Gly Glu Pro Glu Phe Arg Leu Ile Lys Glu Gly
680 685 690

Glu Thr Ile Thr Glu Val Ile His Gly Glu Pro Ile Ile Lys Lys
695 700 705

Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
710 715 720

Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
725 730 735

Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
740 745 750

Lys Leu Leu Gln Glu Glu Val Thr Lys Val Thr Lys Phe Ile Glu
755 760 765

Gly Gly Asp Gly His Leu Phe Glu Asp Glu Glu Ile Lys Arg Leu
770 775 780

Leu Gln Gly Asp Thr Pro Val Arg Lys Leu Gln Ala Asn Lys Lys
785 790 795

Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser Gln
800 805

<210> 121
<211> 751
<212> PRT
<213> Homo sapiens

<400> 121
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

1531

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
125 130 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
140 145 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
155 160 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
170 175 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
185 190 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
275 280 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
290 295 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
305 310 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
320 325 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
335 340 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
350 355 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
365 370 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
380 385 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
395 400 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
410 415 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
425 430 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
440 445 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
455 460 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
470 475 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
485 490 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
500 505 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
515 520 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
530 535 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
545 550 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
560 565 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
575 580 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
590 595 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
605 610 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
620 625 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
635 640 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
650 655 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Lys Pro Ile Ile Lys Lys
665 670 675

Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
680 685 690

Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
695 700 705

Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
710 715 720

Lys Leu Leu Gln Glu Asp Thr Pro Val Arg Lys Leu Gln Ala Asn
725 730 735

Lys Lys Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser
740 745 750

Gln

<210> 122
<211> 781
<212> PRT
<213> Homo sapiens

<400> 122
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys

80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
125 130 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
140 145 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
155 160 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
170 175 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
185 190 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
275 280 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
290 295 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
305 310 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
320 325 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
335 340 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
350 355 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
365 370 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
380 385 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu
395 400 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
410 415 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
425 430 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
440 445 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
455 460 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
470 475 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
485 490 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
500 505 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
515 520 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
530 535 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
545 550 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
560 565 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
575 580 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
590 595 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
605 610 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
620 625 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
635 640 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
650 655 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Arg Pro Thr Leu Thr Lys
665 670 675

Val Lys Ile Glu Gly Glu Pro Glu Phe Arg Leu Ile Lys Glu Gly
680 685 690

Glu Thr Ile Thr Glu Val Ile His Gly Glu Pro Ile Ile Lys Lys
695 700 705

Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
710 715 720

Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
725 730 735

Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
740 745 750

Lys Leu Leu Gln Glu Asp Thr Pro Val Arg Lys Leu Gln Ala Asn
755 760 765

Lys Lys Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser
770 775 780

Gln

<210> 123
<211> 779
<212> PRT
<213> Homo sapiens

<400> 123
Met Ile Pro Phe Leu Pro Met Phe Ser Leu Leu Leu Leu Leu Ile
1 5 10 15

1535/1505

Val Asn Pro Ile Asn Ala Asn Asn His Tyr Asp Lys Ile Leu Ala
20 25 30

His Ser Arg Ile Arg Gly Arg Asp Gln Gly Pro Asn Val Cys Ala
35 40 45

Leu Gln Gln Ile Leu Gly Thr Lys Lys Lys Tyr Phe Ser Thr Cys
50 55 60

Lys Asn Trp Tyr Lys Lys Ser Ile Cys Gly Gln Lys Thr Thr Val
65 70 75

Leu Tyr Glu Cys Cys Pro Gly Tyr Met Arg Met Glu Gly Met Lys
80 85 90

Gly Cys Pro Ala Val Leu Pro Ile Asp His Val Tyr Gly Thr Leu
95 100 105

Gly Ile Val Gly Ala Thr Thr Thr Gln Arg Tyr Ser Asp Ala Ser
110 115 120

Lys Leu Arg Glu Glu Ile Glu Gly Lys Gly Ser Phe Thr Tyr Phe
125 130 135

Ala Pro Ser Asn Glu Ala Trp Asp Asn Leu Asp Ser Asp Ile Arg
140 145 150

Arg Gly Leu Glu Ser Asn Val Asn Val Glu Leu Leu Asn Ala Leu
155 160 165

His Ser His Met Ile Asn Lys Arg Met Leu Thr Lys Asp Leu Lys
170 175 180

Asn Gly Met Ile Ile Pro Ser Met Tyr Asn Asn Leu Gly Leu Phe
185 190 195

Ile Asn His Tyr Pro Asn Gly Val Val Thr Val Asn Cys Ala Arg
200 205 210

Ile Ile His Gly Asn Gln Ile Ala Thr Asn Gly Val Val His Val
215 220 225

Ile Asp Arg Val Leu Thr Gln Ile Gly Thr Ser Ile Gln Asp Phe
230 235 240

Ile Glu Ala Glu Asp Asp Leu Ser Ser Phe Arg Ala Ala Ala Ile
245 250 255

Thr Ser Asp Ile Leu Glu Ala Leu Gly Arg Asp Gly His Phe Thr
260 265 270

Leu Phe Ala Pro Thr Asn Glu Ala Phe Glu Lys Leu Pro Arg Gly
275 280 285

Val Leu Glu Arg Ile Met Gly Asp Lys Val Ala Ser Glu Ala Leu
290 295 300

Met Lys Tyr His Ile Leu Asn Thr Leu Gln Cys Ser Glu Ser Ile
305 310 315

Met Gly Gly Ala Val Phe Glu Thr Leu Glu Gly Asn Thr Ile Glu
320 325 330

Ile Gly Cys Asp Gly Asp Ser Ile Thr Val Asn Gly Ile Lys Met
335 340 345

Val Asn Lys Lys Asp Ile Val Thr Asn Asn Gly Val Ile His Leu
350 355 360

Ile Asp Gln Val Leu Ile Pro Asp Ser Ala Lys Gln Val Ile Glu
365 370 375

Leu Ala Gly Lys Gln Gln Thr Thr Phe Thr Asp Leu Val Ala Gln
380 385 390

Leu Gly Leu Ala Ser Ala Leu Arg Pro Asp Gly Glu Tyr Thr Leu

1536 1546

395 400 405

Leu Ala Pro Val Asn Asn Ala Phe Ser Asp Asp Thr Leu Ser Met
410 415 420

Asp Gln Arg Leu Leu Lys Leu Ile Leu Gln Asn His Ile Leu Lys
425 430 435

Val Lys Val Gly Leu Asn Glu Leu Tyr Asn Gly Gln Ile Leu Glu
440 445 450

Thr Ile Gly Gly Lys Gln Leu Arg Val Phe Val Tyr Arg Thr Ala
455 460 465

Val Cys Ile Glu Asn Ser Cys Met Glu Lys Gly Ser Lys Gln Gly
470 475 480

Arg Asn Gly Ala Ile His Ile Phe Arg Glu Ile Ile Lys Pro Ala
485 490 495

Glu Lys Ser Leu His Glu Lys Leu Lys Gln Asp Lys Arg Phe Ser
500 505 510

Thr Phe Leu Ser Leu Leu Glu Ala Ala Asp Leu Lys Glu Leu Leu
515 520 525

Thr Gln Pro Gly Asp Trp Thr Leu Phe Val Pro Thr Asn Asp Ala
530 535 540

Phe Lys Gly Met Thr Ser Glu Glu Lys Glu Ile Leu Ile Arg Asp
545 550 555

Lys Asn Ala Leu Gln Asn Ile Ile Leu Tyr His Leu Thr Pro Gly
560 565 570

Val Phe Ile Gly Lys Gly Phe Glu Pro Gly Val Thr Asn Ile Leu
575 580 585

Lys Thr Thr Gln Gly Ser Lys Ile Phe Leu Lys Glu Val Asn Asp
590 595 600

Thr Leu Leu Val Asn Glu Leu Lys Ser Lys Glu Ser Asp Ile Met
605 610 615

Thr Thr Asn Gly Val Ile His Val Val Asp Lys Leu Leu Tyr Pro
620 625 630

Ala Asp Thr Pro Val Gly Asn Asp Gln Leu Leu Glu Ile Leu Asn
635 640 645

Lys Leu Ile Lys Tyr Ile Gln Ile Lys Phe Val Arg Gly Ser Thr
650 655 660

Phe Lys Glu Ile Pro Val Thr Val Tyr Lys Pro Ile Ile Lys Lys
665 670 675

Tyr Thr Lys Ile Ile Asp Gly Val Pro Val Glu Ile Thr Glu Lys
680 685 690

Glu Thr Arg Glu Glu Arg Ile Ile Thr Gly Pro Glu Ile Lys Tyr
695 700 705

Thr Arg Ile Ser Thr Gly Gly Gly Glu Thr Glu Glu Thr Leu Lys
710 715 720

Lys Leu Leu Gln Glu Glu Val Thr Lys Val Thr Lys Phe Ile Glu
725 730 735

Gly Gly Asp Gly His Leu Phe Glu Asp Glu Glu Ile Lys Arg Leu
740 745 750

Leu Gln Gly Asp Thr Pro Val Arg Lys Leu Gln Ala Asn Lys Lys
755 760 765

Val Gln Gly Ser Arg Arg Arg Leu Arg Glu Gly Arg Ser Gln
770 775

<210> 124
<211> 283
<212> PRT
<213> Homo sapiens

<400> 124
Met Val Asn Tyr Ala Trp Ala Gly Arg Ser Gln Arg Lys Leu Trp
1 5 10 15

Trp Arg Ser Val Ala Val Leu Thr Cys Lys Ser Val Val Arg Pro
20 25 30

Gly Tyr Arg Gly Gly Leu Gln Ala Arg Arg Ser Thr Leu Leu Lys
35 40 45

Thr Cys Ala Arg Ala Arg Ala Thr Ala Pro Gly Ala Met Lys Met
50 55 60

Val Ala Pro Trp Thr Arg Phe Tyr Ser Asn Ser Cys Cys Leu Cys
65 70 75

Cys His Val Arg Thr Gly Thr Ile Leu Leu Gly Val Trp Tyr Leu
80 85 90

Ile Ile Asn Ala Val Val Leu Leu Ile Leu Leu Ser Ala Leu Ala
95 100 105

Asp Pro Asp Gln Tyr Asn Phe Ser Ser Ser Glu Leu Gly Gly Asp
110 115 120

Phe Glu Phe Met Asp Asp Ala Asn Met Cys Ile Ala Ile Ala Ile
125 130 135

Ser Leu Leu Met Ile Leu Ile Cys Ala Met Ala Thr Tyr Gly Ala
140 145 150

Tyr Lys Gln Arg Ala Ala Trp Ile Ile Pro Phe Phe Cys Tyr Gln
155 160 165

Ile Phe Asp Phe Ala Leu Asn Met Leu Val Ala Ile Thr Val Leu
170 175 180

Ile Tyr Pro Asn Ser Ile Gln Glu Tyr Ile Arg Gln Leu Pro Pro
185 190 195

Asn Phe Pro Tyr Arg Asp Asp Val Met Ser Val Asn Pro Thr Cys
200 205 210

Leu Val Leu Ile Ile Leu Leu Phe Ile Ser Ile Ile Leu Thr Phe
215 220 225

Lys Gly Tyr Leu Ile Ser Cys Val Trp Asn Cys Tyr Arg Tyr Ile
230 235 240

Asn Gly Arg Asn Ser Ser Asp Val Leu Val Tyr Val Thr Ser Asn
245 250 255

Asp Thr Thr Val Leu Leu Pro Pro Tyr Asp Asp Ala Thr Val Asn
260 265 270

Gly Ala Ala Lys Glu Pro Pro Pro Pro Tyr Val Ser Ala
275 280

<210> 125
<211> 254
<212> PRT
<213> Homo sapiens

<400> 125
Met Pro Gly Gln Gly Gly Leu Gln Ala Arg Arg Ser Thr Leu Leu
1 5 10 15

Lys Thr Cys Ala Arg Ala Arg Ala Thr Ala Pro Gly Ala Met Lys
20 25 30

Met Val Ala Pro Trp Thr Arg Phe Tyr Ser Asn Ser Cys Cys Leu
35 40 45

Cys Cys His Val Arg Thr Gly Thr Ile Leu Leu Gly Val Trp Tyr
50 55 60

Leu Ile Ile Asn Ala Val Val Leu Leu Ile Leu Leu Ser Ala Leu
65 70 75

Ala Asp Pro Asp Gln Tyr Asn Phe Ser Ser Ser Glu Leu Gly Gly
80 85 90

Asp Phe Glu Phe Met Asp Asp Ala Asn Met Cys Ile Ala Ile Ala
95 100 105

Ile Ser Leu Leu Met Ile Leu Ile Cys Ala Met Ala Thr Tyr Gly
110 115 120

Ala Tyr Lys Gln Arg Ala Ala Trp Ile Ile Pro Phe Phe Cys Tyr
125 130 135

Gln Ile Phe Asp Phe Ala Leu Asn Met Leu Val Ala Ile Thr Val
140 145 150

Leu Ile Tyr Pro Asn Ser Ile Gln Glu Tyr Ile Arg Gln Leu Pro
155 160 165

Pro Asn Phe Pro Tyr Arg Asp Asp Val Met Ser Val Asn Pro Thr
170 175 180

Cys Leu Val Leu Ile Ile Leu Leu Phe Ile Ser Ile Ile Leu Thr
185 190 195

Phe Lys Gly Tyr Leu Ile Ser Cys Val Trp Asn Cys Tyr Arg Tyr
200 205 210

Ile Asn Gly Arg Asn Ser Ser Asp Val Leu Val Tyr Val Thr Ser
215 220 225

Asn Asp Thr Thr Val Leu Leu Pro Pro Tyr Asp Asp Ala Thr Val
230 235 240

Asn Gly Ala Ala Lys Glu Pro Pro Pro Pro Tyr Val Ser Ala
245 250

<210> 126
<211> 539
<212> PRT
<213> Homo sapiens

<400> 126
Met Gly Arg Ala Gly Gly Gly Gly Pro Gly Arg Gly Pro Pro Pro
1 5 10 15

Leu Leu Leu Phe Leu Gly Ala Ala Leu Val Leu Ala Ser Gly Ala
20 25 30

Val Pro Ala Arg Glu Ala Gly Ser Ala Val Glu Ala Glu Glu Leu
35 40 45

Val Lys Gly Ser Pro Ala Trp Glu Pro Pro Ala Asn Asp Thr Arg
50 55 60

Glu Glu Ala Gly Pro Pro Ala Ala Gly Glu Asp Glu Ala Ser Trp
65 70 75

Thr Ala Pro Gly Gly Glu Leu Ala Gly Pro Glu Glu Val Leu Gln
80 85 90

Glu Ser Ala Ala Val Thr Gly Thr Ala Trp Leu Glu Ala Asp Ser
95 100 105

Pro Gly Leu Gly Gly Val Thr Ala Glu Ala Gly Ser Gly Asp Ala
110 115 120

Gln Ala Leu Pro Ala Thr Leu Gln Ala Pro His Glu Val Leu Gly

```
                125                 130                 135

Gln Ser Ile Met Pro Pro Ala Ile Pro Glu Ala Thr Glu Ala Ser
                140                 145                 150

Gly Pro Pro Ser Pro Thr Pro Gly Asp Lys Leu Ser Pro Ala Ser
                155                 160                 165

Glu Leu Pro Lys Glu Ser Pro Leu Glu Val Trp Leu Asn Leu Gly
                170                 175                 180

Gly Ser Thr Pro Asp Pro Gln Val Pro Glu Leu Thr Tyr Pro Phe
                185                 190                 195

Gln Gly Thr Leu Glu Pro Gln Pro Ala Ser Asp Ile Ile Asp Ile
                200                 205                 210

Asp Tyr Phe Glu Gly Leu Asp Gly Glu Gly Arg Gly Ala Asp Leu
                215                 220                 225

Gly Ser Phe Pro Gly Ser Pro Gly Thr Ser Glu Asn His Pro Asp
                230                 235                 240

Thr Glu Gly Glu Thr Pro Ser Trp Ser Leu Leu Asp Leu Tyr Asp
                245                 250                 255

Asp Phe Thr Pro Phe Asp Glu Ser Asp Phe Tyr Pro Thr Thr Ser
                260                 265                 270

Phe Tyr Asp Asp Leu Asp Glu Glu Glu Glu Glu Glu Glu Asp Asp
                275                 280                 285

Lys Asp Ala Val Gly Gly Gly Asp Leu Glu Asp Glu Asn Glu Leu
                290                 295                 300

Leu Val Pro Thr Gly Lys Pro Gly Leu Gly Pro Gly Thr Gly Gln
                305                 310                 315

Pro Thr Ser Arg Trp His Ala Val Pro Pro Gln His Thr Leu Gly
                320                 325                 330

Ser Val Pro Gly Ser Ser Ile Ala Leu Arg Pro Arg Pro Gly Glu
                335                 340                 345

Pro Gly Arg Asp Leu Ala Ser Ser Glu Asn Gly Thr Glu Cys Arg
                350                 355                 360

Ser Gly Phe Val Arg His Asn Gly Ser Cys Arg Ser Val Cys Asp
                365                 370                 375

Leu Phe Pro Ser Tyr Cys His Asn Gly Gly Gln Cys Tyr Leu Val
                380                 385                 390

Glu Asn Ile Gly Ala Phe Cys Arg Cys Asn Thr Gln Asp Tyr Ile
                395                 400                 405

Trp His Lys Gly Met Arg Cys Glu Ser Ile Ile Thr Asp Phe Gln
                410                 415                 420

Val Met Cys Val Ala Val Gly Ser Ala Ala Leu Val Leu Leu Leu
                425                 430                 435

Leu Phe Met Met Thr Val Phe Phe Ala Lys Lys Leu Tyr Leu Leu
                440                 445                 450

Lys Thr Glu Asn Thr Lys Leu Arg Arg Thr Asn Lys Phe Arg Thr
                455                 460                 465

Pro Ser Glu Leu His Asn Asp Asn Phe Ser Leu Ser Thr Ile Ala
                470                 475                 480

Glu Gly Ser His Pro Asn Asp Asp Pro Ser Ala Pro His Lys Ile
                485                 490                 495

Gln Glu Val Leu Lys Ser Cys Leu Lys Glu Glu Glu Ser Phe Asn
                500                 505                 510
```

1540
1629

Ile Gln Asn Ser Met Ser Pro Lys Leu Glu Gly Gly Lys Gly Asp
515 520 525

Gln Ala Asp Leu Asp Val Asn Cys Leu Gln Asn Asn Leu Thr
530 535

<210> 127
<211> 1325
<212> PRT
<213> Homo sapiens

<400> 127
Met Leu Pro Val Tyr Gln Glu Val Lys Pro Asn Pro Leu Gln Asp
1 5 10 15

Ala Asn Ile Cys Ser Arg Val Phe Phe Trp Trp Leu Asn Pro Leu
20 25 30

Phe Lys Ile Gly His Lys Arg Arg Leu Glu Glu Asp Asp Met Tyr
35 40 45

Ser Val Leu Pro Glu Asp Arg Ser Gln His Leu Gly Glu Glu Leu
50 55 60

Gln Gly Phe Trp Asp Lys Glu Val Leu Arg Ala Glu Asn Asp Ala
65 70 75

Gln Lys Pro Ser Leu Thr Arg Ala Ile Ile Lys Cys Tyr Trp Lys
80 85 90

Ser Tyr Leu Val Leu Gly Ile Phe Thr Leu Ile Glu Glu Ser Ala
95 100 105

Lys Val Ile Gln Pro Ile Phe Leu Gly Lys Ile Ile Asn Tyr Phe
110 115 120

Glu Asn Tyr Asp Pro Met Asp Ser Val Ala Leu Asn Thr Ala Tyr
125 130 135

Ala Tyr Ala Thr Val Leu Thr Phe Cys Thr Leu Ile Leu Ala Ile
140 145 150

Leu His His Leu Tyr Phe Tyr His Val Gln Cys Ala Gly Met Arg
155 160 165

Leu Arg Val Ala Met Cys His Met Ile Tyr Arg Lys Ala Leu Arg
170 175 180

Leu Ser Asn Met Ala Met Gly Lys Thr Thr Thr Gly Gln Ile Val
185 190 195

Asn Leu Leu Ser Asn Asp Val Asn Lys Phe Asp Gln Val Thr Val
200 205 210

Phe Leu His Phe Leu Trp Ala Gly Pro Leu Gln Ala Ile Ala Val
215 220 225

Thr Ala Leu Leu Trp Met Glu Ile Gly Ile Ser Cys Leu Ala Gly
230 235 240

Met Ala Val Leu Ile Ile Leu Leu Pro Leu Gln Ser Cys Phe Gly
245 250 255

Lys Leu Phe Ser Ser Leu Arg Ser Lys Thr Ala Thr Phe Thr Asp
260 265 270

Ala Arg Ile Arg Thr Met Asn Glu Val Ile Thr Gly Ile Arg Ile
275 280 285

Ile Lys Met Tyr Ala Trp Glu Lys Ser Phe Ser Asn Leu Ile Thr
290 295 300

Asn Leu Arg Lys Lys Glu Ile Ser Lys Ile Leu Arg Ser Ser Cys
305 310 315

Leu Arg Gly Met Asn Leu Ala Ser Phe Phe Ser Ala Ser Lys Ile
320 325 330

Ile Val Phe Val Thr Phe Thr Thr Tyr Val Leu Leu Gly Ser Val
335 340 345

Ile Thr Ala Ser Arg Val Phe Val Ala Val Thr Leu Tyr Gly Ala
350 355 360

Val Arg Leu Thr Val Thr Leu Phe Phe Pro Ser Ala Ile Glu Arg
365 370 375

Val Ser Glu Ala Ile Val Ser Ile Arg Arg Ile Gln Thr Phe Leu
380 385 390

Leu Leu Asp Glu Ile Ser Gln Arg Asn Arg Gln Leu Pro Ser Asp
395 400 405

Gly Lys Lys Met Val His Val Gln Asp Phe Thr Ala Phe Trp Asp
410 415 420

Lys Ala Ser Glu Thr Pro Thr Leu Gln Gly Leu Ser Phe Thr Val
425 430 435

Arg Pro Gly Glu Leu Leu Ala Val Val Gly Pro Val Gly Ala Gly
440 445 450

Lys Ser Ser Leu Leu Ser Ala Val Leu Gly Glu Leu Ala Pro Ser
455 460 465

His Gly Leu Val Ser Val His Gly Arg Ile Ala Tyr Val Ser Gln
470 475 480

Gln Pro Trp Val Phe Ser Gly Thr Leu Arg Ser Asn Ile Leu Phe
485 490 495

Gly Lys Lys Tyr Glu Lys Glu Arg Tyr Glu Lys Val Ile Lys Ala
500 505 510

Cys Ala Leu Lys Lys Asp Leu Gln Leu Leu Glu Asp Gly Asp Leu
515 520 525

Thr Val Ile Gly Asp Arg Gly Thr Thr Leu Ser Gly Gly Gln Lys
530 535 540

Ala Arg Val Asn Leu Ala Arg Ala Val Tyr Gln Asp Ala Asp Ile
545 550 555

Tyr Leu Leu Asp Asp Pro Leu Ser Ala Val Asp Ala Glu Val Ser
560 565 570

Arg His Leu Phe Glu Leu Cys Ile Cys Gln Ile Leu His Glu Lys
575 580 585

Ile Thr Ile Leu Val Thr His Gln Leu Gln Tyr Leu Lys Ala Ala
590 595 600

Ser Gln Ile Leu Ile Leu Lys Asp Gly Lys Met Val Gln Lys Gly
605 610 615

Thr Tyr Thr Glu Phe Leu Lys Ser Gly Ile Asp Phe Gly Ser Leu
620 625 630

Leu Lys Lys Asp Asn Glu Glu Ser Glu Gln Pro Pro Val Pro Gly
635 640 645

Thr Pro Thr Leu Arg Asn Arg Thr Phe Ser Glu Ser Ser Val Trp
650 655 660

Ser Gln Gln Ser Ser Arg Pro Ser Leu Lys Asp Gly Ala Leu Glu
665 670 675

Ser Gln Asp Thr Glu Asn Val Pro Val Thr Leu Ser Glu Glu Asn
680 685 690

Arg Ser Glu Gly Lys Val Gly Phe Gln Ala Tyr Lys Asn Tyr Phe

695 700 705

Arg Ala Gly Ala His Trp Ile Val Phe Ile Phe Leu Ile Leu Leu
710 715 720

Asn Thr Ala Ala Gln Val Ala Tyr Val Leu Gln Asp Trp Trp Leu
725 730 735

Ser Tyr Trp Ala Asn Lys Gln Ser Met Leu Asn Val Thr Val Asn
740 745 750

Gly Gly Gly Asn Val Thr Glu Lys Leu Asp Leu Asn Trp Tyr Leu
755 760 765

Gly Ile Tyr Ser Gly Leu Thr Val Ala Thr Val Leu Phe Gly Ile
770 775 780

Ala Arg Ser Leu Leu Val Phe Tyr Val Leu Val Asn Ser Ser Gln
785 790 795

Thr Leu His Asn Lys Met Phe Glu Ser Ile Leu Lys Ala Pro Val
800 805 810

Leu Phe Phe Asp Arg Asn Pro Ile Gly Arg Ile Leu Asn Arg Phe
815 820 825

Ser Lys Asp Ile Gly His Leu Asp Asp Leu Leu Pro Leu Thr Phe
830 835 840

Leu Asp Phe Ile Gln Thr Leu Leu Gln Val Val Gly Val Val Ser
845 850 855

Val Ala Val Ala Val Ile Pro Trp Ile Ala Ile Pro Leu Val Pro
860 865 870

Leu Gly Ile Ile Phe Ile Phe Leu Arg Arg Tyr Phe Leu Glu Thr
875 880 885

Ser Arg Asp Val Lys Arg Leu Glu Ser Thr Thr Arg Ser Pro Val
890 895 900

Phe Ser His Leu Ser Ser Ser Leu Gln Gly Leu Trp Thr Ile Arg
905 910 915

Ala Tyr Lys Ala Glu Glu Arg Cys Gln Glu Leu Phe Asp Ala His
920 925 930

Gln Asp Leu His Ser Glu Ala Trp Phe Leu Phe Leu Thr Thr Ser
935 940 945

Arg Trp Phe Ala Val Arg Leu Asp Ala Ile Cys Ala Met Phe Val
950 955 960

Ile Ile Val Ala Phe Gly Ser Leu Ile Leu Ala Lys Thr Leu Asp
965 970 975

Ala Gly Gln Val Gly Leu Ala Leu Ser Tyr Ala Leu Thr Leu Met
980 985 990

Gly Met Phe Gln Trp Cys Val Arg Gln Ser Ala Glu Val Glu Asn
995 1000 1005

Met Met Ile Ser Val Glu Arg Val Ile Glu Tyr Thr Asp Leu Glu
1010 1015 1020

Lys Glu Ala Pro Trp Glu Tyr Gln Lys Arg Pro Pro Pro Ala Trp
1025 1030 1035

Pro His Glu Gly Val Ile Ile Phe Asp Asn Val Asn Phe Met Tyr
1040 1045 1050

Ser Pro Gly Gly Pro Leu Val Leu Lys His Leu Thr Ala Leu Ile
1055 1060 1065

Lys Ser Gln Glu Lys Val Gly Ile Val Gly Arg Thr Gly Ala Gly
1070 1075 1080

1543/532

Lys Ser Ser Leu Ile Ser Ala Leu Phe Arg Leu Ser Glu Pro Glu
1085 1090 1095

Gly Lys Ile Trp Ile Asp Lys Ile Leu Thr Thr Glu Ile Gly Leu
1100 1105 1110

His Asp Leu Arg Lys Lys Met Ser Ile Ile Pro Gln Glu Pro Val
1115 1120 1125

Leu Phe Thr Gly Thr Met Arg Lys Asn Leu Asp Pro Phe Lys Glu
1130 1135 1140

His Thr Asp Glu Glu Leu Trp Asn Ala Leu Gln Glu Val Gln Leu
1145 1150 1155

Lys Glu Thr Ile Glu Asp Leu Pro Gly Lys Met Asp Thr Glu Leu
1160 1165 1170

Ala Glu Ser Gly Ser Asn Phe Ser Val Gly Gln Arg Gln Leu Val
1175 1180 1185

Cys Leu Ala Arg Ala Ile Leu Arg Lys Asn Gln Ile Leu Ile Ile
1190 1195 1200

Asp Glu Ala Thr Ala Asn Val Asp Pro Arg Thr Asp Glu Leu Ile
1205 1210 1215

Gln Lys Lys Ile Arg Glu Lys Phe Ala His Cys Thr Val Leu Thr
1220 1225 1230

Ile Ala His Arg Leu Asn Thr Ile Ile Asp Ser Asp Lys Ile Met
1235 1240 1245

Val Leu Asp Ser Gly Arg Leu Lys Glu Tyr Asp Glu Pro Tyr Val
1250 1255 1260

Leu Leu Gln Asn Lys Glu Ser Leu Phe Tyr Lys Met Val Gln Gln
1265 1270 1275

Leu Gly Lys Ala Glu Ala Ala Ala Leu Thr Glu Thr Ala Lys Gln
1280 1285 1290

Val Tyr Phe Lys Arg Asn Tyr Pro His Ile Gly His Thr Asp His
1295 1300 1305

Met Val Thr Asn Thr Ser Asn Gly Gln Pro Ser Thr Leu Thr Ile
1310 1315 1320

Phe Glu Thr Ala Leu
1325

<210> 128
<211> 501
<212> PRT
<213> Homo sapiens

<400> 128
Met Ser Ala Leu Arg Arg Lys Phe Gly Asp Asp Tyr Gln Val Val
1 5 10 15

Thr Thr Ser Ser Ser Gly Ser Gly Leu Gln Pro Gln Gly Pro Gly
20 25 30

Gln Asp Pro Gln Gln Gln Leu Val Pro Lys Lys Lys Arg Gln Arg
35 40 45

Phe Val Asp Lys Asn Gly Arg Cys Asn Val Gln His Gly Asn Leu
50 55 60

Gly Ser Glu Thr Ser Arg Tyr Leu Ser Asp Leu Phe Thr Thr Leu
65 70 75

Val Asp Leu Lys Trp Arg Trp Asn Leu Phe Ile Phe Ile Leu Thr
80 85 90

Tyr Thr Val Ala Trp Leu Phe Met Ala Ser Met Trp Trp Val Ile
95 100 105

Ala Tyr Thr Arg Gly Asp Leu Asn Lys Ala His Val Gly Asn Tyr
110 115 120

Thr Pro Cys Val Ala Asn Val Tyr Asn Phe Pro Ser Ala Phe Leu
125 130 135

Phe Phe Ile Glu Thr Glu Ala Thr Ile Gly Tyr Gly Tyr Arg Tyr
140 145 150

Ile Thr Asp Lys Cys Pro Glu Gly Ile Ile Leu Phe Leu Phe Gln
155 160 165

Ser Ile Leu Gly Ser Ile Val Asp Ala Phe Leu Ile Gly Cys Met
170 175 180

Phe Ile Lys Met Ser Gln Pro Lys Lys Arg Ala Glu Thr Leu Met
185 190 195

Phe Ser Glu His Ala Val Ile Ser Met Arg Asp Gly Lys Leu Thr
200 205 210

Leu Met Phe Arg Val Gly Asn Leu Arg Asn Ser His Met Val Ser
215 220 225

Ala Gln Ile Arg Cys Lys Leu Leu Lys Ser Arg Gln Thr Pro Glu
230 235 240

Gly Glu Phe Leu Pro Leu Asp Gln Leu Glu Leu Asp Val Gly Phe
245 250 255

Ser Thr Gly Ala Asp Gln Leu Phe Leu Val Ser Pro Leu Thr Ile
260 265 270

Cys His Val Ile Asp Ala Lys Ser Pro Phe Tyr Asp Leu Ser Gln
275 280 285

Arg Ser Met Gln Thr Glu Gln Phe Glu Ile Val Val Ile Leu Glu
290 295 300

Gly Ile Val Glu Thr Thr Gly Met Thr Cys Gln Ala Arg Thr Ser
305 310 315

Tyr Thr Glu Asp Glu Val Leu Trp Gly His Arg Phe Phe Pro Val
320 325 330

Ile Ser Leu Glu Glu Gly Phe Phe Lys Val Asp Tyr Ser Gln Phe
335 340 345

His Ala Thr Phe Glu Val Pro Thr Pro Pro Tyr Ser Val Lys Glu
350 355 360

Gln Glu Glu Met Leu Leu Met Ser Ser Pro Leu Ile Ala Pro Ala
365 370 375

Ile Thr Asn Ser Lys Glu Arg His Asn Ser Val Glu Cys Leu Asp
380 385 390

Gly Leu Asp Asp Ile Thr Thr Lys Leu Pro Ser Lys Leu Gln Lys
395 400 405

Ile Thr Gly Arg Glu Asp Phe Pro Lys Lys Leu Leu Arg Met Ser
410 415 420

Ser Thr Thr Ser Glu Lys Ala Tyr Ser Leu Gly Asp Leu Pro Met
425 430 435

Lys Leu Gln Arg Ile Ser Ser Val Pro Gly Asn Ser Glu Glu Lys
440 445 450

Leu Val Ser Lys Thr Thr Lys Met Leu Ser Asp Pro Met Ser Gln
455 460 465

Ser Val Ala Asp Leu Pro Pro Lys Leu Gln Lys Met Ala Gly Gly

470 475 480

Ala Ala Arg Met Glu Gly Asn Leu Pro Ala Lys Leu Arg Lys Met
485 490 495

Asn Ser Asp Arg Phe Thr
500

<210> 129
<211> 1148
<212> PRT
<213> Homo sapiens

<400> 129
Met Pro Leu Phe Lys Asn Thr Ser Val Ser Ser Leu Tyr Ser Gly
1 5 10 15

Cys Arg Leu Thr Leu Leu Arg Pro Glu Lys Asp Gly Ala Ala Thr
20 25 30

Arg Val Asp Ala Val Cys Thr His Arg Pro Asp Pro Lys Ser Pro
35 40 45

Gly Leu Asp Arg Glu Arg Leu Tyr Trp Lys Leu Ser Gln Leu Thr
50 55 60

His Gly Ile Thr Glu Leu Gly Pro Tyr Thr Leu Asp Arg His Ser
65 70 75

Leu Tyr Val Asn Gly Phe Thr His Gln Ser Ser Met Thr Thr Thr
80 85 90

Arg Thr Pro Asp Thr Ser Thr Met His Leu Ala Thr Ser Arg Thr
95 100 105

Pro Ala Ser Leu Ser Gly Pro Thr Thr Ala Ser Pro Leu Leu Val
110 115 120

Leu Phe Thr Ile Asn Phe Thr Ile Thr Asn Leu Arg Tyr Glu Glu
125 130 135

Asn Met His His Pro Gly Ser Arg Lys Phe Asn Thr Thr Glu Arg
140 145 150

Val Leu Gln Gly Leu Leu Arg Pro Val Phe Lys Asn Thr Ser Val
155 160 165

Gly Pro Leu Tyr Ser Gly Cys Arg Leu Thr Leu Leu Arg Pro Lys
170 175 180

Lys Asp Gly Ala Ala Thr Lys Val Asp Ala Ile Cys Thr Tyr Arg
185 190 195

Pro Asp Pro Lys Ser Pro Gly Leu Asp Arg Glu Gln Leu Tyr Trp
200 205 210

Glu Leu Ser Gln Leu Thr His Ser Ile Thr Glu Leu Gly Pro Tyr
215 220 225

Thr Leu Asp Arg Asp Ser Leu Tyr Val Asn Gly Phe Thr Gln Arg
230 235 240

Ser Ser Val Pro Thr Thr Ser Ile Pro Gly Thr Pro Thr Val Asp
245 250 255

Leu Gly Thr Ser Gly Thr Pro Val Ser Lys Pro Gly Pro Ser Ala
260 265 270

Ala Ser Pro Leu Leu Val Leu Phe Thr Leu Asn Phe Thr Ile Thr
275 280 285

Asn Leu Arg Tyr Glu Glu Asn Met Gln His Pro Gly Ser Arg Lys
290 295 300

Phe Asn Thr Thr Glu Arg Val Leu Gln Gly Leu Leu Arg Ser Leu
305 310 315

Phe Lys Ser Thr Ser Val Gly Pro Leu Tyr Ser Gly Cys Arg Leu
320 325 330

Thr Leu Leu Arg Pro Glu Lys Asp Gly Thr Ala Thr Gly Val Asp
335 340 345

Ala Ile Cys Thr His His Pro Asp Pro Lys Ser Pro Arg Leu Asp
350 355 360

Arg Glu Gln Leu Tyr Trp Glu Leu Ser Gln Leu Thr His Asn Ile
365 370 375

Thr Glu Leu Gly His Tyr Ala Leu Asp Asn Asp Ser Leu Phe Val
380 385 390

Asn Gly Phe Thr His Arg Ser Ser Val Ser Thr Thr Ser Thr Pro
395 400 405

Gly Thr Pro Thr Val Tyr Leu Gly Ala Ser Lys Thr Pro Ala Ser
410 415 420

Ile Phe Gly Pro Ser Ala Ala Ser His Leu Leu Ile Leu Phe Thr
425 430 435

Leu Asn Phe Thr Ile Thr Asn Leu Arg Tyr Glu Glu Asn Met Trp
440 445 450

Pro Gly Ser Arg Lys Phe Asn Thr Thr Glu Arg Val Leu Gln Gly
455 460 465

Leu Leu Arg Pro Leu Phe Lys Asn Thr Ser Val Gly Pro Leu Tyr
470 475 480

Ser Gly Ser Arg Leu Thr Leu Leu Arg Pro Glu Lys Asp Gly Glu
485 490 495

Ala Thr Gly Val Asp Ala Ile Cys Thr His Arg Pro Asp Pro Thr
500 505 510

Gly Pro Gly Leu Asp Arg Glu Gln Leu Tyr Leu Glu Leu Ser Gln
515 520 525

Leu Thr His Ser Ile Thr Glu Leu Gly Pro Tyr Thr Leu Asp Arg
530 535 540

Asp Ser Leu Tyr Val Asn Gly Phe Thr His Arg Ser Ser Val Pro
545 550 555

Thr Thr Ser Thr Gly Val Val Ser Glu Glu Pro Phe Thr Leu Asn
560 565 570

Phe Thr Ile Asn Asn Leu Arg Tyr Met Ala Asp Met Gly Gln Pro
575 580 585

Gly Ser Leu Lys Phe Asn Ile Thr Asp Asn Val Met Lys His Leu
590 595 600

Leu Ser Pro Leu Phe Gln Arg Ser Ser Leu Gly Ala Arg Tyr Thr
605 610 615

Gly Cys Arg Val Ile Ala Leu Arg Ser Val Lys Asn Gly Ala Glu
620 625 630

Thr Arg Val Asp Leu Leu Cys Thr Tyr Leu Gln Pro Leu Ser Gly
635 640 645

Pro Gly Leu Pro Ile Lys Gln Val Phe His Glu Leu Ser Gln Gln
650 655 660

Thr His Gly Ile Thr Arg Leu Gly Pro Tyr Ser Leu Asp Lys Asp
665 670 675

Ser Leu Tyr Leu Asn Gly Tyr Asn Glu Pro Gly Leu Asp Glu Pro
680 685 690

1546
1547

Pro Thr Thr Pro Lys Pro Ala Thr Thr Phe Leu Pro Pro Leu Ser
695 700 705

Glu Ala Thr Thr Ala Met Gly Tyr His Leu Lys Thr Leu Thr Leu
710 715 720

Asn Phe Thr Ile Ser Asn Leu Gln Tyr Ser Pro Asp Met Gly Lys
725 730 735

Gly Ser Ala Thr Phe Asn Ser Thr Glu Gly Val Leu Gln His Leu
740 745 750

Leu Arg Pro Leu Phe Gln Lys Ser Ser Met Gly Pro Phe Tyr Leu
755 760 765

Gly Cys Gln Leu Ile Ser Leu Arg Pro Glu Lys Asp Gly Ala Ala
770 775 780

Thr Gly Val Asp Thr Thr Cys Thr Tyr His Pro Asp Pro Val Gly
785 790 795

Pro Gly Leu Asp Ile Gln Gln Leu Tyr Trp Glu Leu Ser Gln Leu
800 805 810

Thr His Gly Val Thr Gln Leu Gly Phe Tyr Val Leu Asp Arg Asp
815 820 825

Ser Leu Phe Ile Asn Gly Tyr Ala Pro Gln Asn Leu Ser Ile Arg
830 835 840

Gly Glu Tyr Gln Ile Asn Phe His Ile Val Asn Trp Asn Leu Ser
845 850 855

Asn Pro Asp Pro Thr Ser Ser Glu Tyr Ile Thr Leu Leu Arg Asp
860 865 870

Ile Gln Asp Lys Val Thr Thr Leu Tyr Lys Gly Ser Gln Leu His
875 880 885

Asp Thr Phe Arg Phe Cys Leu Val Thr Asn Leu Thr Met Asp Ser
890 895 900

Val Leu Val Thr Val Lys Ala Leu Phe Ser Ser Asn Leu Asp Pro
905 910 915

Ser Leu Val Glu Gln Val Phe Leu Asp Lys Thr Leu Asn Ala Ser
920 925 930

Phe His Trp Leu Gly Ser Thr Tyr Gln Leu Val Asp Ile His Val
935 940 945

Thr Glu Met Glu Ser Ser Val Tyr Gln Pro Thr Ser Ser Ser Ser
950 955 960

Thr Gln His Phe Tyr Pro Asn Phe Thr Ile Thr Asn Leu Pro Tyr
965 970 975

Ser Gln Asp Lys Ala Gln Pro Gly Thr Thr Asn Tyr Gln Arg Asn
980 985 990

Lys Arg Asn Ile Glu Asp Ala Leu Asn Gln Leu Phe Arg Asn Ser
995 1000 1005

Ser Ile Lys Ser Tyr Phe Ser Asp Cys Gln Val Ser Thr Phe Arg
1010 1015 1020

Ser Val Pro Asn Arg His His Thr Gly Val Asp Ser Leu Cys Asn
1025 1030 1035

Phe Ser Pro Leu Ala Arg Arg Val Asp Arg Val Ala Ile Tyr Glu
1040 1045 1050

Glu Phe Leu Arg Met Thr Arg Asn Gly Thr Gln Leu Gln Asn Phe
1055 1060 1065

Thr Leu Asp Arg Ser Ser Val Leu Val Asp Gly Tyr Ser Pro Asn

1070 1075 1080

Arg Asn Glu Pro Leu Thr Gly Asn Ser Asp Leu Pro Phe Trp Ala
1085 1090 1095

Val Ile Phe Ile Gly Leu Ala Gly Leu Leu Gly Leu Ile Thr Cys
1100 1105 1110

Leu Ile Cys Gly Val Leu Val Thr Thr Arg Arg Arg Lys Lys Glu
1115 1120 1125

Gly Glu Tyr Asn Val Gln Gln Gln Cys Pro Gly Tyr Tyr Gln Ser
1130 1135 1140

His Leu Asp Leu Glu Asp Leu Gln
1145

<210> 130
<211> 564
<212> PRT
<213> Homo sapiens

<400> 130
Met Gly Asn Trp Val Val Asn His Trp Phe Ser Val Leu Phe Leu
1 5 10 15

Val Val Trp Leu Gly Leu Asn Val Phe Leu Phe Val Asp Ala Phe
20 25 30

Leu Lys Tyr Glu Lys Ala Asp Lys Tyr Tyr Tyr Thr Arg Lys Ile
35 40 45

Leu Gly Ser Thr Leu Ala Cys Ala Arg Ala Ser Ala Leu Cys Leu
50 55 60

Asn Phe Asn Ser Thr Leu Ile Leu Leu Pro Val Cys Arg Asn Leu
65 70 75

Leu Ser Phe Leu Arg Gly Thr Cys Ser Phe Cys Ser Arg Thr Leu
80 85 90

Arg Lys Gln Leu Asp His Asn Leu Thr Phe His Lys Leu Val Ala
95 100 105

Tyr Met Ile Cys Leu His Thr Ala Ile His Ile Ile Ala His Leu
110 115 120

Phe Asn Phe Asp Cys Tyr Ser Arg Ser Arg Gln Ala Thr Asp Gly
125 130 135

Ser Leu Ala Ser Ile Leu Ser Ser Leu Ser His Asp Glu Lys Lys
140 145 150

Gly Gly Ser Trp Leu Asn Pro Ile Gln Ser Arg Asn Thr Thr Val
155 160 165

Glu Tyr Val Thr Phe Thr Ser Val Ala Gly Leu Thr Gly Val Ile
170 175 180

Met Thr Ile Ala Leu Ile Leu Met Val Thr Ser Ala Thr Glu Phe
185 190 195

Ile Arg Arg Ser Tyr Phe Glu Val Phe Trp Tyr Thr His His Leu
200 205 210

Phe Ile Phe Tyr Ile Leu Gly Leu Gly Ile His Gly Ile Gly Gly
215 220 225

Ile Val Arg Gly Gln Thr Glu Glu Ser Met Asn Glu Ser His Pro
230 235 240

Arg Lys Cys Ala Glu Ser Phe Glu Met Trp Asp Asp Arg Asp Ser
245 250 255

His Cys Arg Arg Pro Lys Phe Glu Gly His Pro Pro Glu Ser Trp
260 265 270

1528
1549

Lys Trp Ile Leu Ala Pro Val Ile Leu Tyr Ile Cys Glu Arg Ile
275 280 285

Leu Arg Phe Tyr Arg Ser Gln Gln Lys Val Val Ile Thr Lys Val
290 295 300

Val Met His Pro Ser Lys Val Leu Glu Leu Gln Met Asn Lys Arg
305 310 315

Gly Phe Ser Met Glu Val Gly Gln Tyr Ile Phe Val Asn Cys Pro
320 325 330

Ser Ile Ser Leu Leu Glu Trp His Pro Phe Thr Leu Thr Ser Ala
335 340 345

Pro Glu Glu Asp Phe Phe Ser Ile His Ile Arg Ala Ala Gly Asp
350 355 360

Trp Thr Glu Asn Leu Ile Arg Ala Phe Glu Gln Gln Tyr Ser Pro
365 370 375

Ile Pro Arg Ile Glu Val Asp Gly Pro Phe Gly Thr Ala Ser Glu
380 385 390

Asp Val Phe Gln Tyr Glu Val Ala Val Leu Val Gly Ala Gly Ile
395 400 405

Gly Val Thr Pro Phe Ala Ser Ile Leu Lys Ser Ile Trp Tyr Lys
410 415 420

Phe Gln Cys Ala Asp His Asn Leu Lys Thr Lys Lys Ile Tyr Phe
425 430 435

Tyr Trp Ile Cys Arg Glu Thr Gly Ala Phe Ser Trp Phe Asn Asn
440 445 450

Leu Leu Thr Ser Leu Glu Gln Glu Met Glu Glu Leu Gly Lys Val
455 460 465

Gly Phe Leu Asn Tyr Arg Leu Phe Leu Thr Gly Trp Asp Ser Asn
470 475 480

Ile Val Gly His Ala Ala Leu Asn Phe Asp Lys Ala Thr Asp Ile
485 490 495

Val Thr Gly Leu Lys Gln Lys Thr Ser Phe Gly Arg Pro Met Trp
500 505 510

Asp Asn Glu Phe Ser Thr Ile Ala Thr Ser His Pro Lys Ser Val
515 520 525

Val Gly Val Phe Leu Cys Gly Pro Arg Thr Leu Ala Lys Ser Leu
530 535 540

Arg Lys Cys Cys His Arg Tyr Ser Ser Leu Asp Pro Arg Lys Val
545 550 555

Gln Phe Tyr Phe Asn Lys Glu Asn Phe
560

<210> 131
<211> 564
<212> PRT
<213> Homo sapiens

<400> 131
Met Gly Asn Trp Val Val Asn His Trp Phe Ser Val Leu Phe Leu
1 5 10 15

Val Val Trp Leu Gly Leu Asn Val Phe Leu Phe Val Asp Ala Phe
20 25 30

Leu Lys Tyr Glu Lys Ala Asp Lys Tyr Tyr Tyr Thr Arg Lys Ile
35 40 45

1550
1590

Leu Gly Ser Thr Leu Ala Cys Ala Arg Ala Ser Ala Leu Cys Leu
50 55 60

Asn Phe Asn Ser Thr Leu Ile Leu Leu Pro Val Cys Arg Asn Leu
65 70 75

Leu Ser Phe Leu Arg Gly Thr Cys Ser Phe Cys Ser Arg Thr Leu
80 85 90

Arg Lys Gln Leu Asp His Asn Leu Thr Phe His Lys Leu Val Ala
95 100 105

Tyr Met Ile Cys Leu His Thr Ala Ile His Ile Ile Ala His Leu
110 115 120

Phe Asn Phe Asp Cys Tyr Ser Arg Ser Arg Gln Ala Thr Asp Gly
125 130 135

Ser Leu Ala Ser Ile Leu Ser Ser Leu Ser His Asp Glu Lys Lys
140 145 150

Gly Gly Ser Trp Leu Asn Pro Ile Gln Ser Arg Asn Thr Thr Val
155 160 165

Glu Tyr Val Thr Phe Thr Ser Ile Ala Gly Leu Thr Gly Val Ile
170 175 180

Met Thr Ile Ala Leu Ile Leu Met Val Thr Ser Ala Thr Glu Phe
185 190 195

Ile Arg Arg Ser Tyr Phe Glu Val Phe Trp Tyr Thr His His Leu
200 205 210

Phe Ile Phe Tyr Ile Leu Gly Leu Gly Ile His Gly Ile Gly Gly
215 220 225

Ile Val Arg Gly Gln Thr Glu Glu Ser Met Asn Glu Ser His Pro
230 235 240

Arg Lys Cys Ala Glu Ser Phe Glu Met Trp Asp Asp Arg Asp Ser
245 250 255

His Cys Arg Arg Pro Lys Phe Glu Gly His Pro Pro Glu Ser Trp
260 265 270

Lys Trp Ile Leu Ala Pro Val Ile Leu Tyr Ile Cys Glu Arg Ile
275 280 285

Leu Arg Phe Tyr Arg Ser Gln Gln Lys Val Val Ile Thr Lys Val
290 295 300

Val Met His Pro Ser Lys Val Leu Glu Leu Gln Met Asn Lys Arg
305 310 315

Gly Phe Ser Met Glu Val Gly Gln Tyr Ile Phe Val Asn Cys Pro
320 325 330

Ser Ile Ser Leu Leu Glu Trp His Pro Phe Thr Leu Thr Ser Ala
335 340 345

Pro Glu Glu Asp Phe Phe Ser Ile His Ile Arg Ala Ala Gly Asp
350 355 360

Trp Thr Glu Asn Leu Ile Arg Ala Phe Glu Gln Gln Tyr Ser Pro
365 370 375

Ile Pro Arg Ile Glu Val Asp Gly Pro Phe Gly Thr Ala Ser Glu
380 385 390

Asp Val Phe Gln Tyr Glu Val Ala Val Leu Val Gly Ala Gly Ile
395 400 405

Gly Val Thr Pro Phe Ala Ser Ile Leu Lys Ser Ile Trp Tyr Lys
410 415 420

Phe Gln Cys Ala Asp His Asn Leu Lys Thr Lys Lys Ile Tyr Phe

425 430 435

Tyr Trp Ile Cys Arg Glu Thr Gly Ala Phe Ser Trp Phe Asn Asn
440 445 450

Leu Leu Thr Ser Leu Glu Gln Glu Met Glu Glu Leu Gly Lys Val
455 460 465

Gly Phe Leu Asn Tyr Arg Leu Phe Leu Thr Gly Trp Asp Ser Asn
470 475 480

Ile Val Gly His Ala Ala Leu Asn Phe Asp Lys Ala Thr Asp Ile
485 490 495

Val Thr Gly Leu Lys Gln Lys Thr Ser Phe Gly Arg Pro Met Trp
500 505 510

Asp Asn Glu Phe Ser Thr Ile Ala Thr Ser His Pro Lys Ser Val
515 520 525

Val Gly Val Phe Leu Cys Gly Pro Arg Thr Leu Ala Lys Ser Leu
530 535 540

Arg Lys Cys Cys His Arg Tyr Ser Ser Leu Asp Pro Arg Lys Val
545 550 555

Gln Phe Tyr Phe Asn Lys Glu Asn Phe
560

<210> 132
<211> 294
<212> PRT
<213> Homo sapiens

<400> 132
Met Gly Asn Trp Val Val Asn His Trp Phe Ser Val Leu Phe Leu
1 5 10 15

Val Val Trp Leu Gly Leu Asn Val Phe Leu Phe Val Asp Ala Phe
20 25 30

Leu Lys Tyr Glu Lys Ala Asp Lys Tyr Tyr Tyr Thr Arg Lys Ile
35 40 45

Leu Gly Ser Trp Lys Trp Ile Leu Ala Pro Val Ile Leu Tyr Ile
50 55 60

Cys Glu Arg Ile Leu Arg Phe Tyr Arg Ser Gln Gln Lys Val Val
65 70 75

Ile Thr Lys Val Val Met His Pro Ser Lys Val Leu Glu Leu Gln
80 85 90

Met Asn Lys Arg Gly Phe Ser Met Glu Val Gly Gln Tyr Ile Phe
95 100 105

Val Asn Cys Pro Ser Ile Ser Leu Leu Glu Trp His Pro Phe Thr
110 115 120

Leu Thr Ser Ala Pro Glu Glu Asp Phe Phe Ser Ile His Ile Arg
125 130 135

Ala Ala Gly Asp Trp Thr Glu Asn Leu Ile Arg Ala Phe Glu Gln
140 145 150

Gln Tyr Ser Pro Ile Pro Arg Ile Glu Val Asp Gly Pro Phe Gly
155 160 165

Thr Ala Ser Glu Asp Val Phe Gln Tyr Glu Val Ala Val Leu Val
170 175 180

Gly Ala Gly Ile Gly Val Thr Pro Phe Ala Ser Ile Leu Lys Ser
185 190 195

Ile Trp Tyr Lys Phe Gln Cys Ala Asp His Asn Leu Lys Thr Lys
200 205 210

Lys Val Gly His Ala Ala Leu Asn Phe Asp Lys Ala Thr Asp Ile
215 220 225

Val Thr Gly Leu Lys Gln Lys Thr Ser Phe Gly Arg Pro Met Trp
230 235 240

Asp Asn Glu Phe Ser Thr Ile Ala Thr Ser His Pro Lys Ser Val
245 250 255

Val Gly Val Phe Leu Cys Gly Pro Arg Thr Leu Ala Lys Ser Leu
260 265 270

Arg Lys Cys Cys His Arg Tyr Ser Ser Leu Asp Pro Arg Lys Val
275 280 285

Gln Phe Tyr Phe Asn Lys Glu Asn Phe
290

<210> 133
<211> 515
<212> PRT
<213> Homo sapiens

<400> 133
Met Gly Asn Trp Val Val Asn His Trp Phe Ser Val Leu Phe Leu
1 5 10 15

Val Val Trp Leu Gly Leu Asn Val Phe Leu Phe Val Asp Ala Phe
20 25 30

Leu Lys Tyr Glu Lys Ala Asp Lys Tyr Tyr Tyr Thr Arg Lys Ile
35 40 45

Leu Gly Ser Thr Leu Ala Cys Ala Arg Ala Ser Ala Leu Cys Leu
50 55 60

Asn Phe Asn Ser Thr Leu Ile Leu Leu Pro Val Cys Arg Asn Leu
65 70 75

Leu Ser Phe Leu Arg Gly Thr Cys Ser Phe Cys Ser Arg Thr Leu
80 85 90

Arg Lys Gln Leu Asp His Asn Leu Thr Phe His Lys Leu Val Ala
95 100 105

Tyr Met Ile Cys Leu His Thr Ala Ile His Ile Ile Ala His Leu
110 115 120

Phe Asn Phe Asp Cys Tyr Ser Arg Ser Arg Gln Ala Thr Asp Gly
125 130 135

Ser Leu Ala Ser Ile Leu Ser Ser Leu Ser His Asp Glu Lys Lys
140 145 150

Gly Gly Ser Trp Leu Asn Pro Ile Gln Ser Arg Asn Thr Thr Val
155 160 165

Glu Tyr Val Thr Phe Thr Ser Ile Ala Gly Leu Thr Gly Val Ile
170 175 180

Met Thr Ile Ala Leu Ile Leu Met Val Thr Ser Ala Thr Glu Phe
185 190 195

Ile Arg Arg Ser Tyr Phe Glu Val Phe Trp Tyr Thr His His Leu
200 205 210

Phe Ile Phe Tyr Ile Leu Gly Leu Gly Ile His Gly Ile Gly Gly
215 220 225

Ile Val Arg Gly Gln Thr Glu Glu Ser Met Asn Glu Ser His Pro
230 235 240

Arg Lys Cys Ala Glu Ser Phe Glu Met Trp Asp Asp Arg Asp Ser
245 250 255

1553 ISHC

```
His Cys Arg Arg Pro Lys Phe Glu Gly His Pro Pro Glu Ser Trp
                260                 265                 270

Lys Trp Ile Leu Ala Pro Val Ile Leu Tyr Ile Cys Glu Arg Ile
                275                 280                 285

Leu Arg Phe Tyr Arg Ser Gln Gln Lys Val Val Ile Thr Lys Val
                290                 295                 300

Val Met His Pro Ser Lys Val Leu Glu Leu Gln Met Asn Lys Arg
                305                 310                 315

Gly Phe Ser Met Glu Val Gly Gln Tyr Ile Phe Val Asn Cys Pro
                320                 325                 330

Ser Ile Ser Leu Leu Glu Trp His Pro Phe Thr Leu Thr Ser Ala
                335                 340                 345

Pro Glu Glu Asp Phe Phe Ser Ile His Ile Arg Ala Ala Gly Asp
                350                 355                 360

Trp Thr Glu Asn Leu Ile Arg Ala Phe Glu Gln Gln Tyr Ser Pro
                365                 370                 375

Ile Pro Arg Ile Glu Val Asp Gly Pro Phe Gly Thr Ala Ser Glu
                380                 385                 390

Asp Val Phe Gln Tyr Glu Val Ala Val Leu Val Gly Ala Gly Ile
                395                 400                 405

Gly Val Thr Pro Phe Ala Ser Ile Leu Lys Ser Ile Trp Tyr Lys
                410                 415                 420

Phe Gln Cys Ala Asp His Asn Leu Lys Thr Lys Lys Val Gly His
                425                 430                 435

Ala Ala Leu Asn Phe Asp Lys Ala Thr Asp Ile Val Thr Gly Leu
                440                 445                 450

Lys Gln Lys Thr Ser Phe Gly Arg Pro Met Trp Asp Asn Glu Phe
                455                 460                 465

Ser Thr Ile Ala Thr Ser His Pro Lys Ser Val Val Gly Val Phe
                470                 475                 480

Leu Cys Gly Pro Arg Thr Leu Ala Lys Ser Leu Arg Lys Cys Cys
                485                 490                 495

His Arg Tyr Ser Ser Leu Asp Pro Arg Lys Val Gln Phe Tyr Phe
                500                 505                 510

Asn Lys Glu Asn Phe
                515
```

<210> 134
<211> 757
<212> PRT
<213> Homo sapiens

```
<400> 134
 Met Thr Ser Gly Ser Val Phe Phe Tyr Ile Leu Ile Phe Gly Lys
   1               5                  10                  15

 Tyr Phe Ser His Gly Gly Gly Gln Asp Val Lys Cys Ser Leu Gly
                20                  25                  30

 Tyr Phe Pro Cys Gly Asn Ile Thr Lys Cys Leu Pro Gln Leu Leu
                35                  40                  45

 His Cys Asn Gly Val Asp Asp Cys Gly Asn Gln Ala Asp Glu Asp
                50                  55                  60

 Asn Cys Gly Asp Asn Asn Gly Trp Ser Met Gln Phe Asp Lys Tyr
                65                  70                  75

 Phe Ala Ser Tyr Tyr Lys Met Thr Ser Gln Tyr Pro Phe Glu Ala
```

80 85 90

Glu Thr Pro Glu Cys Leu Val Gly Ser Val Pro Val Gln Cys Leu
95 100 105

Cys Gln Gly Leu Glu Leu Asp Cys Asp Glu Thr Asn Leu Arg Ala
110 115 120

Val Pro Ser Val Ser Ser Asn Val Thr Ala Met Ser Leu Gln Trp
125 130 135

Asn Leu Ile Arg Lys Leu Pro Pro Asp Cys Phe Lys Asn Tyr His
140 145 150

Asp Leu Gln Lys Leu Tyr Leu Gln Asn Asn Lys Ile Thr Ser Ile
155 160 165

Ser Ile Tyr Ala Phe Arg Gly Leu Asn Ser Leu Thr Lys Leu Tyr
170 175 180

Leu Ser His Asn Arg Ile Thr Phe Leu Lys Pro Gly Val Phe Glu
185 190 195

Asp Leu His Arg Leu Glu Trp Leu Ile Ile Glu Asp Asn His Leu
200 205 210

Ser Arg Ile Ser Pro Pro Thr Phe Tyr Gly Leu Asn Ser Leu Ile
215 220 225

Leu Leu Val Leu Met Asn Asn Val Leu Thr Arg Leu Pro Asp Lys
230 235 240

Pro Leu Cys Gln His Met Pro Arg Leu His Trp Leu Asp Leu Glu
245 250 255

Gly Asn His Ile His Asn Leu Arg Asn Leu Thr Phe Ile Ser Cys
260 265 270

Ser Asn Leu Thr Val Leu Val Met Arg Lys Asn Lys Ile Asn His
275 280 285

Leu Asn Glu Asn Thr Phe Ala Pro Leu Gln Lys Leu Asp Glu Leu
290 295 300

Asp Leu Gly Ser Asn Lys Ile Glu Asn Leu Pro Pro Leu Ile Phe
305 310 315

Lys Asp Leu Lys Glu Leu Ser Gln Leu Asn Leu Ser Tyr Asn Pro
320 325 330

Ile Gln Lys Ile Gln Ala Asn Gln Phe Asp Tyr Leu Val Lys Leu
335 340 345

Lys Ser Leu Ser Leu Glu Gly Ile Glu Ile Ser Asn Ile Gln Gln
350 355 360

Arg Met Phe Arg Pro Leu Met Asn Leu Ser His Ile Tyr Phe Lys
365 370 375

Lys Phe Gln Tyr Cys Gly Tyr Ala Pro His Val Arg Ser Cys Lys
380 385 390

Pro Asn Thr Asp Gly Ile Ser Ser Leu Glu Asn Leu Leu Ala Ser
395 400 405

Ile Ile Gln Arg Val Phe Val Trp Val Val Ser Ala Val Thr Cys
410 415 420

Phe Gly Asn Ile Phe Val Ile Cys Met Arg Pro Tyr Ile Arg Ser
425 430 435

Glu Asn Lys Leu Tyr Ala Met Ser Ile Ile Ser Leu Cys Cys Ala
440 445 450

Asp Cys Leu Met Gly Ile Tyr Leu Phe Val Ile Gly Gly Phe Asp
455 460 465

Leu Lys Phe Arg Gly Glu Tyr Asn Lys His Ala Gln Leu Trp Met
470 475 480

Glu Ser Thr His Cys Gln Leu Val Gly Ser Leu Ala Ile Leu Ser
485 490 495

Thr Glu Val Ser Val Leu Leu Leu Thr Phe Leu Thr Leu Glu Lys
500 505 510

Tyr Ile Cys Ile Val Tyr Pro Phe Arg Cys Val Arg Pro Gly Lys
515 520 525

Cys Arg Thr Ile Thr Val Leu Ile Leu Ile Trp Ile Thr Gly Phe
530 535 540

Ile Val Ala Phe Ile Pro Leu Ser Asn Lys Glu Phe Phe Lys Asn
545 550 555

Tyr Tyr Gly Thr Asn Gly Val Cys Phe Pro Leu His Ser Glu Asp
560 565 570

Thr Glu Ser Ile Gly Ala Gln Ile Tyr Ser Val Ala Ile Phe Leu
575 580 585

Gly Ile Asn Leu Ala Ala Phe Ile Ile Ile Val Phe Ser Tyr Gly
590 595 600

Ser Met Phe Tyr Ser Val His Gln Ser Ala Ile Thr Ala Thr Glu
605 610 615

Ile Arg Asn Gln Val Lys Lys Glu Met Ile Leu Ala Lys Arg Phe
620 625 630

Phe Phe Ile Val Phe Thr Asp Ala Leu Cys Trp Ile Pro Ile Phe
635 640 645

Val Val Lys Phe Leu Ser Leu Leu Gln Val Glu Ile Pro Gly Thr
650 655 660

Ile Thr Ser Trp Val Val Ile Phe Ile Leu Pro Ile Asn Ser Ala
665 670 675

Leu Asn Pro Ile Leu Tyr Thr Leu Thr Thr Arg Pro Phe Lys Glu
680 685 690

Met Ile His Arg Phe Trp Tyr Asn Tyr Arg Gln Arg Lys Ser Met
695 700 705

Asp Ser Lys Gly Gln Lys Thr Tyr Ala Pro Ser Phe Ile Trp Val
710 715 720

Glu Met Trp Pro Leu Gln Glu Met Pro Pro Glu Leu Met Lys Pro
725 730 735

Asp Leu Phe Thr Tyr Pro Cys Glu Met Ser Leu Ile Ser Gln Ser
740 745 750

Thr Arg Leu Asn Ser Tyr Ser
755

<210> 135
<211> 757
<212> PRT
<213> Homo sapiens

<400> 135
Met Thr Ser Gly Ser Val Phe Phe Tyr Ile Leu Ile Phe Gly Lys
1 5 10 15

Tyr Phe Ser His Gly Gly Gly Gln Asp Val Lys Cys Ser Leu Gly
20 25 30

Tyr Phe Pro Cys Gly Asn Ile Thr Lys Cys Leu Pro Gln Leu Leu
35 40 45

His Cys Asn Gly Val Asp Asp Cys Gly Asn Gln Ala Asp Glu Asp
50 55 60

Asn Cys Gly Asp Asn Asn Gly Trp Ser Leu Gln Phe Asp Lys Tyr
65 70 75

Phe Ala Ser Tyr Tyr Lys Met Thr Ser Gln Tyr Pro Phe Glu Ala
80 85 90

Glu Thr Pro Glu Cys Leu Val Gly Ser Val Pro Val Gln Cys Leu
95 100 105

Cys Gln Gly Leu Glu Leu Asp Cys Asp Glu Thr Asn Leu Arg Ala
110 115 120

Val Pro Ser Val Ser Ser Asn Val Thr Ala Met Ser Leu Gln Trp
125 130 135

Asn Leu Ile Arg Lys Leu Pro Pro Asp Cys Phe Lys Asn Tyr His
140 145 150

Asp Leu Gln Lys Leu Tyr Leu Gln Asn Asn Lys Ile Thr Ser Ile
155 160 165

Ser Ile Tyr Ala Phe Arg Gly Leu Asn Ser Leu Thr Lys Leu Tyr
170 175 180

Leu Ser His Asn Arg Ile Thr Phe Leu Lys Pro Gly Val Phe Glu
185 190 195

Asp Leu His Arg Leu Glu Trp Leu Ile Ile Glu Asp Asn His Leu
200 205 210

Ser Arg Ile Ser Pro Pro Thr Phe Tyr Gly Leu Asn Ser Leu Ile
215 220 225

Leu Leu Val Leu Met Asn Asn Val Leu Thr Arg Leu Pro Asp Lys
230 235 240

Pro Leu Cys Gln His Met Pro Arg Leu His Trp Leu Asp Leu Glu
245 250 255

Gly Asn His Ile His Asn Leu Arg Asn Leu Thr Phe Ile Ser Cys
260 265 270

Ser Asn Leu Thr Val Leu Val Met Arg Lys Asn Lys Ile Asn His
275 280 285

Leu Asn Glu Asn Thr Phe Ala Pro Leu Gln Lys Leu Asp Glu Leu
290 295 300

Asp Leu Gly Ser Asn Lys Ile Glu Asn Leu Pro Pro Leu Ile Phe
305 310 315

Lys Asp Leu Lys Glu Leu Ser Gln Leu Asn Leu Ser Tyr Asn Pro
320 325 330

Ile Gln Lys Ile Gln Ala Asn Gln Phe Asp Tyr Leu Val Lys Leu
335 340 345

Lys Ser Leu Ser Leu Glu Gly Ile Glu Ile Ser Asn Ile Gln Gln
350 355 360

Arg Met Phe Arg Pro Leu Met Asn Leu Ser His Ile Tyr Phe Lys
365 370 375

Lys Phe Gln Tyr Cys Gly Tyr Ala Pro His Val Arg Ser Cys Lys
380 385 390

Pro Asn Thr Asp Gly Ile Ser Ser Leu Glu Asn Leu Leu Ala Ser
395 400 405

Ile Ile Gln Arg Val Phe Val Trp Val Val Ser Ala Val Thr Cys
410 415 420

Phe Gly Asn Ile Phe Val Ile Cys Met Arg Pro Tyr Ile Arg Ser

425 430 435

Glu Asn Lys Leu Tyr Ala Met Ser Ile Ile Ser Leu Cys Cys Ala
440 445 450

Asp Cys Leu Met Gly Ile Tyr Leu Phe Val Ile Gly Gly Phe Asp
455 460 465

Leu Lys Phe Arg Gly Glu Tyr Asn Lys His Ala Gln Leu Trp Met
470 475 480

Glu Ser Thr His Cys Gln Leu Val Gly Ser Leu Ala Ile Leu Ser
485 490 495

Thr Glu Val Ser Val Leu Leu Leu Thr Phe Leu Thr Leu Glu Lys
500 505 510

Tyr Ile Cys Ile Val Tyr Pro Phe Arg Cys Val Arg Pro Gly Lys
515 520 525

Cys Arg Thr Ile Thr Val Leu Ile Leu Ile Trp Ile Thr Gly Phe
530 535 540

Ile Val Ala Phe Ile Pro Leu Ser Asn Lys Glu Phe Phe Lys Asn
545 550 555

Tyr Tyr Gly Thr Asn Gly Val Cys Phe Pro Leu His Ser Glu Asp
560 565 570

Thr Glu Ser Ile Gly Ala Gln Ile Tyr Ser Val Ala Ile Phe Leu
575 580 585

Gly Ile Asn Leu Ala Ala Phe Ile Ile Ile Val Phe Ser Tyr Gly
590 595 600

Ser Met Phe Tyr Ser Val His Gln Ser Ala Ile Thr Ala Thr Glu
605 610 615

Ile Arg Asn Gln Val Lys Lys Glu Met Ile Leu Ala Lys Arg Phe
620 625 630

Phe Phe Ile Val Phe Thr Asp Ala Leu Cys Trp Ile Pro Ile Phe
635 640 645

Val Val Lys Phe Leu Ser Leu Leu Gln Val Glu Ile Pro Gly Thr
650 655 660

Ile Thr Ser Trp Val Val Ile Phe Ile Leu Pro Ile Asn Ser Ala
665 670 675

Leu Asn Pro Ile Leu Tyr Thr Leu Thr Thr Arg Pro Phe Lys Glu
680 685 690

Met Ile His Arg Phe Trp Tyr Asn Tyr Arg Gln Arg Lys Ser Met
695 700 705

Asp Ser Lys Gly Gln Lys Thr Tyr Ala Pro Ser Phe Ile Trp Val
710 715 720

Glu Met Trp Pro Leu Gln Glu Met Pro Pro Glu Leu Met Lys Pro
725 730 735

Asp Leu Phe Thr Tyr Pro Cys Glu Met Ser Leu Ile Ser Gln Ser
740 745 750

Thr Arg Leu Asn Ser Tyr Ser
755

<210> 136
<211> 879
<212> PRT
<213> Homo sapiens

<400> 136
Met Gly Arg Leu Ala Ser Arg Pro Leu Leu Leu Ala Leu Leu Ser
1 5 10 15

Leu Ala Leu Cys Arg Gly Arg Val Val Arg Val Pro Thr Ala Thr
20 25 30

Leu Val Arg Val Val Gly Thr Glu Leu Val Ile Pro Cys Asn Val
35 40 45

Ser Asp Tyr Asp Gly Pro Ser Glu Gln Asn Phe Asp Trp Ser Phe
50 55 60

Ser Ser Leu Gly Ser Ser Phe Val Glu Leu Ala Ser Thr Trp Glu
65 70 75

Val Gly Phe Pro Ala Gln Leu Tyr Gln Glu Arg Leu Gln Arg Gly
80 85 90

Glu Ile Leu Leu Arg Arg Thr Ala Asn Asp Ala Val Glu Leu His
95 100 105

Ile Lys Asn Val Gln Pro Ser Asp Gln Gly His Tyr Lys Cys Ser
110 115 120

Thr Pro Ser Thr Asp Ala Thr Val Gln Gly Asn Tyr Glu Asp Thr
125 130 135

Val Gln Val Lys Val Leu Ala Asp Ser Leu His Val Gly Pro Ser
140 145 150

Ala Arg Pro Pro Pro Ser Leu Ser Leu Arg Glu Gly Glu Pro Phe
155 160 165

Glu Leu Arg Cys Thr Ala Ala Ser Ala Ser Pro Leu His Thr His
170 175 180

Leu Ala Leu Leu Trp Glu Val His Arg Gly Pro Ala Arg Arg Ser
185 190 195

Val Leu Ala Leu Thr His Glu Gly Arg Phe His Pro Gly Leu Gly
200 205 210

Tyr Glu Gln Arg Tyr His Ser Gly Asp Val Arg Leu Asp Thr Val
215 220 225

Gly Ser Asp Ala Tyr Arg Leu Ser Val Ser Arg Ala Leu Ser Ala
230 235 240

Asp Gln Gly Ser Tyr Arg Cys Ile Val Ser Glu Trp Ile Ala Glu
245 250 255

Gln Gly Asn Trp Gln Glu Ile Gln Glu Lys Ala Val Glu Val Ala
260 265 270

Thr Val Val Ile Gln Pro Thr Val Leu Arg Ala Ala Val Pro Lys
275 280 285

Asn Val Ser Val Ala Glu Gly Lys Glu Leu Asp Leu Thr Cys Asn
290 295 300

Ile Thr Thr Asp Arg Ala Asp Asp Val Arg Pro Glu Val Thr Trp
305 310 315

Ser Phe Ser Arg Met Pro Asp Ser Thr Leu Pro Gly Ser Arg Val
320 325 330

Leu Ala Arg Leu Asp Arg Asp Ser Leu Val His Ser Ser Pro His
335 340 345

Val Ala Leu Ser His Val Asp Ala Arg Ser Tyr His Leu Leu Val
350 355 360

Arg Asp Val Ser Lys Glu Asn Ser Gly Tyr Tyr Tyr Cys His Val
365 370 375

Ser Leu Trp Ala Pro Gly His Asn Arg Ser Trp His Lys Val Ala
380 385 390

Glu Ala Val Ser Ser Pro Ala Gly Val Gly Val Thr Trp Leu Glu
395 400 405

Pro Asp Tyr Gln Val Tyr Leu Asn Ala Ser Lys Val Pro Gly Phe
410 415 420

Ala Asp Asp Pro Thr Glu Leu Ala Cys Arg Val Val Asp Thr Lys
425 430 435

Ser Gly Glu Ala Asn Val Arg Phe Thr Val Ser Trp Tyr Tyr Arg
440 445 450

Met Asn Arg Arg Ser Asp Asn Val Val Thr Ser Glu Leu Leu Ala
455 460 465

Val Met Asp Gly Asp Trp Thr Leu Lys Tyr Gly Glu Arg Ser Lys
470 475 480

Gln Arg Ala Gln Asp Gly Asp Phe Ile Phe Ser Lys Glu His Thr
485 490 495

Asp Thr Phe Asn Phe Arg Ile Gln Arg Thr Thr Glu Glu Asp Arg
500 505 510

Gly Asn Tyr Tyr Cys Val Val Ser Ala Trp Thr Lys Gln Arg Asn
515 520 525

Asn Ser Trp Val Lys Ser Lys Asp Val Phe Ser Lys Pro Val Asn
530 535 540

Ile Phe Trp Ala Leu Glu Asp Ser Val Leu Val Val Lys Ala Arg
545 550 555

Gln Pro Lys Pro Phe Phe Ala Ala Gly Asn Thr Phe Glu Met Thr
560 565 570

Cys Lys Val Ser Ser Lys Asn Ile Lys Ser Pro Arg Tyr Ser Val
575 580 585

Leu Ile Met Ala Glu Lys Pro Val Gly Asp Leu Ser Ser Pro Asn
590 595 600

Glu Thr Lys Tyr Ile Ile Ser Leu Asp Gln Asp Ser Val Val Lys
605 610 615

Leu Glu Asn Trp Thr Asp Ala Ser Arg Val Asp Gly Val Val Leu
620 625 630

Glu Lys Val Gln Glu Asp Glu Phe Arg Tyr Arg Met Tyr Gln Thr
635 640 645

Gln Val Ser Asp Ala Gly Leu Tyr Arg Cys Met Val Thr Ala Trp
650 655 660

Ser Pro Val Arg Gly Ser Leu Trp Arg Glu Ala Ala Thr Ser Leu
665 670 675

Ser Asn Pro Ile Glu Ile Asp Phe Gln Thr Ser Gly Pro Ile Phe
680 685 690

Asn Ala Ser Val His Ser Asp Thr Pro Ser Val Ile Arg Gly Asp
695 700 705

Leu Ile Lys Leu Phe Cys Ile Ile Thr Val Glu Gly Ala Ala Leu
710 715 720

Asp Pro Asp Asp Met Ala Phe Asp Val Ser Trp Phe Ala Val His
725 730 735

Ser Phe Gly Leu Asp Lys Ala Pro Val Leu Leu Ser Ser Leu Asp
740 745 750

Arg Lys Gly Ile Val Thr Thr Ser Arg Arg Asp Trp Lys Ser Asp
755 760 765

Leu Ser Leu Glu Arg Val Ser Val Leu Glu Phe Leu Leu Gln Val

770 775 780

His Gly Ser Glu Asp Gln Asp Phe Gly Asn Tyr Tyr Cys Ser Val
785 790 795

Thr Pro Trp Val Lys Ser Pro Thr Gly Ser Trp Gln Lys Glu Ala
800 805 810

Glu Ile His Ser Lys Pro Val Phe Ile Thr Val Lys Met Asp Val
815 820 825

Leu Asn Ala Phe Lys Tyr Pro Leu Leu Ile Gly Ile Gly Leu Ser
830 835 840

Thr Val Ile Gly Leu Leu Ser Cys Leu Ile Gly Tyr Cys Ser Ser
845 850 855

His Trp Cys Cys Lys Lys Glu Val Gln Glu Thr Arg Arg Glu Arg
860 865 870

Arg Arg Leu Met Ser Met Glu Met Asp
875

<210> 137
<211> 434
<212> PRT
<213> Homo sapiens

<400> 137
Met Ser Arg Ser Arg His Leu Gly Lys Ile Arg Lys Arg Leu Glu
1 5 10 15

Asp Val Lys Ser Gln Trp Val Arg Pro Ala Arg Ala Asp Phe Ser
20 25 30

Asp Asn Glu Ser Ala Arg Leu Ala Thr Asp Ala Leu Leu Asp Gly
35 40 45

Gly Ser Glu Ala Tyr Trp Arg Val Leu Ser Gln Glu Gly Glu Val
50 55 60

Asp Phe Leu Ser Ser Val Glu Ala Gln Tyr Ile Gln Ala Gln Ala
65 70 75

Arg Glu Pro Pro Cys Pro Pro Asp Thr Leu Gly Gly Ala Glu Ala
80 85 90

Gly Pro Lys Gly Leu Asp Ser Ser Ser Leu Gln Ser Gly Thr Tyr
95 100 105

Phe Pro Val Ala Ser Glu Gly Ser Glu Pro Ala Leu Leu His Ser
110 115 120

Trp Ala Ser Ala Glu Lys Pro Tyr Leu Lys Glu Lys Ser Ser Ala
125 130 135

Thr Val Tyr Phe Gln Thr Val Lys His Asn Asn Ile Arg Asp Leu
140 145 150

Val Arg Arg Cys Ile Thr Arg Thr Ser Gln Val Leu Val Ile Leu
155 160 165

Met Asp Val Phe Thr Asp Val Glu Ile Phe Cys Asp Ile Leu Glu
170 175 180

Ala Ala Asn Lys Arg Gly Val Phe Val Cys Val Leu Leu Asp Gln
185 190 195

Gly Gly Val Lys Leu Phe Gln Glu Met Cys Asp Lys Val Gln Ile
200 205 210

Ser Asp Ser His Leu Lys Asn Ile Ser Ile Arg Ser Val Glu Gly
215 220 225

Glu Ile Tyr Cys Ala Lys Ser Gly Arg Lys Phe Ala Gly Gln Ile
230 235 240

Arg Glu Lys Phe Ile Ile Ser Asp Trp Arg Phe Val Leu Ser Gly
245 250 255

Ser Tyr Ser Phe Thr Trp Leu Cys Gly His Val His Arg Asn Ile
260 265 270

Leu Ser Lys Phe Thr Gly Gln Ala Val Glu Leu Phe Asp Glu Glu
275 280 285

Phe Arg His Leu Tyr Ala Ser Ser Lys Pro Val Met Gly Leu Lys
290 295 300

Ser Pro Arg Leu Val Ala Pro Val Pro Pro Gly Ala Ala Pro Ala
305 310 315

Asn Gly Arg Leu Ser Ser Ser Ser Gly Ser Ala Ser Asp Arg Thr
320 325 330

Ser Ser Asn Pro Phe Ser Gly Arg Ser Ala Gly Ser His Pro Gly
335 340 345

Thr Arg Ser Val Ser Ala Ser Ser Gly Pro Cys Ser Pro Ala Ala
350 355 360

Pro His Pro Pro Pro Pro Pro Arg Phe Gln Pro His Gln Gly Pro
365 370 375

Trp Gly Ala Pro Ser Pro Gln Ala His Leu Ser Pro Arg Pro His
380 385 390

Asp Gly Pro Pro Ala Ala Val Tyr Ser Asn Leu Gly Ala Tyr Arg
395 400 405

Pro Thr Arg Leu Gln Leu Glu Gln Leu Gly Leu Val Pro Arg Leu
410 415 420

Thr Pro Thr Trp Arg Pro Phe Leu Gln Ala Ser Pro His Phe
425 430

<210> 138
<211> 933
<212> PRT
<213> Homo sapiens

<400> 138
Met Ala Leu Pro Gln Pro Pro Pro Pro Ala Gly Ala Arg Asp Pro
1 5 10 15

Val Thr Pro Glu His Trp Ala Ser Gly Pro Pro Ser Gln Ala Gln
20 25 30

Pro Leu Leu Arg Gln Glu Ala Lys Glu Glu Glu Glu Gly Glu Glu
35 40 45

Thr Gly Val Gln Gly Ala Trp Gly Thr Gly Thr Ala Glu Gln Arg
50 55 60

Arg Arg Gly Trp Gly Glu Ala Ala Glu Ser Ala Ala Ala Glu Glu
65 70 75

Gly Gln Ala Glu Val Gly Gly Ala Ala Ala Ala Gly Ser Gly Ser
80 85 90

Pro Ala Gly Gly Ala Gly Gly Gly Leu Gly Ser Trp Arg Pro Leu
95 100 105

Leu Ala Trp Leu Gln Arg Arg Gln Pro Gln Cys Cys Pro Cys Ala
110 115 120

Ala Pro Leu Ser Arg Ser Ala Ala His Cys Cys His Gly Gly Thr
125 130 135

Lys Met Ala Ala Leu Ala Tyr Asn Leu Gly Lys Arg Glu Ile Asn
140 145 150

His Tyr Phe Ser Val Arg Ser Ala Lys Val Leu Ala Leu Val Ala
155 160 165

Val Leu Leu Leu Ala Ala Cys His Leu Ala Ser Arg Arg Tyr Arg
170 175 180

Gly Asn Asp Ser Cys Glu Tyr Leu Leu Ser Ser Gly Arg Phe Leu
185 190 195

Gly Glu Lys Val Trp Gln Pro His Ser Cys Met Met His Lys Tyr
200 205 210

Lys Ile Ser Glu Ala Lys Asn Cys Leu Val Asp Lys His Ile Ala
215 220 225

Phe Ile Gly Asp Ser Arg Ile Arg Gln Leu Phe Tyr Ser Phe Val
230 235 240

Lys Ile Ile Asn Pro Gln Phe Lys Glu Glu Gly Asn Lys His Glu
245 250 255

Asn Ile Pro Phe Glu Asp Lys Thr Ala Ser Val Lys Val Asp Phe
260 265 270

Leu Trp His Pro Glu Val Asn Gly Ser Met Lys Gln Cys Ile Lys
275 280 285

Val Trp Thr Glu Asp Ser Ile Ala Lys Pro His Val Ile Val Ala
290 295 300

Gly Ala Ala Thr Trp Ser Ile Lys Ile His Asn Gly Ser Ser Glu
305 310 315

Ala Leu Ser Gln Tyr Lys Met Asn Ile Thr Ser Ile Ala Pro Leu
320 325 330

Leu Glu Lys Leu Ala Lys Thr Ser Asp Val Tyr Trp Val Leu Gln
335 340 345

Asp Pro Val Tyr Glu Asp Leu Leu Ser Glu Asn Arg Lys Met Ile
350 355 360

Thr Asn Glu Lys Ile Asp Ala Tyr Asn Glu Ala Ala Val Ser Ile
365 370 375

Leu Asn Ser Ser Thr Arg Asn Ser Lys Ser Asn Val Lys Met Phe
380 385 390

Ser Val Ser Lys Leu Ile Ala Gln Glu Thr Ile Met Glu Ser Leu
395 400 405

Asp Gly Leu His Leu Pro Glu Ser Ser Arg Glu Thr Thr Ala Met
410 415 420

Ile Leu Met Asn Val Tyr Cys Asn Lys Ile Leu Lys Pro Val Asp
425 430 435

Gly Ser Cys Cys Gln Pro Arg Pro Pro Val Thr Leu Ile Gln Lys
440 445 450

Leu Ala Ala Cys Phe Phe Thr Leu Ser Ile Ile Gly Tyr Leu Ile
455 460 465

Phe Tyr Ile Ile His Arg Asn Ala His Arg Lys Asn Lys Pro Cys
470 475 480

Thr Asp Leu Glu Ser Gly Glu Glu Lys Lys Asn Ile Ile Asn Thr
485 490 495

Pro Val Ser Ser Leu Glu Ile Leu Leu Gln Ser Phe Cys Lys Leu
500 505 510

Gly Leu Ile Met Ala Tyr Phe Tyr Met Cys Asp Arg Ala Asn Leu
515 520 525

Phe Met Lys Glu Asn Lys Phe Tyr Thr His Ser Ser Phe Phe Ile

1563552

530 535 540

Pro Ile Ile Tyr Ile Leu Val Leu Gly Val Phe Tyr Asn Glu Asn
545 550 555

Thr Lys Glu Thr Lys Val Leu Asn Arg Glu Gln Thr Asp Glu Trp
560 565 570

Lys Gly Trp Met Gln Leu Val Ile Leu Ile Tyr His Ile Ser Gly
575 580 585

Ala Ser Thr Phe Leu Pro Val Tyr Met His Ile Arg Val Leu Val
590 595 600

Ala Ala Tyr Leu Phe Gln Thr Gly Tyr Gly His Phe Ser Tyr Phe
605 610 615

Trp Ile Lys Gly Asp Phe Gly Ile Tyr Arg Val Cys Gln Val Leu
620 625 630

Phe Arg Leu Asn Phe Leu Val Val Val Leu Cys Ile Val Met Asp
635 640 645

Arg Pro Tyr Gln Phe Tyr Tyr Phe Val Pro Leu Val Thr Val Trp
650 655 660

Phe Met Val Ile Tyr Val Thr Leu Ala Leu Trp Pro Gln Ile Ile
665 670 675

Gln Lys Lys Ala Asn Gly Asn Cys Phe Trp His Phe Gly Leu Leu
680 685 690

Leu Lys Leu Gly Phe Leu Leu Leu Phe Ile Cys Phe Leu Ala Tyr
695 700 705

Ser Gln Gly Ala Phe Glu Lys Ile Phe Ser Leu Trp Pro Leu Ser
710 715 720

Lys Cys Phe Glu Leu Lys Gly Asn Val Tyr Glu Trp Trp Phe Arg
725 730 735

Trp Arg Leu Asp Arg Tyr Val Val Phe His Gly Met Leu Phe Ala
740 745 750

Phe Ile Tyr Leu Ala Leu Gln Lys Arg Gln Ile Leu Ser Glu Gly
755 760 765

Lys Gly Glu Pro Leu Phe Ser Asn Lys Ile Ser Asn Phe Leu Leu
770 775 780

Phe Ile Ser Val Val Ser Phe Leu Thr Tyr Ser Ile Trp Ala Ser
785 790 795

Ser Cys Lys Asn Lys Ala Glu Cys Asn Glu Leu His Pro Ser Val
800 805 810

Ser Val Val Gln Ile Leu Ala Phe Ile Leu Ile Arg Asn Ile Pro
815 820 825

Gly Tyr Ala Arg Ser Val Tyr Ser Ser Phe Phe Ala Trp Phe Gly
830 835 840

Lys Ile Ser Leu Glu Leu Phe Ile Cys Gln Tyr His Ile Trp Leu
845 850 855

Ala Ala Asp Thr Arg Gly Ile Leu Val Leu Ile Pro Gly Asn Pro
860 865 870

Met Leu Asn Ile Ile Val Ser Thr Phe Ile Phe Val Cys Val Ala
875 880 885

His Glu Ile Ser Gln Ile Thr Asn Asp Leu Ala Gln Ile Ile Ile
890 895 900

Pro Lys Asp Asn Ser Ser Leu Leu Lys Arg Leu Ala Cys Ile Ala
905 910 915

Ala Phe Phe Cys Gly Leu Leu Ile Leu Ser Ser Ile Gln Asp Lys
920 925 930

Ser Lys His

<210> 139
<211> 320
<212> PRT
<213> Homo sapiens

<400> 139
Met Tyr His Cys His Ser Gly Ser Lys Pro Thr Glu Lys Gly Ala
1 5 10 15

Asn Glu Tyr Ala Tyr Ala Lys Trp Lys Leu Cys Ser Ala Ser Ala
20 25 30

Ile Cys Phe Ile Phe Met Ile Ala Glu Val Val Gly Gly His Ile
35 40 45

Ala Gly Ser Leu Ala Val Val Thr Asp Ala Ala His Leu Leu Ile
50 55 60

Asp Leu Thr Ser Phe Leu Leu Ser Leu Phe Ser Leu Trp Leu Ser
65 70 75

Ser Lys Pro Pro Ser Lys Arg Leu Thr Phe Gly Trp His Arg Ala
80 85 90

Glu Ile Leu Gly Ala Leu Leu Ser Ile Leu Cys Ile Trp Val Val
95 100 105

Thr Gly Val Leu Val Tyr Leu Ala Cys Glu Arg Leu Leu Tyr Pro
110 115 120

Asp Tyr Gln Ile Gln Ala Thr Val Met Ile Ile Val Ser Ser Cys
125 130 135

Ala Val Ala Ala Asn Ile Val Leu Thr Val Val Leu His Gln Arg
140 145 150

Cys Leu Gly His Asn His Lys Glu Val Gln Ala Asn Ala Ser Val
155 160 165

Arg Ala Ala Phe Val His Ala Leu Gly Asp Leu Phe Gln Ser Ile
170 175 180

Ser Val Leu Ile Ser Ala Leu Ile Ile Tyr Phe Lys Pro Glu Tyr
185 190 195

Lys Ile Ala Asp Pro Ile Cys Thr Phe Ile Phe Ser Ile Leu Val
200 205 210

Leu Ala Ser Thr Ile Thr Ile Leu Lys Asp Phe Ser Ile Leu Leu
215 220 225

Met Glu Gly Val Pro Lys Ser Leu Asn Tyr Ser Gly Val Lys Glu
230 235 240

Leu Ile Leu Ala Val Asp Gly Val Leu Ser Val His Ser Leu His
245 250 255

Ile Trp Ser Leu Thr Met Asn Gln Val Ile Leu Ser Ala His Val
260 265 270

Ala Thr Ala Ala Ser Arg Asp Ser Gln Val Val Arg Arg Glu Ile
275 280 285

Ala Lys Ala Leu Ser Lys Ser Phe Thr Met His Ser Leu Thr Ile
290 295 300

Gln Met Glu Ser Pro Val Asp Gln Asp Pro Asp Cys Leu Phe Cys
305 310 315

Glu Asp Pro Cys Asp
320

<210> 140
<211> 426
<212> PRT
<213> Homo sapiens

<400> 140
Met Val Leu Ser Val Pro Val Ile Ala Leu Gly Ala Thr Leu Gly
1 5 10 15

Thr Ala Thr Ser Ile Leu Ala Leu Cys Gly Val Thr Cys Leu Cys
20 25 30

Arg His Met His Pro Lys Lys Gly Leu Leu Pro Arg Asp Gln Asp
35 40 45

Pro Asp Leu Glu Lys Ala Lys Pro Ser Leu Leu Gly Ser Ala Gln
50 55 60

Gln Phe Asn Val Lys Lys Ser Thr Glu Pro Val Gln Pro Arg Ala
65 70 75

Leu Leu Lys Phe Pro Asp Ile Tyr Gly Pro Arg Pro Ala Val Thr
80 85 90

Ala Pro Glu Val Ile Asn Tyr Ala Asp Tyr Ser Leu Arg Ser Thr
95 100 105

Glu Glu Pro Thr Ala Pro Ala Ser Pro Gln Pro Pro Asn Asp Ser
110 115 120

Arg Leu Lys Arg Gln Val Thr Glu Glu Leu Phe Ile Leu Pro Gln
125 130 135

Asn Gly Val Val Glu Asp Val Cys Val Met Glu Thr Trp Asn Pro
140 145 150

Glu Lys Ala Ala Ser Trp Asn Gln Ala Pro Lys Leu His Tyr Cys
155 160 165

Leu Asp Tyr Asp Cys Gln Lys Ala Glu Leu Phe Val Thr Arg Leu
170 175 180

Glu Ala Val Thr Ser Asn His Asp Gly Gly Cys Asp Cys Tyr Val
185 190 195

Gln Gly Ser Val Ala Asn Arg Thr Gly Ser Val Glu Ala Gln Thr
200 205 210

Ala Leu Lys Lys Arg Gln Leu His Thr Thr Trp Glu Glu Gly Leu
215 220 225

Val Leu Pro Leu Ala Glu Glu Glu Leu Pro Thr Ala Thr Leu Thr
230 235 240

Leu Thr Leu Arg Thr Cys Asp Arg Phe Ser Arg His Ser Val Ala
245 250 255

Gly Glu Leu Arg Leu Gly Leu Asp Gly Thr Ser Val Pro Leu Gly
260 265 270

Ala Ala Gln Trp Gly Glu Leu Lys Thr Ser Ala Lys Glu Pro Ser
275 280 285

Ala Gly Ala Gly Glu Val Leu Leu Ser Ile Ser Tyr Leu Pro Ala
290 295 300

Ala Asn Arg Leu Leu Val Val Leu Ile Lys Ala Lys Asn Leu His
305 310 315

Ser Asn Gln Ser Lys Glu Leu Leu Gly Lys Asp Val Ser Val Lys
320 325 330

Val Thr Leu Lys His Gln Ala Arg Lys Leu Lys Lys Lys Gln Thr

335 340 345

Lys Arg Ala Lys His Lys Ile Asn Pro Val Trp Asn Glu Met Ile
350 355 360

Met Phe Glu Leu Pro Asp Asp Leu Leu Gln Ala Ser Ser Val Glu
365 370 375

Leu Glu Val Leu Gly Gln Asp Asp Ser Gly Gln Ser Cys Ala Leu
380 385 390

Gly His Cys Ser Leu Gly Leu His Thr Ser Gly Ser Glu Arg Ser
395 400 405

His Trp Glu Glu Met Leu Lys Asn Pro Arg Arg Gln Ile Ala Met
410 415 420

Trp His Gln Leu His Leu
425

<210> 141
<211> 883
<212> PRT
<213> Homo sapiens

<400> 141
Met Gln Lys Ile Met His Ile Ser Val Leu Leu Ser Pro Val Leu
1 5 10 15

Trp Gly Leu Ile Phe Gly Val Ser Ser Asn Ser Ile Gln Ile Gly
20 25 30

Gly Leu Phe Pro Arg Gly Ala Asp Gln Glu Tyr Ser Ala Phe Arg
35 40 45

Val Gly Met Val Gln Phe Ser Thr Ser Glu Phe Arg Leu Thr Pro
50 55 60

His Ile Asp Asn Leu Glu Val Ala Asn Ser Phe Ala Val Thr Asn
65 70 75

Ala Phe Cys Ser Gln Phe Ser Arg Gly Val Tyr Ala Ile Phe Gly
80 85 90

Phe Tyr Asp Lys Lys Ser Val Asn Thr Ile Thr Ser Phe Cys Gly
95 100 105

Thr Leu His Val Ser Phe Ile Thr Pro Ser Phe Pro Thr Asp Gly
110 115 120

Thr His Pro Phe Val Ile Gln Met Arg Pro Asp Leu Lys Gly Ala
125 130 135

Leu Leu Ser Leu Ile Glu Tyr Tyr Gln Trp Asp Lys Phe Ala Tyr
140 145 150

Leu Tyr Asp Ser Asp Arg Gly Leu Ser Thr Leu Gln Ala Val Leu
155 160 165

Asp Ser Ala Ala Glu Lys Lys Trp Gln Val Thr Ala Ile Asn Val
170 175 180

Gly Asn Ile Asn Asn Asp Lys Lys Asp Glu Met Tyr Arg Ser Leu
185 190 195

Phe Gln Asp Leu Glu Leu Lys Lys Glu Arg Arg Val Ile Leu Asp
200 205 210

Cys Glu Arg Asp Lys Val Asn Asp Ile Val Asp Gln Val Ile Thr
215 220 225

Ile Gly Lys His Val Lys Gly Tyr His Tyr Ile Ile Ala Asn Leu
230 235 240

Glu Phe Thr Asp Gly Asp Leu Leu Lys Ile Gln Phe Gly Gly Ala
245 250 255

1567 X 956

Asn Val Ser Gly Phe Gln Ile Val Asp Tyr Asp Asp Ser Leu Val
260 265 270

Ser Lys Phe Ile Glu Arg Trp Ser Thr Leu Glu Glu Lys Glu Tyr
275 280 285

Pro Gly Ala His Thr Thr Thr Ile Lys Tyr Thr Ser Ala Leu Thr
290 295 300

Tyr Asp Ala Val Gln Val Met Thr Glu Ala Phe Arg Asn Leu Arg
305 310 315

Lys Gln Arg Ile Glu Ile Ser Arg Arg Gly Asn Ala Gly Asp Cys
320 325 330

Leu Ala Asn Pro Ala Val Pro Trp Gly Gln Gly Val Glu Ile Glu
335 340 345

Arg Ala Leu Lys Gln Val Gln Val Glu Gly Leu Ser Gly Asn Ile
350 355 360

Lys Phe Asp Gln Asn Gly Lys Arg Ile Asn Tyr Thr Ile Asn Ile
365 370 375

Met Glu Leu Lys Thr Asn Gly Pro Arg Lys Ile Gly Tyr Trp Ser
380 385 390

Glu Val Asp Lys Met Val Val Thr Leu Thr Glu Leu Pro Ser Gly
395 400 405

Asn Asp Thr Ser Gly Leu Glu Asn Lys Thr Val Val Val Thr Thr
410 415 420

Ile Leu Glu Ser Pro Tyr Val Met Met Lys Lys Asn His Glu Met
425 430 435

Leu Glu Gly Asn Glu Arg Tyr Glu Gly Tyr Cys Val Asp Leu Ala
440 445 450

Ala Glu Ile Ala Lys His Cys Gly Phe Lys Tyr Lys Leu Thr Ile
455 460 465

Val Gly Asp Gly Lys Tyr Gly Ala Arg Asp Ala Asp Thr Lys Ile
470 475 480

Trp Asn Gly Met Val Gly Glu Leu Val Tyr Gly Lys Ala Asp Ile
485 490 495

Ala Ile Ala Pro Leu Thr Ile Thr Leu Val Arg Glu Glu Val Ile
500 505 510

Asp Phe Ser Lys Pro Phe Met Ser Leu Gly Ile Ser Ile Met Ile
515 520 525

Lys Lys Pro Gln Lys Ser Lys Pro Gly Val Phe Ser Phe Leu Asp
530 535 540

Pro Leu Ala Tyr Glu Ile Trp Met Cys Ile Val Phe Ala Tyr Ile
545 550 555

Gly Val Ser Val Val Leu Phe Leu Val Ser Arg Phe Ser Pro Tyr
560 565 570

Glu Trp His Thr Glu Glu Phe Glu Asp Gly Arg Glu Thr Gln Ser
575 580 585

Ser Glu Ser Thr Asn Glu Phe Gly Ile Phe Asn Ser Leu Trp Phe
590 595 600

Ser Leu Gly Ala Phe Met Arg Gln Gly Cys Asp Ile Ser Pro Arg
605 610 615

Ser Leu Ser Gly Arg Ile Val Gly Gly Val Trp Trp Phe Phe Thr
620 625 630

Leu Ile Ile Ile Ser Ser Tyr Thr Ala Asn Leu Ala Ala Phe Leu
635 640 645

Thr Val Glu Arg Met Val Ser Pro Ile Glu Ser Ala Glu Asp Leu
650 655 660

Ser Lys Gln Thr Glu Ile Ala Tyr Gly Thr Leu Asp Ser Gly Ser
665 670 675

Thr Lys Glu Phe Phe Arg Arg Ser Lys Ile Ala Val Phe Asp Lys
680 685 690

Met Trp Thr Tyr Met Arg Ser Ala Glu Pro Ser Val Phe Val Arg
695 700 705

Thr Thr Ala Glu Gly Val Ala Arg Val Arg Lys Ser Lys Gly Lys
710 715 720

Tyr Ala Tyr Leu Leu Glu Ser Thr Met Asn Glu Tyr Ile Glu Gln
725 730 735

Arg Lys Pro Cys Asp Thr Met Lys Val Gly Gly Asn Leu Asp Ser
740 745 750

Lys Gly Tyr Gly Ile Ala Thr Pro Lys Gly Ser Ser Leu Gly Thr
755 760 765

Pro Val Asn Leu Ala Val Leu Lys Leu Ser Glu Gln Gly Val Leu
770 775 780

Asp Lys Leu Lys Asn Lys Trp Trp Tyr Asp Lys Gly Glu Cys Gly
785 790 795

Ala Lys Asp Ser Gly Ser Lys Glu Lys Thr Ser Ala Leu Ser Leu
800 805 810

Ser Asn Val Ala Gly Val Phe Tyr Ile Leu Val Gly Gly Leu Gly
815 820 825

Leu Ala Met Leu Val Ala Leu Ile Glu Phe Cys Tyr Lys Ser Arg
830 835 840

Ala Glu Ala Lys Arg Met Lys Val Ala Lys Asn Ala Gln Asn Ile
845 850 855

Asn Pro Ser Ser Ser Gln Asn Ser Gln Asn Phe Ala Thr Tyr Lys
860 865 870

Glu Gly Tyr Asn Val Tyr Gly Ile Glu Ser Val Lys Ile
875 880

<210> 142
<211> 357
<212> PRT
<213> Homo sapiens

<400> 142
Met Ala Thr Thr Val Pro Asp Gly Cys Arg Asn Gly Leu Lys Ser
1 5 10 15

Lys Tyr Tyr Arg Leu Cys Asp Lys Ala Glu Ala Trp Gly Ile Val
20 25 30

Leu Glu Thr Val Ala Thr Ala Gly Val Val Thr Ser Val Ala Phe
35 40 45

Met Leu Thr Leu Pro Ile Leu Val Cys Lys Val Gln Asp Ser Asn
50 55 60

Arg Arg Lys Met Leu Pro Thr Gln Phe Leu Phe Leu Leu Gly Val
65 70 75

Leu Gly Ile Phe Gly Leu Thr Phe Ala Phe Ile Ile Gly Leu Asp
80 85 90

Gly Ser Thr Gly Pro Thr Arg Phe Phe Leu Phe Gly Ile Leu Phe

95 100 105

Ser Ile Cys Phe Ser Cys Leu Leu Ala His Ala Val Ser Leu Thr
110 115 120

Lys Leu Val Arg Gly Arg Lys Pro Leu Ser Leu Leu Val Ile Leu
125 130 135

Gly Leu Ala Val Gly Phe Ser Leu Val Gln Asp Val Ile Ala Ile
140 145 150

Glu Tyr Ile Val Leu Thr Met Asn Arg Thr Asn Val Asn Val Phe
155 160 165

Ser Glu Leu Ser Ala Pro Arg Arg Asn Glu Asp Phe Val Leu Leu
170 175 180

Leu Thr Tyr Val Leu Phe Leu Met Ala Leu Thr Phe Leu Met Ser
185 190 195

Ser Phe Thr Phe Cys Gly Ser Phe Thr Gly Trp Lys Arg His Gly
200 205 210

Ala His Ile Tyr Leu Thr Met Leu Leu Ser Ile Ala Ile Trp Val
215 220 225

Ala Trp Ile Thr Leu Leu Met Leu Pro Asp Phe Asp Arg Arg Trp
230 235 240

Asp Asp Thr Ile Leu Ser Ser Ala Leu Ala Ala Asn Gly Trp Val
245 250 255

Phe Leu Leu Ala Tyr Val Ser Pro Glu Phe Trp Leu Leu Thr Lys
260 265 270

Gln Arg Asn Pro Met Asp Tyr Pro Val Glu Asp Ala Phe Cys Lys
275 280 285

Pro Gln Leu Val Lys Lys Ser Tyr Gly Val Glu Asn Arg Ala Tyr
290 295 300

Ser Gln Glu Glu Ile Thr Gln Gly Phe Glu Glu Thr Gly Asp Thr
305 310 315

Leu Tyr Ala Pro Tyr Ser Thr His Phe Gln Leu Gln Asn Gln Pro
320 325 330

Pro Gln Lys Glu Phe Ser Ile Pro Arg Ala His Ala Trp Pro Ser
335 340 345

Pro Tyr Lys Asp Tyr Glu Val Lys Lys Glu Gly Ser
350 355

<210> 143
<211> 222
<212> PRT
<213> Homo sapiens

<400> 143
Met Leu Pro Ala Gln Glu Ala Ala Lys Leu Tyr His Thr Asn Tyr
1 5 10 15

Val Arg Asn Ser Arg Ala Ile Gly Val Leu Trp Ala Ile Phe Thr
20 25 30

Ile Cys Phe Ala Ile Val Asn Val Val Cys Phe Ile Gln Pro Tyr
35 40 45

Trp Ile Gly Asp Gly Val Asp Thr Pro Gln Ala Gly Tyr Phe Gly
50 55 60

Leu Phe His Tyr Cys Ile Gly Asn Gly Phe Ser Arg Glu Leu Thr
65 70 75

Cys Arg Gly Ser Phe Thr Asp Phe Ser Thr Leu Pro Ser Gly Ala
80 85 90

Phe Lys Ala Ala Ser Phe Phe Ile Gly Leu Ser Met Met Leu Ile
95 100 105

Ile Ala Cys Ile Ile Cys Phe Thr Leu Phe Phe Phe Cys Asn Thr
110 115 120

Ala Thr Val Tyr Lys Ile Cys Ala Trp Met Gln Leu Thr Ser Ala
125 130 135

Ala Cys Leu Val Leu Gly Cys Met Ile Phe Pro Asp Gly Trp Asp
140 145 150

Ser Asp Glu Val Lys Arg Met Cys Gly Glu Lys Thr Asp Lys Tyr
155 160 165

Thr Leu Gly Ala Cys Ser Val Arg Trp Ala Tyr Ile Leu Ala Ile
170 175 180

Ile Gly Ile Leu Asp Ala Leu Ile Leu Ser Phe Leu Ala Phe Val
185 190 195

Leu Gly Asn Arg Gln Asp Ser Leu Met Ala Glu Glu Leu Lys Ala
200 205 210

Glu Asn Lys Val Leu Leu Ser Gln Tyr Ser Leu Glu
215 220

<210> 144
<211> 236
<212> PRT
<213> Homo sapiens

<400> 144
Met Pro Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Met
1 5 10 15

Leu Pro Ala Gln Glu Ala Ala Lys Leu Tyr His Thr Asn Tyr Val
20 25 30

Arg Asn Ser Arg Ala Ile Gly Val Leu Trp Ala Ile Phe Thr Ile
35 40 45

Cys Phe Ala Ile Val Asn Val Val Cys Phe Ile Gln Pro Tyr Trp
50 55 60

Ile Gly Asp Gly Val Asp Thr Pro Gln Ala Gly Tyr Phe Gly Leu
65 70 75

Phe His Tyr Cys Ile Gly Asn Gly Phe Ser Arg Glu Leu Thr Cys
80 85 90

Arg Gly Ser Phe Thr Asp Phe Ser Thr Leu Pro Ser Gly Ala Phe
95 100 105

Lys Ala Ala Ser Phe Phe Ile Gly Leu Ser Met Met Leu Ile Ile
110 115 120

Ala Cys Ile Ile Cys Phe Thr Leu Phe Phe Phe Cys Asn Thr Ala
125 130 135

Thr Val Tyr Lys Ile Cys Ala Trp Met Gln Leu Thr Ser Ala Ala
140 145 150

Cys Leu Val Leu Gly Cys Met Ile Phe Pro Asp Gly Trp Asp Ser
155 160 165

Asp Glu Val Lys Arg Met Cys Gly Glu Lys Thr Asp Lys Tyr Thr
170 175 180

Leu Gly Ala Cys Ser Val Arg Trp Ala Tyr Ile Leu Ala Ile Ile
185 190 195

Gly Ile Leu Asp Ala Leu Ile Leu Ser Phe Leu Ala Phe Val Leu
200 205 210

Gly Asn Arg Gln Asp Ser Leu Met Ala Glu Glu Leu Lys Ala Glu
215 220 225

Asn Lys Val Leu Leu Ser Gln Tyr Ser Leu Glu
230 235

<210> 145
<211> 236
<212> PRT
<213> Homo sapiens

<400> 145
Met Pro Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Met
1 5 10 15

Leu Pro Ala Gln Glu Ala Ala Lys Leu Tyr His Thr Asn Tyr Val
20 25 30

Arg Asn Ser Arg Ala Ile Gly Val Leu Trp Ala Ile Phe Thr Ile
35 40 45

Cys Phe Ala Ile Val Asn Val Val Cys Phe Ile Gln Pro Tyr Trp
50 55 60

Ile Gly Asp Gly Val Asp Thr Pro Gln Ala Gly Tyr Phe Gly Leu
65 70 75

Phe His Tyr Cys Ile Gly Asn Gly Phe Ser Arg Glu Leu Thr Cys
80 85 90

Arg Gly Ser Phe Thr Asp Phe Ser Thr Leu Pro Ser Gly Ala Phe
95 100 105

Lys Ala Ala Ser Phe Phe Ile Gly Leu Ser Met Met Leu Ile Ile
110 115 120

Ala Cys Ile Ile Cys Phe Thr Leu Phe Phe Phe Cys Asn Thr Ala
125 130 135

Thr Val Tyr Lys Ile Cys Ala Trp Met Gln Leu Thr Ser Ala Ala
140 145 150

Cys Leu Val Leu Gly Cys Met Ile Phe Pro Asp Gly Trp Asp Ser
155 160 165

Asp Glu Val Lys Arg Met Cys Gly Glu Lys Thr Asp Lys Tyr Thr
170 175 180

Leu Gly Ala Cys Ser Val Arg Trp Ala Tyr Ile Leu Ala Ile Ile
185 190 195

Gly Ile Leu Asp Ala Leu Ile Leu Ser Phe Leu Ala Phe Val Leu
200 205 210

Gly Asn Arg Gln Asp Ser Leu Met Ala Glu Glu Leu Lys Ala Glu
215 220 225

Asn Lys Val Leu Leu Ser Gln Tyr Ser Leu Glu
230 235

<210> 146
<211> 408
<212> PRT
<213> Homo sapiens

<400> 146
Arg Pro Leu Arg Leu Pro Cys Arg Gly Pro Arg Ala Ala Gly Thr
1 5 10 15

Val Ser Ser Glu Ser Leu Gly His Pro Ala Pro Gly Gln Val Thr
20 25 30

Asn Arg Ser Pro Gln Phe Pro Met Glu Leu Phe Gln Ser Ser Pro
35 40 45

Arg Pro Pro Leu Gln Pro Val Cys Val Cys Pro Gly Leu Glu Leu

50 55 60

Cys Asp Pro Pro His Arg Leu Ala Phe Ser Met Gly Ala Gly Leu
65 70 75

Phe Ser Val Val Gly Thr Leu Leu Leu Pro Gly Leu Ala Ala Leu
80 85 90

Val Gln Asp Trp Arg Leu Leu Gln Gly Leu Gly Ala Leu Met Ser
95 100 105

Gly Leu Leu Leu Leu Phe Trp Gly Phe Pro Ala Leu Phe Pro Glu
110 115 120

Ser Pro Cys Trp Leu Leu Ala Thr Gly Gln Val Ala Arg Ala Arg
125 130 135

Lys Ile Leu Trp Arg Phe Ala Glu Ala Ser Gly Val Gly Pro Gly
140 145 150

Asp Ser Ser Leu Glu Glu Asn Ser Leu Ala Thr Glu Leu Thr Met
155 160 165

Leu Ser Ala Arg Ser Pro Gln Pro Arg Tyr His Ser Pro Leu Gly
170 175 180

Leu Leu Arg Thr Arg Val Thr Trp Arg Asn Gly Leu Ile Leu Gly
185 190 195

Phe Ser Ser Leu Val Gly Gly Gly Ile Arg Ala Ser Phe Arg Arg
200 205 210

Ser Leu Ala Pro Gln Val Pro Thr Phe Tyr Leu Pro Tyr Phe Leu
215 220 225

Glu Ala Gly Leu Glu Ala Ala Ala Leu Val Phe Leu Leu Leu Thr
230 235 240

Ala Asp Cys Cys Gly Arg Arg Pro Val Leu Leu Leu Gly Thr Met
245 250 255

Val Thr Gly Leu Ala Ser Leu Leu Leu Leu Ala Gly Ala Gln Tyr
260 265 270

Leu Pro Gly Trp Thr Val Leu Phe Leu Ser Val Leu Gly Leu Leu
275 280 285

Ala Ser Arg Ala Val Ser Ala Leu Ser Ser Leu Phe Ala Ala Glu
290 295 300

Val Phe Pro Thr Val Ile Arg Gly Ala Gly Leu Gly Leu Val Leu
305 310 315

Gly Ala Gly Phe Leu Gly Gln Ala Ala Gly Pro Leu Asp Thr Leu
320 325 330

His Gly Arg Gln Gly Phe Phe Leu Gln Gln Val Val Phe Ala Ser
335 340 345

Leu Ala Val Leu Ala Leu Leu Cys Val Leu Leu Leu Pro Glu Ser
350 355 360

Arg Ser Arg Gly Leu Pro Gln Ser Leu Gln Asp Ala Asp Arg Leu
365 370 375

Arg Arg Ser Pro Leu Leu Arg Gly Arg Pro Arg Gln Asp His Leu
380 385 390

Pro Leu Leu Pro Pro Ser Asn Ser Tyr Trp Ala Gly His Thr Pro
395 400 405

Glu Gln His

<210> 147
<211> 1152

<212> PRT
<213> Homo sapiens

<220>
<221> unsure
<222> 382
<223> unknown amino acid

<400> 147

Met	Pro	Phe	Glu	Gln	Cys	His	Leu	Ser	Trp	Lys	Asn	Val	Arg	Arg	
1				5					10						15
Leu	Tyr	Asn	Ala	Val	Ser	Lys	Glu	Leu	Val	Gly	Glu	Phe	Leu	Gln	
				20					25						30
Phe	Val	Gln	Leu	Asp	Lys	Glu	Ala	Ser	Asp	Pro	Phe	Ser	Leu	Asn	
				35					40						45
Glu	Leu	Leu	Asp	Glu	Leu	Ser	Arg	Lys	Gln	Lys	Glu	Glu	Leu	Trp	
				50					55						60
Gln	Arg	Leu	Lys	Asn	Leu	Leu	Thr	Asp	Val	Leu	Leu	Glu	Ser	Pro	
				65					70						75
Val	Asp	Gly	Trp	Gln	Val	Val	Glu	Ala	Gln	Gly	Glu	Asp	Asn	Met	
				80					85						90
Glu	Thr	Glu	His	Gly	Ser	Lys	Met	Arg	Lys	Ser	Ile	Glu	Ile	Ile	
				95					100						105
Tyr	Ala	Ile	Thr	Ser	Val	Ile	Leu	Ala	Ser	Val	Ser	Val	Ile	Asn	
				110					115						120
Glu	Ser	Glu	Asn	Tyr	Glu	Ala	Leu	Leu	Glu	Cys	Val	Ile	Ile	Leu	
				125					130						135
Asn	Gly	Ile	Leu	Tyr	Ala	Leu	Pro	Glu	Ser	Glu	Arg	Lys	Leu	Gln	
				140					145						150
Ser	Ser	Ile	Gln	Asp	Leu	Cys	Val	Thr	Trp	Trp	Glu	Lys	Gly	Leu	
				155					160						165
Pro	Ala	Lys	Glu	Asp	Thr	Gly	Lys	Thr	Ala	Phe	Val	Met	Leu	Leu	
				170					175						180
Arg	Arg	Ser	Leu	Glu	Thr	Lys	Thr	Gly	Ala	Asp	Val	Cys	Arg	Leu	
				185					190						195
Trp	Arg	Ile	His	Gln	Ala	Leu	Tyr	Cys	Phe	Asp	Tyr	Asp	Leu	Glu	
				200					205						210
Glu	Ser	Gly	Glu	Ile	Lys	Asp	Met	Leu	Leu	Glu	Cys	Phe	Ile	Asn	
				215					220						225
Ile	Asn	Tyr	Ile	Lys	Lys	Glu	Glu	Gly	Arg	Arg	Phe	Leu	Ser	Cys	
				230					235						240
Leu	Phe	Asn	Trp	Asn	Ile	Asn	Phe	Ile	Lys	Met	Ile	His	Gly	Thr	
				245					250						255
Ile	Lys	Asn	Gln	Leu	Gln	Gly	Leu	Gln	Lys	Ser	Leu	Met	Val	Tyr	
				260					265						270
Ile	Ala	Glu	Ile	Tyr	Phe	Arg	Ala	Trp	Lys	Lys	Ala	Ser	Gly	Lys	
				275					280						285
Ile	Leu	Glu	Ala	Ile	Glu	Asn	Asp	Cys	Ile	Gln	Asp	Phe	Met	Phe	
				290					295						300
His	Gly	Ile	His	Leu	Pro	Arg	Arg	Ser	Pro	Val	His	Ser	Lys	Val	
				305					310						315
Arg	Glu	Val	Leu	Ser	Tyr	Phe	His	His	Gln	Lys	Lys	Val	Arg	Gln	
				320					325						330
Gly	Val	Glu	Glu	Met	Leu	Tyr	Arg	Leu	Tyr	Lys	Pro	Ile	Leu	Trp	

335 340 345

Arg Gly Leu Lys Ala Arg Asn Ser Glu Val Arg Ser Asn Ala Ala
350 355 360

Leu Leu Phe Val Glu Ala Phe Pro Ile Arg Asp Pro Asn Leu His
365 370 375

Ala Ile Glu Met Asp Ser Xaa Ile Gln Lys Gln Phe Glu Glu Leu
380 385 390

Tyr Ser Leu Leu Glu Asp Pro Tyr Pro Met Val Arg Ser Thr Gly
395 400 405

Ile Leu Gly Val Cys Lys Ile Thr Ser Lys Tyr Trp Glu Met Met
410 415 420

Pro Pro Thr Ile Leu Ile Asp Leu Leu Lys Lys Val Thr Gly Glu
425 430 435

Leu Ala Phe Asp Thr Ser Ser Ala Asp Val Arg Cys Ser Val Phe
440 445 450

Lys Cys Leu Pro Met Ile Leu Asp Asn Lys Leu Ser His Pro Leu
455 460 465

Leu Glu Gln Leu Leu Pro Ala Leu Arg Tyr Ser Leu His Asp Asn
470 475 480

Ser Glu Lys Val Arg Val Ala Phe Val Asp Met Leu Leu Lys Ile
485 490 495

Lys Ala Val Arg Ala Ala Lys Phe Trp Lys Ile Cys Pro Met Glu
500 505 510

His Ile Leu Val Arg Leu Glu Thr Asp Ser Arg Pro Val Ser Arg
515 520 525

Arg Leu Val Ser Leu Ile Phe Asn Ser Phe Leu Pro Val Asn Gln
530 535 540

Pro Glu Glu Val Trp Cys Glu Arg Cys Val Thr Leu Val Gln Met
545 550 555

Asn His Ala Ala Ala Arg Arg Phe Tyr Gln Tyr Ala His Glu His
560 565 570

Thr Ala Cys Thr Asn Ile Ala Lys Leu Ile His Val Ile Arg His
575 580 585

Cys Leu Asn Ala Cys Ile Gln Arg Ala Val Arg Glu Pro Pro Glu
590 595 600

Asp Glu Glu Glu Glu Asp Gly Arg Glu Lys Glu Asn Val Thr Val
605 610 615

Leu Asp Lys Thr Leu Ser Val Asn Asp Val Ala Cys Met Ala Gly
620 625 630

Leu Leu Glu Ile Ile Val Ile Leu Trp Lys Ser Ile Asp Arg Ser
635 640 645

Met Glu Asn Asn Lys Glu Ala Lys Leu Tyr Thr Ile Asn Lys Phe
650 655 660

Ala Ser Val Leu Pro Glu Tyr Leu Lys Val Phe Lys Asp Asp Arg
665 670 675

Cys Lys Ile Pro Leu Phe Met Leu Met Ser Phe Met Pro Ala Ser
680 685 690

Ala Val Pro Pro Phe Ser Cys Gly Val Ile Ser Thr Leu Arg Ser
695 700 705

Arg Glu Glu Gly Ala Val Asp Lys Ser Tyr Cys Thr Leu Leu Asp
710 715 720

Cys Leu Cys Ser Trp Gly Gln Val Gly His Ile Leu Glu Leu Val
725 730 735

Asp Asn Trp Leu Pro Thr Glu His Ala Gln Ala Lys Ser Asn Thr
740 745 750

Ala Ser Lys Gly Arg Val Gln Ile His Asp Thr Arg Pro Val Lys
755 760 765

Pro Glu Leu Ala Leu Val Tyr Ile Glu Tyr Leu Leu Thr His Pro
770 775 780

Lys Asn Arg Glu Cys Leu Leu Ser Ala Pro Arg Lys Lys Leu Asn
785 790 795

His Leu Leu Lys Ala Leu Glu Thr Ser Lys Ala Asp Leu Glu Ser
800 805 810

Leu Leu Gln Thr Pro Gly Gly Lys Pro Arg Gly Phe Ser Glu Ala
815 820 825

Ala Ala Pro Arg Ala Phe Gly Leu His Cys Arg Leu Ser Ile His
830 835 840

Leu Gln His Lys Phe Cys Ser Glu Gly Lys Val Tyr Leu Ser Met
845 850 855

Leu Glu Asp Thr Gly Phe Trp Leu Glu Ser Lys Ile Leu Ser Phe
860 865 870

Ile Gln Asp Gln Glu Glu Asp Tyr Leu Lys Leu His Arg Val Ile
875 880 885

Tyr Gln Gln Ile Ile Gln Thr Tyr Leu Thr Val Cys Lys Asp Val
890 895 900

Val Met Val Gly Leu Gly Asp His Gln Phe Gln Met Gln Leu Leu
905 910 915

Gln Arg Ser Leu Gly Ile Met Gln Thr Val Lys Gly Phe Phe Tyr
920 925 930

Val Ser Leu Leu Leu Asp Ile Leu Lys Glu Ile Thr Gly Ser Ser
935 940 945

Leu Ile Gln Lys Thr Asp Ser Asp Glu Glu Val Ala Met Leu Leu
950 955 960

Asp Thr Val Gln Lys Val Phe Gln Lys Met Leu Glu Cys Ile Ala
965 970 975

Arg Ser Phe Arg Lys Gln Pro Glu Glu Gly Leu Arg Leu Leu Tyr
980 985 990

Ser Val Gln Arg Pro Leu His Glu Phe Ile Thr Ala Val Gln Ser
995 1000 1005

Arg His Thr Asp Thr Pro Val His Arg Gly Val Leu Ser Thr Leu
1010 1015 1020

Ile Ala Gly Pro Val Val Glu Ile Ser His Gln Leu Arg Lys Val
1025 1030 1035

Ser Asp Val Glu Glu Leu Thr Pro Pro Glu His Leu Ser Asp Leu
1040 1045 1050

Pro Pro Phe Ser Arg Cys Leu Ile Gly Ile Ile Ile Lys Ser Ser
1055 1060 1065

Asn Val Val Arg Ser Phe Leu Asp Glu Leu Lys Ala Cys Val Ala
1070 1075 1080

Ser Asn Asp Ile Glu Gly Ile Val Cys Leu Thr Ala Ala Val His
1085 1090 1095

Ile Ile Leu Val Ile Asn Ala Gly Lys His Lys Ser Ser Lys Val
1100 1105 1110

Arg Glu Val Ala Ala Thr Val His Arg Lys Leu Lys Thr Phe Met
1115 1120 1125

Glu Ile Thr Leu Glu Glu Asp Ser Ile Glu Arg Phe Leu Tyr Glu
1130 1135 1140

Ser Ser Ser Arg Thr Leu Gly Glu Leu Leu Asn Ser
1145 1150

<210> 148
<211> 1143
<212> PRT
<213> Homo sapiens

<400> 148
Met Glu Lys Arg Glu Thr Phe Val Gln Ala Val Ser Lys Glu Leu
1 5 10 15

Val Gly Glu Phe Leu Gln Phe Val Gln Leu Asp Lys Glu Ala Ser
20 25 30

Asp Pro Phe Ser Leu Asn Glu Leu Leu Asp Glu Leu Ser Arg Lys
35 40 45

Gln Lys Glu Glu Leu Trp Gln Arg Leu Lys Asn Leu Leu Thr Asp
50 55 60

Val Leu Leu Glu Ser Pro Val Asp Gly Trp Gln Val Val Glu Ala
65 70 75

Gln Gly Glu Asp Asn Met Glu Thr Glu His Gly Ser Lys Met Arg
80 85 90

Lys Ser Ile Glu Ile Ile Tyr Ala Ile Thr Ser Val Ile Leu Ala
95 100 105

Ser Val Ser Val Ile Asn Glu Ser Glu Asn Tyr Glu Ala Leu Leu
110 115 120

Glu Cys Val Ile Ile Leu Asn Gly Ile Leu Tyr Ala Leu Pro Glu
125 130 135

Ser Glu Arg Lys Leu Gln Ser Ser Ile Gln Asp Leu Cys Val Thr
140 145 150

Trp Trp Glu Lys Gly Leu Pro Ala Lys Glu Asp Thr Gly Lys Thr
155 160 165

Ala Phe Val Met Leu Leu Arg Arg Ser Leu Glu Thr Lys Thr Gly
170 175 180

Ala Asp Val Cys Arg Leu Trp Arg Ile His Gln Ala Leu Tyr Cys
185 190 195

Phe Asp Tyr Asp Leu Glu Glu Ser Gly Glu Ile Lys Asp Met Leu
200 205 210

Leu Glu Cys Phe Ile Asn Ile Asn Tyr Ile Lys Lys Glu Glu Gly
215 220 225

Arg Arg Phe Leu Ser Cys Leu Phe Asn Trp Asn Ile Asn Phe Ile
230 235 240

Lys Met Ile His Gly Thr Ile Lys Asn Gln Leu Gln Gly Leu Gln
245 250 255

Lys Ser Leu Met Val Tyr Ile Ala Glu Ile Tyr Phe Arg Ala Trp
260 265 270

Lys Lys Ala Ser Gly Lys Ile Leu Glu Ala Ile Glu Asn Asp Cys
275 280 285

Ile Gln Asp Phe Met Phe His Gly Ile His Leu Pro Arg Arg Ser

290 295 300

Pro Val His Ser Lys Val Arg Glu Val Leu Ser Tyr Phe His His
305 310 315

Gln Lys Lys Val Arg Gln Gly Val Glu Glu Met Leu Tyr Arg Leu
320 325 330

Tyr Lys Pro Ile Leu Trp Arg Gly Leu Lys Ala Arg Asn Ser Glu
335 340 345

Val Arg Ser Asn Ala Ala Leu Leu Phe Val Glu Ala Phe Pro Ile
350 355 360

Arg Asp Pro Asn Leu His Ala Ile Glu Met Asp Ser Glu Ile Gln
365 370 375

Lys Gln Phe Glu Glu Leu Tyr Ser Leu Leu Glu Asp Pro Tyr Pro
380 385 390

Met Val Arg Ser Thr Gly Ile Leu Gly Val Cys Lys Ile Thr Ser
395 400 405

Lys Tyr Trp Glu Met Met Pro Pro Thr Ile Leu Ile Asp Leu Leu
410 415 420

Lys Lys Val Thr Gly Glu Leu Ala Phe Asp Thr Ser Ser Ala Asp
425 430 435

Val Arg Cys Ser Val Phe Lys Cys Leu Pro Met Ile Leu Asp Asn
440 445 450

Lys Leu Ser His Pro Leu Leu Glu Gln Leu Leu Pro Ala Leu Arg
455 460 465

Tyr Ser Leu His Asp Asn Ser Glu Lys Val Arg Val Ala Phe Val
470 475 480

Asp Met Leu Leu Lys Ile Lys Ala Val Arg Ala Ala Lys Phe Trp
485 490 495

Lys Ile Cys Pro Met Glu His Ile Leu Val Arg Leu Glu Thr Asp
500 505 510

Ser Arg Pro Val Ser Arg Arg Leu Val Ser Leu Ile Phe Asn Ser
515 520 525

Phe Leu Pro Val Asn Gln Pro Glu Glu Val Trp Cys Glu Arg Cys
530 535 540

Val Thr Leu Val Gln Met Asn His Ala Ala Ala Arg Arg Phe Tyr
545 550 555

Gln Tyr Ala His Glu His Thr Ala Cys Thr Asn Ile Ala Lys Leu
560 565 570

Ile His Val Ile Arg His Cys Leu Asn Ala Cys Ile Gln Arg Ala
575 580 585

Val Arg Glu Pro Pro Glu Asp Glu Glu Glu Glu Asp Gly Arg Glu
590 595 600

Lys Glu Asn Val Thr Val Leu Asp Lys Thr Leu Ser Val Asn Asp
605 610 615

Val Ala Cys Met Ala Gly Leu Leu Glu Ile Ile Val Ile Leu Trp
620 625 630

Lys Ser Ile Asp Arg Ser Met Glu Asn Asn Lys Glu Ala Lys Leu
635 640 645

Tyr Thr Ile Asn Lys Phe Ala Ser Val Leu Pro Glu Tyr Leu Lys
650 655 660

Val Phe Lys Asp Asp Arg Cys Lys Ile Pro Leu Phe Met Leu Met
665 670 675

Ser Phe Met Pro Ala Ser Ala Val Pro Pro Phe Ser Cys Gly Val
680 685 690

Ile Ser Thr Leu Arg Ser Arg Glu Glu Gly Ala Val Asp Lys Ser
695 700 705

Tyr Cys Thr Leu Leu Asp Cys Leu Cys Ser Trp Gly Gln Val Gly
710 715 720

His Ile Leu Glu Leu Val Asp Asn Trp Leu Pro Thr Glu His Ala
725 730 735

Gln Ala Lys Ser Asn Thr Ala Ser Lys Gly Arg Val Gln Ile His
740 745 750

Asp Thr Arg Pro Val Lys Pro Glu Leu Ala Leu Val Tyr Ile Glu
755 760 765

Tyr Leu Leu Thr His Pro Lys Asn Arg Glu Cys Leu Leu Ser Ala
770 775 780

Pro Arg Lys Lys Leu Asn His Leu Leu Lys Ala Leu Glu Thr Ser
785 790 795

Lys Ala Asp Leu Glu Ser Leu Leu Gln Thr Pro Gly Gly Lys Pro
800 805 810

Arg Gly Phe Ser Glu Ala Ala Ala Pro Arg Ala Phe Gly Leu His
815 820 825

Cys Arg Leu Ser Ile His Leu Gln His Lys Phe Cys Ser Glu Gly
830 835 840

Lys Val Tyr Leu Ser Met Leu Glu Asp Thr Gly Phe Trp Leu Glu
845 850 855

Ser Lys Ile Leu Ser Phe Ile Gln Asp Gln Glu Glu Asp Tyr Leu
860 865 870

Lys Leu His Arg Val Ile Tyr Gln Gln Ile Ile Gln Thr Tyr Leu
875 880 885

Thr Val Cys Lys Asp Val Val Met Val Gly Leu Gly Asp His Gln
890 895 900

Phe Gln Met Gln Leu Leu Gln Arg Ser Leu Gly Ile Met Gln Thr
905 910 915

Val Lys Gly Phe Phe Tyr Val Ser Leu Leu Leu Asp Ile Leu Lys
920 925 930

Glu Ile Thr Gly Ser Ser Leu Ile Gln Lys Thr Asp Ser Asp Glu
935 940 945

Glu Val Ala Met Leu Leu Asp Thr Val Gln Lys Val Phe Gln Lys
950 955 960

Met Leu Glu Cys Ile Ala Arg Ser Phe Arg Lys Gln Pro Glu Glu
965 970 975

Gly Leu Arg Leu Leu Tyr Ser Val Gln Arg Pro Leu His Glu Phe
980 985 990

Ile Thr Ala Val Gln Ser Arg His Thr Asp Thr Pro Val His Arg
995 1000 1005

Gly Val Leu Ser Thr Leu Ile Ala Gly Pro Val Val Glu Ile Ser
1010 1015 1020

His Gln Leu Arg Lys Val Ser Asp Val Glu Glu Leu Thr Pro Pro
1025 1030 1035

Glu His Leu Ser Asp Leu Pro Pro Phe Ser Arg Cys Leu Ile Gly
1040 1045 1050

Ile Ile Ile Lys Ser Ser Asn Val Val Arg Ser Phe Leu Asp Glu
1055 1060 1065

Leu Lys Ala Cys Val Ala Ser Asn Asp Ile Glu Gly Ile Val Cys
1070 1075 1080

Leu Thr Ala Ala Val His Ile Ile Leu Val Ile Asn Ala Gly Lys
1085 1090 1095

His Lys Ser Ser Lys Val Arg Glu Val Ala Ala Thr Val His Arg
1100 1105 1110

Lys Leu Lys Thr Phe Met Glu Ile Thr Leu Glu Glu Asp Ser Ile
1115 1120 1125

Glu Arg Phe Leu Tyr Glu Ser Ser Ser Arg Thr Leu Gly Glu Leu
1130 1135 1140

Leu Asn Ser

<210> 149
<211> 384
<212> PRT
<213> Homo sapiens

<400> 149
Met Asn Ser Thr Leu Phe Ser Gln Val Glu Asn His Ser Val His
1 5 10 15

Ser Asn Phe Ser Glu Lys Asn Ala Gln Leu Leu Ala Phe Glu Asn
20 25 30

Asp Asp Cys His Leu Pro Leu Ala Met Ile Phe Thr Leu Ala Leu
35 40 45

Ala Tyr Gly Ala Val Ile Ile Leu Gly Val Ser Gly Asn Leu Ala
50 55 60

Leu Ile Ile Ile Ile Leu Lys Gln Lys Glu Met Arg Asn Val Thr
65 70 75

Asn Ile Leu Ile Val Asn Leu Ser Phe Ser Asp Leu Leu Val Ala
80 85 90

Ile Met Cys Leu Pro Phe Thr Phe Val Tyr Thr Leu Met Asp His
95 100 105

Trp Val Phe Gly Glu Ala Met Cys Lys Leu Asn Pro Phe Val Gln
110 115 120

Cys Val Ser Ile Thr Val Ser Ile Phe Ser Leu Val Leu Ile Ala
125 130 135

Val Glu Arg His Gln Leu Ile Ile Asn Pro Arg Gly Trp Arg Pro
140 145 150

Asn Asn Arg His Ala Tyr Val Gly Ile Ala Val Ile Trp Val Leu
155 160 165

Ala Val Ala Ser Ser Leu Pro Phe Leu Ile Tyr Gln Val Met Thr
170 175 180

Asp Glu Pro Phe Gln Asn Val Thr Leu Asp Ala Tyr Lys Asp Lys
185 190 195

Tyr Val Cys Phe Asp Gln Phe Pro Ser Asp Ser His Arg Leu Ser
200 205 210

Tyr Thr Thr Leu Leu Leu Val Leu Gln Tyr Phe Gly Pro Leu Cys
215 220 225

Phe Ile Phe Ile Cys Tyr Phe Lys Ile Tyr Ile Arg Leu Lys Arg
230 235 240

Arg Asn Asn Met Met Asp Lys Met Arg Asp Asn Lys Tyr Arg Ser

158C

245 250 255

Ser Glu Thr Lys Arg Ile Asn Ile Met Leu Leu Ser Ile Val Val
260 265 270

Ala Phe Ala Val Cys Trp Leu Pro Leu Thr Ile Phe Asn Thr Val
275 280 285

Phe Asp Trp Asn His Gln Ile Ile Ala Thr Cys Asn His Asn Leu
290 295 300

Leu Phe Leu Leu Cys His Leu Thr Ala Met Ile Ser Thr Cys Val
305 310 315

Asn Pro Ile Phe Tyr Gly Phe Leu Asn Lys Asn Phe Gln Arg Asp
320 325 330

Leu Gln Phe Phe Phe Asn Phe Cys Asp Phe Arg Ser Arg Asp Asp
335 340 345

Asp Tyr Glu Thr Ile Ala Met Ser Thr Met His Thr Asp Val Ser
350 355 360

Lys Thr Ser Leu Lys Gln Ala Ser Pro Val Ala Phe Lys Lys Ile
365 370 375

Asn Asn Asn Asp Asp Asn Glu Lys Ile
380

<210> 150
<211> 327
<212> PRT
<213> Homo sapiens

<400> 150
Met Val Arg Cys Asp Arg Gly Leu Gln Met Leu Leu Thr Thr Ala
1 5 10 15

Gly Ala Phe Ala Ala Phe Ser Leu Met Ala Ile Ala Ile Gly Thr
20 25 30

Asp Tyr Trp Leu Tyr Ser Ser Ala His Ile Cys Asn Gly Thr Asn
35 40 45

Leu Thr Met Asp Asp Gly Pro Pro Pro Arg Arg Ala Arg Gly Asp
50 55 60

Leu Thr His Ser Gly Leu Trp Arg Val Cys Cys Ile Glu Gly Ile
65 70 75

Tyr Lys Gly His Cys Phe Arg Ile Asn His Phe Pro Glu Asp Asn
80 85 90

Asp Tyr Asp His Asp Ser Ser Glu Tyr Leu Leu Arg Ile Val Arg
95 100 105

Ala Ser Ser Val Phe Pro Ile Leu Ser Thr Ile Leu Leu Leu Leu
110 115 120

Gly Gly Leu Cys Ile Gly Ala Gly Arg Ile Tyr Ser Arg Lys Asn
125 130 135

Asn Ile Val Leu Ser Ala Gly Ile Leu Phe Val Ala Ala Gly Leu
140 145 150

Ser Asn Ile Ile Gly Ile Ile Val Tyr Ile Ser Ser Asn Thr Gly
155 160 165

Asp Pro Ser Asp Lys Arg Asp Glu Asp Lys Lys Asn His Tyr Asn
170 175 180

Tyr Gly Trp Ser Phe Tyr Phe Gly Ala Leu Ser Phe Ile Val Ala
185 190 195

Glu Thr Val Gly Val Leu Ala Val Asn Ile Tyr Ile Glu Lys Asn
200 205 210

1570
1581

Lys Glu Leu Arg Phe Lys Thr Lys Arg Glu Phe Leu Lys Ala Ser
215 220 225

Ser Ser Ser Pro Tyr Ala Arg Met Pro Ser Tyr Arg Tyr Arg Arg
230 235 240

Arg Arg Ser Arg Ser Ser Ser Arg Ser Thr Glu Ala Ser Pro Ser
245 250 255

Arg Asp Val Ser Pro Met Gly Leu Lys Ile Thr Gly Ala Ile Pro
260 265 270

Met Gly Glu Leu Ser Met Tyr Thr Leu Ser Arg Glu Pro Leu Lys
275 280 285

Val Thr Thr Ala Ala Ser Tyr Ser Pro Asp Gln Glu Ala Ser Phe
290 295 300

Leu Gln Val His Asp Phe Phe Gln Gln Asp Leu Lys Glu Gly Phe
305 310 315

His Val Ser Met Leu Asn Arg Arg Thr Thr Pro Val
320 325

<210> 151
<211> 323
<212> PRT
<213> Homo sapiens

<400> 151
Met Ala Arg Gly Pro Gly Leu Ala Pro Pro Pro Leu Arg Leu Pro
1 5 10 15

Leu Leu Leu Leu Val Leu Ala Ala Val Thr Gly His Thr Ala Ala
20 25 30

Gln Asp Asn Cys Thr Cys Pro Thr Asn Lys Met Thr Val Cys Ser
35 40 45

Pro Asp Gly Pro Gly Gly Arg Cys Gln Cys Arg Ala Leu Gly Ser
50 55 60

Gly Met Ala Val Asp Cys Ser Thr Leu Thr Ser Lys Cys Leu Leu
65 70 75

Leu Lys Ala Arg Met Ser Ala Pro Lys Asn Ala Arg Thr Leu Val
80 85 90

Arg Pro Ser Glu His Ala Leu Val Asp Asn Asp Gly Leu Tyr Asp
95 100 105

Pro Asp Cys Asp Pro Glu Gly Arg Phe Lys Ala Arg Gln Cys Asn
110 115 120

Gln Thr Ser Val Cys Trp Cys Val Asn Ser Val Gly Val Arg Arg
125 130 135

Thr Asp Lys Gly Asp Leu Ser Leu Arg Cys Asp Glu Leu Val Arg
140 145 150

Thr His His Ile Leu Ile Asp Leu Arg His Arg Pro Thr Ala Gly
155 160 165

Ala Phe Asn His Ser Asp Leu Asp Ala Glu Leu Arg Arg Leu Phe
170 175 180

Arg Glu Arg Tyr Arg Leu His Pro Lys Phe Val Ala Ala Val His
185 190 195

Tyr Glu Gln Pro Thr Ile Gln Ile Glu Leu Arg Gln Asn Thr Ser
200 205 210

Gln Lys Ala Ala Gly Glu Val Asp Ile Gly Asp Ala Ala Tyr Tyr
215 220 225

Phe Glu Arg Asp Ile Lys Gly Glu Ser Leu Phe Gln Gly Arg Gly
230 235 240

Gly Leu Asp Leu Arg Val Arg Gly Glu Pro Leu Gln Val Glu Arg
245 250 255

Thr Leu Ile Tyr Tyr Leu Asp Glu Ile Pro Pro Lys Phe Ser Met
260 265 270

Lys Arg Leu Thr Ala Gly Leu Ile Ala Val Ile Val Val Val Val
275 280 285

Val Ala Leu Val Ala Gly Met Ala Val Leu Val Ile Thr Asn Arg
290 295 300

Arg Lys Ser Gly Lys Tyr Lys Lys Val Glu Ile Lys Glu Leu Gly
305 310 315

Glu Leu Arg Lys Glu Pro Ser Leu
320

<210> 152
<211> 207
<212> PRT
<213> Homo sapiens

<400> 152
Met Gly Lys Cys Ser Gly Arg Cys Thr Leu Val Ala Phe Cys Cys
1 5 10 15

Leu Gln Leu Val Ala Ala Leu Glu Arg Gln Ile Phe Asp Phe Leu
20 25 30

Gly Tyr Gln Trp Ala Pro Ile Leu Ala Asn Phe Leu His Ile Met
35 40 45

Ala Val Ile Leu Gly Ile Phe Gly Thr Val Gln Tyr Arg Ser Arg
50 55 60

Tyr Leu Ile Leu Tyr Ala Ala Trp Leu Val Leu Trp Val Gly Trp
65 70 75

Asn Ala Phe Ile Ile Cys Phe Tyr Leu Glu Val Gly Gln Leu Ser
80 85 90

Gln Asp Arg Asp Phe Ile Met Thr Phe Asn Thr Ser Leu His Arg
95 100 105

Ser Trp Trp Met Glu Asn Gly Pro Gly Cys Leu Val Thr Pro Val
110 115 120

Leu Asn Ser Arg Leu Ala Leu Glu Asp His His Val Ile Ser Val
125 130 135

Thr Gly Cys Leu Leu Asp Tyr Pro Tyr Ile Glu Ala Leu Ser Ser
140 145 150

Ala Leu Gln Ile Phe Leu Ala Leu Phe Gly Phe Val Phe Ala Cys
155 160 165

Tyr Val Ser Lys Val Phe Leu Glu Glu Glu Asp Ser Phe Asp Phe
170 175 180

Ile Gly Gly Phe Asp Ser Tyr Gly Tyr Gln Ala Pro Gln Lys Thr
185 190 195

Ser His Leu Gln Leu Gln Pro Leu Tyr Thr Ser Gly
200 205

<210> 153
<211> 581
<212> PRT
<213> Homo sapiens

<400> 153
Met Pro Ala Pro Arg Ala Arg Glu Gln Pro Arg Val Pro Gly Glu

1542
1583

```
1               5                  10                  15

Arg Gln Pro Leu Leu Pro Arg Gly Ala Arg Gly Pro Arg Arg Trp
            20                  25                  30

Arg Arg Ala Ala Gly Ala Ala Val Leu Leu Val Glu Met Leu Glu
        35                  40                  45

Arg Ala Ala Phe Phe Gly Val Thr Ala Asn Leu Val Leu Tyr Leu
    50                  55                  60

Asn Ser Thr Asn Phe Asn Trp Thr Gly Glu Gln Ala Thr Arg Ala
65                  70                  75

Ala Leu Val Phe Leu Gly Ala Ser Tyr Leu Leu Ala Pro Val Gly
            80                  85                  90

Gly Trp Leu Ala Asp Val Tyr Leu Gly Arg Tyr Arg Ala Val Ala
        95                  100                 105

Leu Ser Leu Leu Leu Tyr Leu Ala Ala Ser Gly Leu Leu Pro Ala
    110                 115                 120

Thr Ala Phe Pro Asp Gly Arg Ser Ser Phe Cys Gly Glu Met Pro
125                 130                 135

Ala Ser Pro Leu Gly Pro Ala Cys Pro Ser Ala Gly Cys Pro Arg
            140                 145                 150

Ser Ser Pro Ser Pro Tyr Cys Ala Pro Val Leu Tyr Ala Gly Leu
        155                 160                 165

Leu Leu Leu Gly Leu Ala Ala Ser Ser Val Arg Ser Asn Leu Thr
    170                 175                 180

Ser Phe Gly Ala Asp Gln Val Met Asp Leu Gly Arg Asp Ala Thr
185                 190                 195

Arg Arg Phe Phe Asn Trp Phe Tyr Trp Ser Ile Asn Leu Gly Ala
            200                 205                 210

Val Leu Ser Leu Leu Val Val Ala Phe Ile Gln Gln Asn Ile Ser
        215                 220                 225

Phe Leu Leu Gly Tyr Ser Ile Pro Val Gly Cys Val Gly Leu Ala
    230                 235                 240

Phe Phe Ile Phe Leu Phe Ala Thr Pro Val Phe Ile Thr Lys Pro
245                 250                 255

Pro Met Gly Ser Gln Val Ser Ser Met Leu Lys Leu Ala Leu Gln
            260                 265                 270

Asn Cys Cys Pro Gln Leu Trp Gln Arg His Ser Ala Arg Asp Arg
        275                 280                 285

Gln Cys Ala Arg Val Leu Ala Asp Glu Arg Ser Pro Gln Pro Gly
    290                 295                 300

Ala Ser Pro Gln Glu Asp Ile Ala Asn Phe Gln Val Leu Val Lys
305                 310                 315

Ile Leu Pro Val Met Val Thr Leu Val Pro Tyr Trp Met Val Tyr
            320                 325                 330

Phe Gln Met Gln Ser Thr Tyr Val Leu Gln Gly Leu His Leu His
        335                 340                 345

Ile Pro Asn Ile Phe Pro Ala Asn Pro Ala Asn Ile Ser Val Ala
    350                 355                 360

Leu Arg Ala Gln Gly Ser Ser Tyr Thr Ile Pro Glu Ala Trp Leu
365                 370                 375

Leu Leu Ala Asn Val Val Val Val Leu Ile Leu Val Pro Leu Lys
            380                 385                 390
```

```
Asp Arg Leu Ile Asp Pro Leu Leu Leu Arg Cys Lys Leu Leu Pro
                395                 400                 405

Ser Ala Leu Gln Lys Met Ala Leu Gly Met Phe Phe Gly Phe Thr
                410                 415                 420

Ser Val Ile Val Ala Gly Val Leu Glu Met Glu Arg Leu His Tyr
                425                 430                 435

Ile His His Asn Glu Thr Val Ser Gln Gln Ile Gly Glu Val Leu
                440                 445                 450

Tyr Asn Ala Ala Pro Leu Ser Ile Trp Trp Gln Ile Pro Gln Tyr
                455                 460                 465

Leu Leu Ile Gly Ile Ser Glu Ile Phe Ala Ser Ile Pro Gly Leu
                470                 475                 480

Glu Phe Ala Tyr Ser Glu Ala Pro Arg Ser Met Gln Gly Ala Ile
                485                 490                 495

Met Gly Ile Phe Phe Cys Leu Ser Gly Val Gly Ser Leu Leu Gly
                500                 505                 510

Ser Ser Leu Val Ala Leu Leu Ser Leu Pro Gly Gly Trp Leu His
                515                 520                 525

Cys Pro Lys Asp Phe Gly Asn Ile Asn Asn Cys Arg Met Asp Leu
                530                 535                 540

Tyr Phe Phe Leu Leu Ala Gly Ile Gln Ala Val Thr Ala Leu Leu
                545                 550                 555

Phe Val Trp Ile Ala Gly Arg Tyr Glu Arg Ala Ser Gln Gly Pro
                560                 565                 570

Ala Ser His Ser Arg Phe Ser Arg Asp Arg Gly
                575                 580
```

<210> 154
<211> 1266
<212> PRT
<213> Homo sapiens

<400> 154

```
Met Trp Val Asn Pro Glu Glu Val Leu Leu Ala Asn Ala Leu Trp
  1                 5                  10                  15

Ile Thr Glu Arg Ala Asn Pro Tyr Phe Ile Leu Gln Arg Arg Lys
                 20                  25                  30

Gly His Ala Gly Asp Gly Gly Gly Gly Gly Gly Leu Ala Gly Leu
                 35                  40                  45

Leu Val Gly Thr Leu Asp Val Val Leu Asp Ser Ser Ala Arg Val
                 50                  55                  60

Ala Pro Tyr Arg Ile Leu Tyr Gln Thr Pro Asp Ser Leu Val Tyr
                 65                  70                  75

Trp Thr Ile Ala Cys Gly Gly Ser Arg Lys Glu Ile Thr Glu His
                 80                  85                  90

Trp Glu Trp Leu Glu Gln Asn Leu Leu Gln Thr Leu Ser Ile Phe
                 95                 100                 105

Glu Asn Glu Asn Asp Ile Thr Thr Phe Val Arg Gly Lys Ile Gln
                110                 115                 120

Gly Ile Ile Ala Glu Tyr Asn Lys Ile Asn Asp Val Lys Glu Asp
                125                 130                 135

Asp Asp Thr Glu Lys Phe Lys Glu Ala Ile Val Lys Phe His Arg
                140                 145                 150
```

Leu Phe Gly Met Pro Glu Glu Glu Lys Leu Val Asn Tyr Tyr Ser
155 160 165

Cys Ser Tyr Trp Lys Gly Lys Val Pro Arg Gln Gly Trp Met Tyr
170 175 180

Leu Ser Ile Asn His Leu Cys Phe Tyr Ser Phe Leu Met Gly Arg
185 190 195

Glu Ala Lys Leu Val Ile Arg Trp Val Asp Ile Thr Gln Leu Glu
200 205 210

Lys Asn Ala Thr Leu Leu Leu Pro Asp Val Ile Lys Val Ser Thr
215 220 225

Arg Ser Ser Glu His Phe Phe Ser Val Phe Leu Asn Ile Asn Glu
230 235 240

Thr Phe Lys Leu Met Glu Gln Leu Ala Asn Ile Ala Met Arg Gln
245 250 255

Leu Leu Asp Asn Glu Gly Phe Glu Gln Asp Arg Ser Leu Pro Lys
260 265 270

Leu Lys Arg Lys Ser Pro Lys Lys Val Ser Ala Leu Lys Arg Asp
275 280 285

Leu Asp Ala Arg Ala Lys Ser Glu Arg Tyr Arg Ala Leu Phe Arg
290 295 300

Leu Pro Lys Asp Glu Lys Leu Asp Gly His Thr Asp Cys Thr Leu
305 310 315

Trp Thr Pro Phe Asn Lys Met His Ile Leu Gly Gln Met Phe Val
320 325 330

Ser Thr Asn Tyr Ile Cys Phe Thr Ser Lys Glu Glu Asn Leu Cys
335 340 345

Ser Leu Ile Ile Pro Leu Arg Glu Val Thr Ile Val Glu Lys Ala
350 355 360

Asp Ser Ser Ser Val Leu Pro Ser Pro Leu Ser Ile Ser Thr Arg
365 370 375

Asn Arg Met Thr Phe Leu Phe Ala Asn Leu Lys Asp Arg Asp Phe
380 385 390

Leu Val Gln Arg Ile Ser Asp Phe Leu Gln Gln Thr Thr Ser Lys
395 400 405

Ile Tyr Ser Asp Lys Glu Phe Ala Gly Ser Tyr Asn Ser Ser Asp
410 415 420

Asp Glu Val Tyr Ser Arg Pro Ser Ser Leu Val Ser Ser Ser Pro
425 430 435

Gln Arg Ser Thr Ser Ser Asp Ala Asp Gly Glu Arg Gln Phe Asn
440 445 450

Leu Asn Gly Asn Ser Val Pro Thr Ala Thr Gln Thr Leu Met Thr
455 460 465

Met Tyr Arg Arg Arg Ser Pro Glu Glu Phe Asn Pro Lys Leu Ala
470 475 480

Lys Glu Phe Leu Lys Glu Gln Ala Trp Lys Ile His Phe Ala Glu
485 490 495

Tyr Gly Gln Gly Ile Cys Met Tyr Arg Thr Glu Lys Thr Arg Glu
500 505 510

Leu Val Leu Lys Gly Ile Pro Glu Ser Met Arg Gly Glu Leu Trp
515 520 525

Leu Leu Leu Ser Gly Ala Ile Asn Glu Lys Ala Thr His Pro Gly

530 535 540

Tyr Tyr Glu Asp Leu Val Glu Lys Ser Met Gly Lys Tyr Asn Leu
545 550 555

Ala Thr Glu Glu Ile Glu Arg Asp Leu His Arg Ser Leu Pro Glu
560 565 570

His Pro Ala Phe Gln Asn Glu Met Gly Ile Ala Ala Leu Arg Arg
575 580 585

Val Leu Thr Ala Tyr Ala Phe Arg Asn Pro Asn Ile Gly Tyr Cys
590 595 600

Gln Ala Met Asn Ile Val Thr Ser Val Leu Leu Leu Tyr Ala Lys
605 610 615

Glu Glu Glu Ala Phe Trp Leu Leu Val Ala Leu Cys Glu Arg Met
620 625 630

Leu Pro Asp Tyr Tyr Asn Thr Arg Val Val Gly Ala Leu Val Asp
635 640 645

Gln Gly Val Phe Glu Glu Leu Ala Arg Asp Tyr Val Pro Gln Leu
650 655 660

Tyr Asp Cys Met Gln Asp Leu Gly Val Ile Ser Thr Ile Ser Leu
665 670 675

Ser Trp Phe Leu Thr Leu Phe Leu Ser Val Met Pro Phe Glu Ser
680 685 690

Ala Val Val Val Val Asp Cys Phe Phe Tyr Glu Gly Ile Lys Val
695 700 705

Ile Phe Gln Leu Ala Leu Ala Val Leu Asp Ala Asn Val Asp Lys
710 715 720

Leu Leu Asn Cys Lys Asp Asp Gly Glu Ala Met Thr Val Leu Gly
725 730 735

Arg Tyr Leu Asp Ser Val Thr Asn Lys Asp Ser Thr Leu Pro Pro
740 745 750

Ile Pro His Leu His Ser Leu Leu Ser Asp Asp Val Glu Pro Tyr
755 760 765

Pro Glu Val Asp Ile Phe Arg Leu Ile Arg Thr Ser Tyr Glu Lys
770 775 780

Phe Gly Thr Ile Arg Ala Asp Leu Ile Glu Gln Met Arg Phe Lys
785 790 795

Gln Arg Leu Lys Val Ile Gln Thr Leu Glu Asp Thr Thr Lys Arg
800 805 810

Asn Val Val Arg Thr Ile Val Thr Glu Thr Ser Phe Thr Ile Asp
815 820 825

Glu Leu Glu Glu Leu Tyr Ala Leu Phe Lys Ala Glu His Leu Thr
830 835 840

Ser Cys Tyr Trp Gly Gly Ser Ser Asn Ala Leu Asp Arg His Asp
845 850 855

Pro Ser Leu Pro Tyr Leu Glu Gln Tyr Arg Ile Asp Phe Glu Gln
860 865 870

Phe Lys Gly Met Phe Ala Leu Leu Phe Pro Trp Ala Cys Gly Thr
875 880 885

His Ser Asp Val Leu Ala Ser Arg Leu Phe Gln Leu Leu Asp Glu
890 895 900

Asn Gly Asp Ser Leu Ile Asn Phe Arg Glu Phe Val Ser Gly Leu
905 910 915

Ser Ala Ala Cys His Gly Asp Leu Thr Glu Lys Leu Lys Leu Leu
920 925 930

Tyr Lys Met His Val Leu Pro Glu Pro Ser Ser Asp Gln Asp Glu
935 940 945

Pro Asp Ser Ala Phe Glu Ala Thr Gln Tyr Phe Phe Glu Asp Ile
950 955 960

Thr Pro Glu Cys Thr His Val Val Gly Leu Asp Ser Arg Ser Lys
965 970 975

Gln Gly Ala Asp Asp Gly Phe Val Thr Val Ser Leu Lys Pro Asp
980 985 990

Lys Gly Lys Arg Ala Asn Ser Gln Glu Asn Arg Asn Tyr Leu Arg
995 1000 1005

Leu Trp Thr Pro Glu Asn Lys Ser Lys Ser Lys Asn Ala Lys Asp
1010 1015 1020

Leu Pro Lys Leu Asn Gln Gly Gln Phe Ile Glu Leu Cys Lys Thr
1025 1030 1035

Met Tyr Asn Met Phe Ser Glu Asp Pro Asn Glu Gln Glu Leu Tyr
1040 1045 1050

His Ala Thr Ala Ala Val Thr Ser Leu Leu Leu Glu Ile Gly Glu
1055 1060 1065

Val Gly Lys Leu Phe Val Ala Gln Pro Ala Lys Glu Gly Gly Ser
1070 1075 1080

Gly Gly Ser Gly Pro Ser Cys His Gln Gly Ile Pro Gly Val Leu
1085 1090 1095

Phe Pro Lys Lys Gly Pro Gly Gln Pro Tyr Val Val Glu Ser Val
1100 1105 1110

Glu Pro Leu Pro Ala Ser Leu Ala Pro Asp Ser Glu Glu His Ser
1115 1120 1125

Leu Gly Gly Gln Met Glu Asp Ile Lys Leu Glu Asp Ser Ser Pro
1130 1135 1140

Arg Asp Asn Gly Ala Cys Ser Ser Met Leu Ile Ser Asp Asp Asp
1145 1150 1155

Thr Lys Asp Asp Ser Ser Met Ser Ser Tyr Ser Val Leu Ser Ala
1160 1165 1170

Gly Ser His Glu Glu Asp Lys Leu His Cys Glu Asp Ile Gly Glu
1175 1180 1185

Asp Thr Val Leu Val Arg Ser Gly Gln Gly Thr Ala Ala Leu Pro
1190 1195 1200

Arg Ser Thr Ser Leu Asp Arg Asp Trp Ala Ile Thr Phe Glu Gln
1205 1210 1215

Phe Leu Ala Ser Leu Leu Thr Glu Pro Ala Leu Val Lys Tyr Phe
1220 1225 1230

Asp Lys Pro Val Cys Met Met Ala Arg Ile Thr Ser Ala Lys Asn
1235 1240 1245

Ile Arg Met Met Gly Lys Pro Leu Thr Ser Ala Ser Asp Tyr Glu
1250 1255 1260

Ile Ser Ala Met Ser Gly
1265

<210> 155
<211> 21
<212> DNA

<213> Homo sapiens

<400> 155
aaggaagagg cgcgggagaa g 21

<210> 156
<211> 21
<212> DNA
<213> Homo sapiens

<400> 156
ggaagaggcg cgggagaagt t 21

<210> 157
<211> 21
<212> DNA
<213> Homo sapiens

<400> 157
cuucucccgc gccucuucct t 21

<210> 158
<211> 21
<212> DNA
<213> Homo sapiens

<400> 158
cuggacuucc agaagaacat t 21

<210> 159
<211> 18
<212> DNA
<213> Homo sapiens

<400> 159
uguucuggaa guccagtt 18

1589 X 5770

From the
INTERNATIONAL SEARCHING AUTHORITY

To:

see form PCT/ISA/220

PCT

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY
(PCT Rule 43*bis*.1)

Date of mailing *(day/month/year)* see form PCT/ISA/210 (second sheet)	
Applicant's or agent's file reference see form PCT/ISA/220	FOR FURTHER ACTION See paragraph 2 below

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US2005/031798	07.09.2005	08.09.2004

International Patent Classification (IPC) or both national classification and IPC
INV. C07K16/30 A61K47/48 A61K39/395 A61P35/00

Applicant
GENENTECH, INC.

1. This opinion contains indications relating to the following items:

- ☒ Box No. I Basis of the opinion
- ☐ Box No. II Priority
- ☒ Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- ☒ Box No. IV Lack of unity of invention
- ☒ Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- ☐ Box No. VI Certain documents cited
- ☐ Box No. VII Certain defects in the international application
- ☐ Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will usually be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

3. For further details, see notes to Form PCT/ISA/220.

Name and mailing address of the ISA:	Date of completion of this opinion	Authorized Officer
European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	see form PCT/ISA/210	Bayer, A Telephone No. +49 89 2399-7103

159P5438

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US2005/031798

Box No. I Basis of the opinion

1. With regard to the language, this opinion has been established on the basis of:

   ☒ the international application in the language in which it was filed

   ☐ a translation of the international application into , which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1 (b)).

2. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, this opinion has been established on the basis of:

   a. type of material:

      ☒ a sequence listing

      ☐ table(s) related to the sequence listing

   b. format of material:

      ☒ on paper

      ☒ in electronic form

   c. time of filing/furnishing:

      ☐ contained in the international application as filed.

      ☐ filed together with the international application in electronic form.

      ☒ furnished subsequently to this Authority for the purposes of search.

3. ☒ In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

4. Additional comments:

   see separate sheet

1591
+5a

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US2005/031798

Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non obvious), or to be industrially applicable have not been examined in respect of

☐ the entire international application

☒ claims Nos. 1-21, 51-54, 58 (all partially); 28, 29, 38, 47-50, 55, 56 (all complete); 33-37, 58-61 (regarding IA)

because:

☒ the said international application, or the said claims Nos. 33-37, 58-61 (regarding IA) relate to the following subject matter which does not require an international search *(specify)*:

see separate sheet

☒ the description, claims or drawings *(indicate particular elements below)* or said claims Nos. 58 (partially) are so unclear that no meaningful opinion could be formed *(specify)*:

see separate sheet

☒ the claims, or said claims Nos. 58 (partially) are so inadequately supported by the description that no meaningful opinion could be formed *(specify)*:

see separate sheet

☒ no international search report has been established for the whole application or for said claims Nos. 1-21, 51-54 (all partially); 28, 29, 38, 47-50, 55, 56 (all complete)

☐ a meaningful opinion could not be formed without the sequence listing; the applicant did not, within the prescribed time limit:

- ☐ furnish a sequence listing on paper complying with the standard provided for in Annex C of the Administrative Instructions, and such listing was not available to the International Searching Authority in a form and manner acceptable to it.
- ☐ furnish a sequence listing in electronic form complying with the standard provided for in Annex C of the Administrative Instructions, and such listing was not available to the International Searching Authority in a form and manner acceptable to it.
- ☐ pay the required late furnishing fee for the furnishing of a sequence listing in response to an invitation under Rules 13*ter*.1(a) or (b).

☐ a meaningful opinion could not be formed without the tables related to the sequence listings; the applicant did not, within the prescribed time limit, furnish such tables in electronic form complying with the technical requirements provided for in Annex C-*bis* of the Administrative Instructions, and such tables were not available to the International Searching Authority in a form and manner acceptable to it.

☐ the tables related to the nucleotide and/or amino acid sequence listing, if in electronic form only, do not comply with the technical requirements provided for in Annex C-*bis* of the Administrative Instructions.

☐ See Supplemental Box for further details

1592 X50

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US2005/031798

Box No. IV Lack of unity of invention

1. ☒ In response to the invitation (Form PCT/ISA/206) to pay additional fees, the applicant has, within the applicable time limit:

   ☒ paid additional fees

   ☐ paid additional fees under protest and, where applicable, the protest fee

   ☐ paid additional fees under protest but the applicable protest fee was not paid

   ☐ not paid additional fees

2. ☐ This Authority found that the requirement of unity of invention is not complied with and chose not to invite the applicant to pay additional fees.

3. This Authority considers that the requirement of unity of invention in accordance with Rule 13.1, 13.2 and 13.3 is

   ☐ complied with

   ☒ not complied with for the following reasons:

   see separate sheet

4. Consequently, this report has been established in respect of the following parts of the international application:

   ☐ all parts.

   ☒ the parts relating to claims Nos. 1-21, 51-54 (all partially), 22-46, 57-61 (all complete)

Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	16-21
	No:	Claims	1-15, 22-27,30-37,39-46,51-54,58-61
Inventive step (IS)	Yes:	Claims	
	No:	Claims	1-27, 30-37,39-46,51-54,58-61
Industrial applicability (IA)	Yes:	Claims	1-27,30-32,39-46, 51-54,57
	No:	Claims	—

2. Citations and explanations

   see separate sheet

1593 X50

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY (SEPARATE SHEET)

International application No.
PCT/US2005/031798

Re Item I

1. Sequence listings (pages 1-303) corresponding to figures 1-154 as filed are considered to be part of the application and will be considered for the examination.

Re Item III

1. The international search report has been drawn up in respect to only part of the claims of the international application as filed. Since only claims searched will be the subject of the international preliminary examination (Rule 66.1(e) PCT), this written opinion is based on claims 1-21, 51-54, 58 (all partially) and claims 22-27, 30-37, 39-46, 57, 59-61 (all complete) (see also Item IV below).

2. Claims 58-61 and claims 33-37 as far as in vivo is concerned relate to subject-matter considered by this Authority to be covered by the provisions of Rule 67.1(iv) PCT. Consequently, no opinion will be formulated with respect to the industrial applicability of the subject-matter of these claims (Article 34(4)(a)(I) PCT).

3. Present claim 58 (partially) encompass compounds defined only by their desired function, contrary to the requirements of clarity of Article 6 PCT, because the result-to-be-achieved type of definition does not allow the scope of the claim to be ascertained. Arguing that any compound could be screened does not overcome this objection, as the skilled person would not have knowledge beforehand as to whether it would fall within the scope claimed, except for the antibodies against TAT188 disclosed in the description. It is therefore considered that an equally unquantifiable and thus unreasonable amount of experimentation would be required to screen peptides randomly, imposing a severe and undue burden on all those wishing to ascertain the scope of the claims, which is not in compliance with the clarity requirements of Article 6 PCT. Furthermore, the description does not provide support and disclosure in the sense of Articles 5 and 6 PCT for all imaginable compounds having the desired function. This noncompliance with the substantive provisions is to such an extent, that the search was performed taking into consideration the noncompliance in determining the extent of the search for said claim (PCT Guidelines, 9.19 and 9.24).

The search of said claim was consequently restricted to antibodies as defined in present claim 59.

Accordingly, this written opinion is based on the subject-matter of claim 58 only with regard to antibodies.

4. Present claims 22, 23, 26, 27, 30, 33, 39, 44, 45, 58 and 59 in which the polypeptide is designated by its internal denomination, i.e. TAT188, are not clear (Article 6 PCT). The polypeptide should be defined by its sequence.

Re Item IV

The separate inventions are:

Inventions 1-74: Claims 1-21, 51-54 (all partially)
Each of the antibodies binding to the polypeptides defined by SEQ-ID NOs: 79-114 and 116-154 and the related subject-matter define one independent invention.

Invention 75: Claims 1-21, 51-54 (all partially); claims 22-46, 57-61 (all complete)
An isolated antibody binding to the polypeptide defined by SEQ-ID NO: 115, i.e. TAT188 and the related subject-matter.

Invention 76: Claims 47-50, 55, 56 (all complete); claims 51-54 (all partially)
TAT188 siRNA and the related subject-matter.

They are not so linked as to form a single general inventive concept (Rule 13.1 PCT) for the following reasons:

The present application relates to the identification of various TAT (Tumour-associated Antigenic Target) polypeptides defined by SEQ-ID NOs: 79-154, wherein SEQ-ID NO: 115 corresponds to TAT188, and discloses isolated antibodies against such polypeptides as well as uses thereof. Thus, the problem underlying the present application can be seen as the provision of antibodies against TAT polypeptides.
Additionally, the present application discloses TAT188 binding interfering RNA (siRNA)

and uses thereof. Thus, another problem underlying the present application can be seen as the provision of TAT siRNA.

The technical relationship and thus common concept within the meaning of Rule 13.2 PCT a priori linking the antibodies is considered as all relating to TAT polypeptides. The technical relationship and thus common concept within the meaning of Rule 13.2 PCT a priori linking the anti-TAT188 antibody and the TAT188 siRNA is considered as both relating to TAT188.

The prior art, however, already discloses said common concepts. WO03000113 (D1) discloses the identical TAT polypeptides of the present application, i.e. SEQ-ID NOs: 79-154 including TAT188 as well as antibodies against such polypeptides.

Thus, in view of the prior art, the problem underlying the present application can be formulated as the provision of alternative TAT modulators. All inventions listed above, represent an independent solution concerning the foregoing problem of the underlying application.

Since no other technical feature can be distinguished which, in the light of the prior art could be regarded as special common technical feature, the ISA is of the opinion that there is no single inventive concept (Rule 13.1 PCT) underlying the plurality of different inventions in the sense of Rule 13.2 PCT.

Consequently there is lack of unity and the different inventions not belonging to a common inventive concept are formulated as different inventions pursuant to Article 17(3)(a) PCT.

Since the applicant did not pay search fees for all of the inventions, an opinion will only be formulated for the subject-matter of the searched inventions, i.e. invention 1 (an isolated antibody binding to ta polypeptide defined by SEQ-ID NO: 79 and the related subject-matter) and 75 (an isolated antibody binding to the polypeptide defined by SEQ-ID NO:115, i.e. TAT188 and the related subject-matter).

Re Item V

1596 1 Sheet

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY (SEPARATE SHEET)

International application No.
PCT/US2005/031798

1. Reference is made to the following documents:

D1: WO 03/000113 A (GENENTECH, INC) 3 January 2003 (2003-01-03)

D2: HENDERSON MARILYN ET AL: "Microfibril-associated glycoprotein-1 (MAGP-1) is specifically located on the beads of the beaded-filament structure for fibrillin-containing microfibrils as visualized by the rotary shadowing technique" JOURNAL OF HISTOCHEMISTRY AND CYTOCHEMISTRY, vol. 44, no. 12, 1996, pages 1389-1397, XP002379484 ISSN: 0022-1554

D3: MANNION BRIAN A ET AL: "The light chain of CD98 is identified at E16/TA1 protein" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 273, no. 50, 11 December 1998 (1998-12-11), pages 33127-33129, XP002392000 ISSN: 0021-9258

D4: EP-A-1 111 048 (JAPAN SCIENCE AND TECHNOLOGY CORPORATION; JAPAN SCIENCE AND TECHNOLOGY) 27 June 2001 (2001-06-27)

D5: DORONINA S O ET AL: "Development of potent monoclonal antibody auristatin conjugates for cancer therapy" NATURE BIOTECHNOLOGY, NATURE PUBLISHING, US, vol. 21, no. 7, July 2003 (2003-07), pages 778-784, XP002280966 ISSN: 1067-0156 cited in the application

D6: DORONINA SVETLANA ET AL: "Immunoconjugates comprised of drugs with impaired cellular permeability: A new approach to targeted therapy." ABSTRACTS OF PAPERS AMERICAN CHEMICAL SOCIETY, vol. 228, no. Part 1, August 2004 (2004-08), page U908, XP009066098 & MEETING OF THE DIVISION OF CHEMICAL TOXICOLOGY OF THE AMERICAN-CHEMICAL-SOCIETY HELD AT THE 228TH N; PHILADELPHIA, PA, USA; AUGUST 22 -26, 2004 ISSN: 0065-7727

As not otherwise indicated, reference is made to the passages cited in the International Search Report.

2. Novelty

The present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of claims 1-15, 22-27, 30-37, 39-46, 51-54 and 58-61 is not new in the sense of Article 33(2) PCT.

Document D1 discloses the identical subject-matter as the present application, including SEQ-ID NOs: 1, 79, 37 and 115 and antibodies against in general as well as uses thereof. Only missing from D1 are the toxins and linkers specified in dependent claims 16-21 and the specific antibodies 3B5.1, 12B9.1 and 12G12.1. Accordingly, in view of D1, the subject-matter of present claims 1-15, 22-27, 30-37, 39-46, 51-54 and 58-61 is not novel.

Document D2 discloses the preparation and use of MAGP-1 monoclonal antibody (mAb 11B). Since MAGP-1 is synonymous to MFAP2 (microfibrillar-associated protein 2) which is identical to TAT161 presented by SEQ-ID 79 of the present application, D2 anticipates the novelty of present claims 1-3.

Document D3 discloses monoclonal antibodies against E16 which is identical to TAT188 and their use in immunoblots of colon carcinomas compared to normal colon cells. Since any antibody against E16 inherently has the functional features of present claims 23, 26, 27 and 30, claims 1-3, 22-27, 30-32 and 39-46 are not novel in view of D3.

Document D4 discloses human LAT1 (L-amino acid transporter 1 which is identical to TAT188) and antibodies against as well as their use in therapy, diagnosis and quantification of LAT1. Accordingly, in view of D4, the subject-matter of present claims 1-5, 12-15, 22-27, 30-37, 39-46, 51-54 and 58-61 is not novel.

For the specific anti-TAT188 antibodies 3B5.1, 12B9.1 and 12G12.1 presented by their ATCC Accession Numbers novelty can be acknowledged since none of the prior art available discloses such antibodies (Article 33(2) PCT).

3. Inventive step

The present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of claims 1-27, 30-37, 39-46, 51-54, 57-61 does not involve an inventive step in the sense of Article 33(3) PCT.

3.1 Taking D1 as the closest prior art, the problem underlying the present claims 16-21

may be considered as the provision of alternative immunoconjugates, the solution being the use of the toxins and carriers recited in claims 16-21.

However such toxins and carriers have already been employed for the same purpose, i.e. for the preparation of mAb immunoconjugates, see, e.g. D5 and D6. It would therefore be obvious to the person skilled in the art, to apply these compounds with corresponding effect to the antibody of D1, thereby arriving at the subject-matter of present claims 16-21.

3.2 The antibody of present claims 1-27, 30-32 is only defined by its binding to TAT188 without giving any technical feature of the antibody. Since an antibody against a known protein represents the outcome of routine experimentation that does not involve inventive skill, said claims are not considered as inventive.

3.3 The specific antibodies produced by hybridomas 3B5.1, 12B9.1 and 12G12.1 as well as their uses are not considered as inventive either.

Taking D1 as the closest prior art the problem to be solved by the present application was to provide monoclonal antibodies against T188. D1, however very detailed teaches how to obtain such antibodies using methods well-known to the skilled person as well as uses thereof and thus the generation and uses of said antibodies do not involve inventive skill.

Taking D3 as the closest prior art the problem to be solved by the present application was to provide alternative monoclonal antibodies against T188. However, the generation of monoclonal antibodies against a known protein represents the outcome of routine experimentation not involving inventive skill. Additionally, there seems to be no surprising effect/feature of the specific antibodies of the present application in comparison to the antibodies disclosed in D3.

Thus, no inventive step for claims 1-27, 30-37, 39-46, 51-54 and 57-61 can be acknowledged.

4. For the assessment of the present claims 33-37 and 58-61 on the question whether

1599 158X

they are industrially applicable, no unified criteria exist in the PCT Contracting States. The patentability can also be dependent upon the formulation of the claims. The EPO, for example, does not recognize as industrially applicable the subject-matter of claims to the use of a compound in medical treatment, but may allow, however, claims to a known compound for first use in medical treatment and the use of such a compound for the manufacture of a medicament for a new medical treatment.

1600

PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION CONCERNING
SUBMISSION OR TRANSMITTAL
OF PRIORITY DOCUMENT

(PCT Administrative Instructions, Section 411)

To:

CONLEY, Deirdre
MS49 1 DNA Way
South San Francisco, CA 94080-4990
ETATS-UNIS D'AMERIQUE

Date of mailing *(day/month/year)* 18 November 2005 (18.11.2005)	
Applicant's or agent's file reference P5001R1P1	IMPORTANT NOTIFICATION
International application No. PCT/US2005/031798	International filing date *(day/month/year)* 07 September 2005 (07.09.2005)
International publication date *(day/month/year)*	Priority date *(day/month/year)* 08 September 2004 (08.09.2004)
Applicant GENENTECH, INC. et al	

1. By means of this Form, which replaces any previously issued notification concerning submission or transmittal of priority documents, the applicant is hereby notified of the date of receipt by the International Bureau of the priority document(s) relating to all earlier application(s) whose priority is claimed. Unless otherwise indicated by the letters "NR", in the right-hand column or by an asterisk appearing next to a date of receipt, the priority document concerned was submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b).

2. *(If applicable)* The letters "NR" appearing in the right-hand column denote a priority document which, on the date of mailing of this Form, had not yet been received by the International Bureau under Rule 17.1(a) or (b). Where, under Rule 17.1(a), the priority document must be submitted by the applicant to the receiving Office or the International Bureau, but the applicant fails to submit the priority document within the applicable time limit under that Rule, the attention of the applicant is directed to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.

3. *(If applicable)* An asterisk (*) appearing next to a date of receipt, in the right-hand column, denotes a priority document submitted or transmitted to the International Bureau but not in compliance with Rule 17.1(a) or (b) (the priority document was received after the time limit prescribed in Rule 17.1(a) or the request to prepare and transmit the priority document was submitted to the receiving Office after the applicable time limit under Rule 17.1(b)). Even though the priority document was not furnished in compliance with Rule 17.1(a) or (b), the International Bureau will nevertheless transmit a copy of the document to the designated Offices, for their consideration. In case such a copy is not accepted by the designated Office as the priority document, Rule 17.1(c) provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.

Priority date	Priority application No.	Country or regional Office or PCT receiving Office	Date of receipt of priority document
08 September 2004 (08.09.2004)	10/936,628	US	07 November 2005 (07.11.2005)

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland Facsimile No. +41 22 338 82 70	Authorized officer SNELL Matthew Facsimile No. +41 22 338 7090 Telephone No. +41 22 338 7117

Form PCT/IB/304 (January 2004)

1601 2584

P5001R1P1

PCT REQUEST

Original (for SUBMISSION)

0	For receiving Office use only	
0-1	International Application No.	
0-2	International Filing Date	
0-3	Name of receiving Office and "PCT International Application"	

0-4	Form PCT/RO/101 PCT Request	
0-4-1	Prepared Using	PCT-SAFE [EASY mode] Version 3.50 (Build 0002.173)
0-5	Petition The undersigned requests that the present international application be processed according to the Patent Cooperation Treaty	
0-6	Receiving Office (specified by the applicant)	United States Patent and Trademark Office (USPTO) (RO/US)
0-7	Applicant's or agent's file reference	P5001R1P1
I	Title of invention	COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF TUMOR
II	Applicant	
II-1	This person is	applicant only
II-2	Applicant for	all designated States except US
II-4	Name	GENENTECH, INC.
II-5	Address	MS 49 1 DNA Way South San Francisco, California 94080-4990 United States of America
II-6	State of nationality	US
II-7	State of residence	US
II-8	Telephone No.	650-225-1000
II-9	Facsimile No.	650-952-9881

1602 1560

P5001R1P1

PCT REQUEST

Original (for SUBMISSION)

III-1	Applicant and/or Inventor	
III-1-1	This person is	applicant and inventor
III-1-2	Applicant for	US only
III-1-4	Name (LAST, First)	CAIRNS, Belinda
III-1-5	Address	1229 Paloma Avenue Burlingame, California 94010 United States of America
III-1-6	State of nationality	GB
III-1-7	State of residence	US
III-2	Applicant and/or Inventor	
III-2-1	This person is	applicant and inventor
III-2-2	Applicant for	US only
III-2-4	Name (LAST, First)	CHEN, Ruihuan
III-2-5	Address	1110 Polynesia Drive #301 Foster City, California 94404 United States of America
III-2-6	State of nationality	CN
III-2-7	State of residence	US
III-3	Applicant and/or Inventor	
III-3-1	This person is	applicant and inventor
III-3-2	Applicant for	US only
III-3-4	Name (LAST, First)	CHUNTHARAPAI, Anan
III-3-5	Address	826 Ellis Drive Colma, California 94015 United States of America
III-3-6	State of nationality	US
III-3-7	State of residence	US
III-4	Applicant and/or Inventor	
III-4-1	This person is	applicant and inventor
III-4-2	Applicant for	US only
III-4-4	Name (LAST, First)	FRANTZ, Gretchen
III-4-5	Address	135 San Benito Way San Francisco, California 94127 United States of America
III-4-6	State of nationality	US
III-4-7	State of residence	US

III-5	Applicant and/or inventor	
III-5-1	This person is	applicant and inventor
III-5-2	Applicant for	US only
III-5-4	Name (LAST, First)	HILLAN, Kenneth
III-5-5	Address	64 Seward Street San Francisco, California 94114 United States of America
III-5-6	State of nationality	GB
III-5-7	State of residence	US
III-6	Applicant and/or inventor	
III-6-1	This person is	applicant and inventor
III-6-2	Applicant for	US only
III-6-4	Name (LAST, First)	KOEPPEN, Kartmut
III-6-5	Address	1746 Milvia Street Berkeley, California 94709 United States of America
III-6-6	State of nationality	DE
III-6-7	State of residence	US
III-7	Applicant and/or inventor	
III-7-1	This person is	applicant and inventor
III-7-2	Applicant for	US only
III-7-4	Name (LAST, First)	PHILLIPS, Heidi
III-7-5	Address	1511 Walnut Drive Palo Alto, California 94303 United States of America
III-7-6	State of nationality	US
III-7-7	State of residence	US
III-8	Applicant and/or inventor	
III-8-1	This person is	applicant and inventor
III-8-2	Applicant for	US only
III-8-4	Name (LAST, First)	POLAKIS, Paul
III-8-5	Address	1449 Cortez Avenue Burlingame, California 94010 United States of America
III-8-6	State of nationality	US
III-8-7	State of residence	US

1604 1497

P5001R1P1

PCT REQUEST

Original (for SUBMISSION)

III-9	Applicant and/or Inventor	
III-9-1	This person is	applicant and inventor
III-9-2	Applicant for	US only
III-9-4	Name (LAST, First)	REED, Chae Janeka
III-9-5	Address	2396 Post Street San Francisco, California 94115 United States of America
III-9-6	State of nationality	US
III-9-7	State of residence	US
III-10	Applicant and/or Inventor	
III-10-1	This person is	applicant and inventor
III-10-2	Applicant for	US only
III-10-4	Name (LAST, First)	SAKANAKA, Chie
III-10-5	Address	1481 8th Avenue San Francisco, California 94122 United States of America
III-10-6	State of nationality	JP
III-10-7	State of residence	US
III-11	Applicant and/or Inventor	
III-11-1	This person is	applicant and inventor
III-11-2	Applicant for	US only.
III-11-4	Name (LAST, First)	SMITH, Victoria
III-11-5	Address	19 Dwight Road Burlingame, California 94010 United States of America
III-11-6	State of nationality	AU
III-11-7	State of residence	US
III-12	Applicant and/or Inventor	
III-12-1	This person is	applicant and inventor
III-12-2	Applicant for	US only
III-12-4	Name (LAST, First)	SPENCER, Susan
III-12-5	Address	2395 Paradise Drive Tiburon, California 94920 United States of America
III-12-6	State of nationality	US
III-12-7	State of residence	US

1605 ts13

P5001R1P1

PCT REQUEST

Original (for SUBMISSION)

III-13	Applicant and/or inventor	
III-13-1	This person is	applicant and inventor
III-13-2	Applicant for	US only
III-13-4	Name (LAST, First)	WILLIAMS, P. Mickey
III-13-5	Address	509 Alto Avenue Half Moon Bay, California 94019 United States of America
III-13-6	State of nationality	US
III-13-7	State of residence	US
III-14	Applicant and/or inventor	
III-14-1	This person is	applicant and inventor
III-14-2	Applicant for	US only
III-14-4	Name (LAST, First)	WU, Thomas
III-14-5	Address	41 Nevada Street San Francisco, California 94110 United States of America
III-14-6	State of nationality	US
III-14-7	State of residence	US
III-15	Applicant and/or inventor	
III-15-1	This person is	applicant and inventor
III-15-2	Applicant for	US only
III-15-4	Name (LAST, First)	ZHANG, Zemin
III-15-5	Address	876 Taurus Drive Foster City, California 94404 United States of America
III-15-6	State of nationality	US
III-15-7	State of residence	US
IV-1	Agent or common representative; or address for correspondence	
	The person identified below is hereby/ has been appointed to act on behalf of the applicant(s) before the competent International Authorities as:	agent
IV-1-1	Name (LAST, First)	CONLEY, Deirdre
IV-1-2	Address	MS49 1 DNA Way South San Francisco, California 94080-4990 United States of America
IV-1-3	Telephone No.	650-225-2066
IV-1-4	Facsimile No.	650-952-9881
IV-1-5	e-mail	conley@gene.com
IV-2	Additional agent(s)	additional agent(s) with same address as

1606 15 Sept

IV-2-1	Name(s)	first named agent AGARWAL,Atulya; ANDRUS,Alex; BARNES,Elizabeth; CARPENTER,David; COBURN,Cara; CUI,Steven; ELLIOTT,Jennifer; EVANS,David; HASAK,Janet; JOHNSTON,Sean; KALINOWSKI,Grant; KRESNAK,Mark; KUBINEC,Jeffrey; LEE,Wendy; MARSCHANG,Diane; NAIK,Paul; PLEASURE,Irene; SCHWARTZ,Timothy; SHIN,Elinor; SVOBODA,Craig; TAN,Lee; YEUNG,Bonny	
V	DESIGNATIONS		
V-1	The filing of this request constitutes under Rule 4.9(a), the designation of all Contracting States bound by the PCT on the international filing date, for the grant of every kind of protection available and, where applicable, for the grant of both regional and national patents.		
VI-1	Priority claim of earlier national application		
VI-1-1	Filing date	08 September 2004 (08.09.2004)	
VI-1-2	Number	10/936,626	
VI-1-3	Country	US	
VI-2	Priority document request The receiving Office is requested to prepare and transmit to the International Bureau a certified copy of the earlier application(s) identified above as item(s):	VI-1	
VII-1	International Searching Authority Chosen	European Patent Office (EPO) (ISA/EP)	
VIII	Declarations	Number of declarations	
VIII-1	Declaration as to the identity of the inventor	-	
VIII-2	Declaration as to the applicant's entitlement, as at the international filing date, to apply for and be granted a patent	-	
VIII-3	Declaration as to the applicant's entitlement, as at the international filing date, to claim the priority of the earlier application	-	
VIII-4	Declaration of inventorship (only for the purposes of the designation of the United States of America)	-	
VIII-5	Declaration as to non-prejudicial disclosures or exceptions to lack of novelty	-	

1607 ES

P5001R1P1

PCT REQUEST

Original (for SUBMISSION)

IX	Check list	number of sheets	electronic file(s) attached
IX-1	Request (including declaration sheets)	7	✓
IX-2	Description	197	-
IX-3	Claims	6	-
IX-4	Abstract	1	✓
IX-5	Drawings	249	-
IX-7	TOTAL	460	
	Accompanying items	paper document(s) attached	electronic file(s) attached
IX-8	Fee calculation sheet	✓	-
IX-17	PCT-SAFE physical media	-	✓
IX-19	Figure of the drawings which should accompany the abstract	None	
IX-20	Language of filing of the international application	English	
X-1	Signature of applicant, agent or common representative		
X-1-1	Name (LAST, First)	CONLEY, Deirdre	
X-1-2	Name of signatory		
X-1-3	Capacity		

FOR RECEIVING OFFICE USE ONLY

10-1	Date of actual receipt of the purported international application	
10-2	Drawings:	
10-2-1	Received	
10-2-2	Not received	
10-3	Corrected date of actual receipt due to later but timely received papers or drawings completing the purported international application	
10-4	Date of timely receipt of the required corrections under PCT Article 11(2)	
10-5	International Searching Authority	ISA/EP
10-6	Transmittal of search copy delayed until search fee is paid	

FOR INTERNATIONAL BUREAU USE ONLY

11-1	Date of receipt of the record copy by the International Bureau	

1608

P5001R1P1

PCT (ANNEX - FEE CALCULATION SHEET)

Duplicate of Original

(This sheet is not part of and does not count as a sheet of the International application)

0	For receiving Office use only				
0-1	International Application No.				
0-2	Date stamp of the receiving Office				
0-4	Form PCT/RO/101 (Annex) PCT Fee Calculation Sheet				
0-4-1	Prepared Using		PCT-SAFE [EASY mode] Version 3.50 (Build 0002.173)		
0-5	Applicant's or agent's file reference		P5001R1P1		
2	Applicant		GENENTECH, INC.		
12	Calculation of prescribed fees		fee amount/multiplier	Total amounts (USD)	
12-1	Transmittal fee	T	⇨	300	
12-2-1	Search fee	S	⇨	2075	
12-2-2	International search to be carried out by		EP		
12-3	International filing fee (first 30 sheets)	I1	1211		
12-4	Remaining sheets		430		
12-5	Additional amount	(X)	13		
12-6	Total additional amount	I2	5590		
12-7	I1 + I2 =	I	6801		
12-12	EASY Filing reduction	R	-87		
12-13	Total international filing fee (I-R)	I	⇨	6714	
12-14	Fee for priority document Number of priority documents requested		1		
12-15	Fee per document	(X)	20		
12-16	Total priority document fee:	P	⇨	20	
12-17	TOTAL FEES PAYABLE (T+S+I+P)		⇨	9109	
12-19	Mode of payment		authorization to charge deposit account		
12-20	Deposit account instructions The receiving Office		United States Patent and Trademark Office (USPTO) (RO/US)		
12-20-1	Authorization to charge the total fees indicated above		✓		
12-20-2	Authorization to charge any deficiency or credit any overpayment in the total fees indicated above		✓		
12-20-3	Authorization to charge the fee for priority document		✓		
12-21	Deposit account No.		07-0630		
12-22	Date		07 September 2005 (07.09.2005)		
12-23	Name and signature		CONLEY, Deirdre		

PCT

Original (for SUBMISSION)

13-8-7	Validation messages Contents	Yellow! The power of attorney or a copy of the general power of attorney will need to be furnished unless all applicants sign the request form.
13-8-8	Validation messages Payment	Green? Please ensure that you have a valid deposit account with the receiving Office selected.

1610

State of California

SECRETARY OF STATE

APOSTILLE

(Convention de La Haye du 5 octobre 1961)

1. *Country: United States of America*
   This public document
2. *has been signed by AIDA MICLAT GUIAM*
3. *acting in the capacity of Notary Public, State of California*
4. *bears the seal/stamp of AIDA MICLAT GUIAM, Notary Public, State of California*

CERTIFIED

5. *At Sacramento, California*
6. *the 15th day of December 2005*
7. *by Deputy Secretary of State, State of California*
8. *No. 267083*
9. *Seal/Stamp:*

10. *Signature*

Secretary of State

BY

NP-84 A (REV. 4-06)

14 FEB 2007

OSCAR ALARCON NUÑEZ

NOTARIO

1611

CALIFORNIA ALL-PURPOSE ACKNOWLEDGMENT

State of California
County of San Mateo } ss.

On December 6, 2005 (Date) before me, Aida Miclat Guiam (Name and Title of Officer (e.g., "Jane Doe, Notary Public"))

personally appeared Timothy R. Schwartz (Name(s) of Signer(s))

☐ personally known to me
☒ proved to me on the basis of satisfactory evidence

to be the person(s) whose name(s) is/are subscribed to the within instrument and acknowledged to me that he/she/they executed the same in his/her/their authorized capacity(ies), and that by his/her/their signature(s) on the instrument the person(s), or the entity upon behalf of which the person(s) acted, executed the instrument.

AIDA MICLAT GUIAM
Commission # 1455662
Notary Public - California
San Mateo County
My Comm. Expires Dec 9, 2007

WITNESS my hand and official seal.

Aida Miclat Guiam
(Signature of Notary Public)

OPTIONAL

Though the information below is not required by law, it may prove valuable to persons relying on the document and could prevent fraudulent removal and reattachment of this form to another document.

Description of Attached Document

Title or Type of Document: _____

Document Date: _____ Number of Pages: _____

Signer(s) Other Than Named Above: _____

Capacity(ies) Claimed by Signer

Signer's Name: _____

☐ Individual
☐ Corporate Officer — Title(s): _____
☐ Partner — ☐ Limited ☐ General
☐ Attorney-In-Fact
☐ Trustee
☐ Guardian or Conservator
☐ Other: _____

Signer Is Representing: _____

RIGHT THUMBPRINT OF SIGNER
Top of thumb here

© 1999 National Notary Association • 9350 De Soto Ave., P.O. Box 2402 • Chatsworth, CA 91313-2402 • www.nationalnotary.org

14 FEB 2007
OMAR ALARCON NUÑEZ

1612

Brigard & Castro
BOGOTA - COLOMBIA

PODER

Yo (nosotros)

domiciliado(s) en

Nombro(amos) a: **JUAN PABLO CADENA SARMIENTO** y/o **BEATRIZ JARAMILLO** y/o **CARLOS H. PALACIO EASTMAN** y/o **HECTOR GOMEZ MEJIA**, de Bogotá, Colombia, apoderados verdaderos y legales con poder especial, amplio y suficiente, para recabar de las oficinas y autoridades colombianas administrativas, judiciales y de policía, en cualquier instancia o recurso, la obtención de patentes de invención, registros de marcas, diseños industriales, modelos de utilidad, depósitos de nombres comerciales y enseñas, registros de derechos de autor y contratos sobre los mismos, obtención de certificados de obtentor de variedades vegetales, así como sus prórrogas, renovaciones, anotaciones de traspaso, cambios de nombre y/o domicilio; elaborar, tramitar y registrar contratos de licencia y licencias de cualesquiera clases; intervenir como demandante o como demandado en cualquier instancia y recurso ante cualquier juez, corporación, funcionario público o autoridad en acciones judiciales, administrativas y de policía tales como: oposiciones al registro de marcas; acciones de caducidad, cancelación y nulidad; acciones reivindicatorias, acciones derivadas de la obtención o explotación de licencias; acciones de competencia desleal; demandas penales; acciones de indemnización de perjuicios, reclamos y observaciones a solicitudes de patentes, diseños industriales, modelos de utilidad, en el ejercicio de medidas cautelares y en otras acciones o medidas similares.

A este efecto, lo(a) faculto(amos) para elevar solicitudes, formular descripciones, enmiendas, protestas, declaraciones, oposiciones, cancelaciones, nulidades, apelaciones y reclamos; pagar impuestos, cuotas y gravámenes prescritos por la ley; renunciar o desistir a derechos concedidos o en curso de concesión; recibir toda clase de documentos y valores, dando el descargo respectivo, y, en general, para dar ante dichas autoridades todos los pasos necesarios al efecto indicado y tomar todas las medidas que crayeren conducentes al resguardo de mis (nuestros intereses), con todas las demás facultades legales necesarias.

Este poder comprende la facultad de ratificar o confirmar en mi (nuestro) nombre lo actuado así como la facultad de desistir, transigir, comprometer, y la de sustituirlo en todo o parte y revocar sustituciones. Asimismo nombro(amos) a: **JUAN PABLO CADENA SARMIENTO** y/o **BEATRIZ JARAMILLO** y/o **CARLOS H. PALACIO EASTMAN** y/o **HECTOR GOMEZ MEJIA**, domiciliado(s) en la Carrera 7 No. 16-56 Piso 9, mis (nuestros) apoderados en Colombia, con facultades para recibir notificaciones y designar apoderados judiciales o extrajudiciales.

Otorgado y firmado en

hoy de de 200_

Firma

Nombre
Cargo

POWER OF ATTORNEY

I (We) **Genentech, Inc.**
1 DNA Way
domiciled in **South San Francisco, California 94080-4990**
USA

Hereby appoints **JUAN PABLO CADENA SARMIENTO** and/or **BEATRIZ JARAMILLO** and/or **CARLOS H. PALACIO EASTMAN** and/or **HECTOR GOMEZ MEJIA**, of Bogotá, Colombia, my (our) true and lawful attorneys with special, ample and sufficient power to obtain from the Colombian administrative, judicial and police offices and authorities, in any instance, privileges of patents of invention, registration of trademarks, industrial designs and utility models; registration of commercial names and emblems, registration of copyright and contracts related thereto, certificates of obtainers of vegetables varieties, as well as extensions, renewals and recordal of assignments and changes of name and domicile related to the above; to prepare, prosecute and register licenses of any kind; to act as plaintiff or defendant in any instance or recourse, and before any judge, corporation, public official or authority in judicial, administrative and police actions, such as: oppositions to the registration of trademarks; caducity, cancellation and nullity actions; claim actions derived from the obtaining or exploitation of licenses, action of unfair competition; criminal proceedings, actions for payment of damages: claims or observations on patent applications, models, industrial designs and utility models; in the exercise of preventive measures and in other actions or similar proceedings.

To that effect they hereby empowered to file applications, formulate descriptions, amendments, protests, declarations, oppositions, cancellations, nullities, appeals and claims; to pay taxes, quotas and assessments prescribed by law, renounce or waive the rights granted or in the process of granting; to receive all kinds of documents and valuables, giving the respective release of receipt, and, in general, to take before said authorities the necessary steps to such effect, and any measures that they may deem pertinent to the safeguard of my (our) interest(s) with all the other necessary legal powers.

This power includes the faculty to ratify or confirm desist, settle and compromise on my (our) behalf, and to substitute this power in whole or in part and to revoke substitutions. At the same time I (We) hereby appoints **JUAN PABLO CADENA SARMIENTO** and/or **BEATRIZ JARAMILLO** and/or **CARLOS H. PALACIO EASTMAN** and/or **HECTOR GOMEZ MEJIA**, domiciled in Carrera 7 No. 16-56, Piso 9, my (our) attorney(s) in fact, with powers to receive notifications and to designate judicial and extrajudicial attorneys.

Granted and signed in **South San Francisco, California USA**

this **6th** day of **December** 2005

Signature

Name **Timothy R. Schwartz, Ph**
Title **Section Patent Counsel**
Authorized Corporate Signatory
Genentech, Inc.

NOTARIO CUARENTA Y SEIS DEL CIRCULO DE SANTAFE DE BOGOTÁ HAGO CONSTAR QUE ESTA FIRMA POR ESTÁTICA CORRESPONDE CON LA QUE HA TENIDO A LA VISTA EN EL REGISTRO DE SANTAFE DE BOGOTÁ D.C. REPUBLICA DE COLOMBIA
14 FEB 2007
OSCAR ALARCON NUÑEZ
NOTARIA CUARENTA Y SEIS

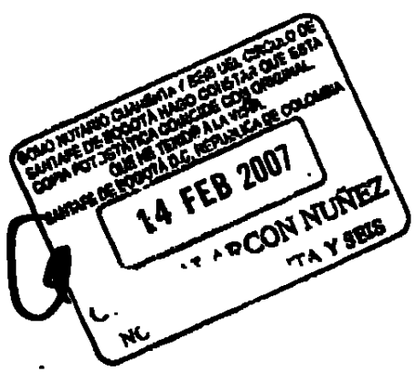
1613

AUTHORIZATION

Genentech, Inc. hereby authorizes any of Janet E. Hasak, Mark T. Kresnak, Jeffrey S. Kubinec and Timothy R. Schwartz, for so long as they remain employees of Genentech, Inc., to execute, on its behalf, any submissions associated with the filing, prosecution, granting, and extension and maintenance of grant of patent and trademark applications in the U.S. and foreign countries, contracts relating to biological deposits with the American Type Culture Collection, and assignments to publishers of copyrights covering works authored by Genentech employees.

August 27, 2002
Dated

Sean Johnston
Sean A. Johnston
Vice President, Intellectual Property
Genentech, Inc.



1614.

TRADUCCION AL ESPAÑOL del Documento de Poder otorgado por **Genentech, Inc.** domiciliada en **1 DNA Way, South San Francisco, California 94080-4990, Estados Unidos de América.**

ESTADO DE CALIFORNIA
SECRETARIA DE ESTADO

APOSTILLA
(Convención de La Haya del 5 de Octubre de 1961)

1 País: Estados Unidos de América
Esta documento público
2. ha sido firmado por AIDA MICLAT GUIAM
3. actuando en capacidad de Notario Público, Estado de California
4. lleva el sello/timbre de AIDA MICLAT GUIAM, Notario Público
Estado de California

CERTIFICADO

5. en Sacramento, California
6. el 15 de diciembre de 2005
7. por el Secretario de Estado Delegado, Estado de California
8. No. 267083
9. Sello/Timbre

10. Firma

(Firma)
Secretario de Estado

NOTARIA CUARENTA Y SEIS DEL CIRCULO DE SANTAFE DE BOGOTA D.C. HAGO CONSTAR QUE ESTA COPIA POR: SU FOTOCOPIA CONCUERDA CON EL ORIGINAL QUE ME FUERA PRESENTADO A LA VISTA.
SANTAFE DE BOGOTA D.C. REPUBLICA DE COLOMBIA
14 FEB 2007
OSCAR ALARCON NUÑEZ
NOTARIO CUARENTA Y SEIS

16803
1615

RECONOCIMIENTO TODO-PROPOSITO DE CALIFORNIA

Estado de California)
} ss.
Condado de San Mateo)

El 6 de diciembre de 2005, ante mi, AIDA MICLAT GUIAM, apareció personalmente Thimoty R. Schwartz

☐ a quien conozco personalmente

☒ quien demostró ante mi en base a evidencia satisfactoria ser la persona cuyo nombre esta suscrito en el instrumento y quien reconoció ante mi que ejecutó el mismo en su capacidad autorizada, y que por su firma en el instrumento la persona o la entidad a nombre de la cual la persona actuó, ejecutó el instrumento.

Hay un sello que dice:
AIDA MICLAT GUIAM
COMISIÓN # 14556612
Notario Público-California
Condado de San Mateo
Mi comisión expira Dic. 9 de 2007

Atestigua mi firma y sello oficial,

AIDA MICLAT GUIAM

OPCIONAL

Aunque la información en esta sección no es requerida por la ley, puede probar ser valiosa para las personas que se basan en el documento y puede prevenir la remoción y readhesión fraudulenta de esta forma a otro documento

Descripción adicional de cualquier documento anexo

Título o tipo de documento: _____

Fecha del Documento: _____ Número de páginas: _____

Firmante(s) diferente(s) a (los) nombrado(s) arriba: _____

Capacidad(es) Reivindicada(s) por el (los) firmante(s)

Nombre del firmante: _____

☐ Individuo
☐ Funcionario Corporativo – Título(s): _____
☐ Socio - ☐ Limitado ☐ General
☐ Abogado de hecho
☐ Fideicomisario
☐ Guardián o conservador
☐ Otro: _____

El firmante representa: _____

Huella índice derecho del firmante
Parte alta del dedo aquí

ESTADO MODERNO GUARDIOLA Y SEIS NOTARIOS DEL CIRCULO DE BANTAME DE BOGOTÁ HAGO CONSTAR QUE ESTA COPIA POTOSTÁTICA CORRESPONDE CON LA ORIGINAL QUE ME FUE PRESENTADA A LA VISTA DE BOGOTÁ D.C. REPÚBLICA DE COLOMBIA.

14 FEB 2007

OSCAR ALARCON NUÑEZ
NOTARIO CUARENTA Y SEIS

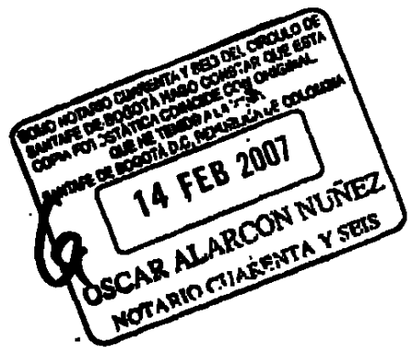
1616

AUTORIZACION

Genentech, Inc. por la presente autoriza a cualquiera de Janet E. Hasak, Mark T. Kresnak, Jeffrey S. Kubinec y Thomity R. Schwartz, mientras permanezcan como empleados de Genentech, Inc., a ejecutar, a su nombre, cualquier solicitud asociada con presentación, procesamiento, consecución, y extensión y mantenimiento de concesión de solicitudes de patente y marca registrada en los Estados Unidos y países extranjeros, contratos relacionados con depósitos biológicos con la Colección Americana de Tipo de Cultivo, y cesiones a publicistas de derechos de autor que cubran los trabajos de autoría de los empleados de Genentech.

27 de Agosto de 2002
Fechado

(Firma)
Sean A. Johnston
Vicepresidente, Propiedad Intelectual
Genentech, Inc.



97. 200
1617

Apostille

(Convention de La Haye du 5 Octobre 1961)

1. Country: United States of America

This public document:

2. has been signed by Harriet Smith Windsor

3. acting in the capacity of Secretary of State of Delaware

4. bears the seal/stamp of Office of Secretary of State

Certified

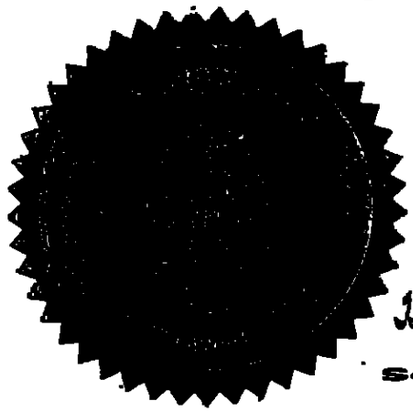
5. at Dover, Delaware

6. the twenty-fourth day of May, A.D. 2005

7. by Secretary of State, Delaware Department of State

8. No. 0254030

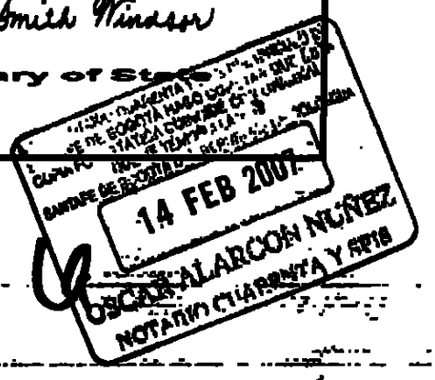
9. Seal/Stamp:



10. Signature:

Harriet Smith Windsor

Secretary of State



1618
1606

Delaware

The First State

I, HARRIET SMITH WINDSOR, SECRETARY OF STATE OF THE STATE OF DELAWARE, DO HEREBY CERTIFY THE ATTACHED IS A TRUE AND CORRECT COPY OF THE RESTATED CERTIFICATE OF "GENENTECH, INC.", FILED IN THIS OFFICE ON THE TWENTY-SECOND DAY OF JULY, A.D. 1999, AT 9 O'CLOCK A.M.

NOTARIA CUARENTA Y SEIS DEL CIRCULO DE SANTAFE DE BOGOTA D.C. COTEJA QUE ESTA COPIA FOTOSTATICA CORRESPONDE CON EL ORIGINAL QUE HE TENIDO A LA VISTA. SANTAFE DE BOGOTA D.C. COLOMBIA
14 FEB 2007
OSCAR ALARCON NUÑEZ
NOTARIA CUARENTA Y SEIS

Harriet Smith Windsor, Secretary of State

2112549

050426506

AUTHENTICATION: 3900384

DATE: 05-24-05

160X
1619

AMENDED AND RESTATED

CERTIFICATE OF INCORPORATION

GENENTECH, INC.

Genentech, Inc. (the "Corporation"), a corporation organized and existing under and by virtue of the General Corporation Law of the State of Delaware, does hereby amend the Certificate of Incorporation of the Corporation, which was originally filed on December 23, 1986 under the name "Genentech Delaware, Inc.", and subsequently amended.

The Certificate of Incorporation of the Corporation is hereby restated and further amended to read in its entirety as follows:

ARTICLE 1

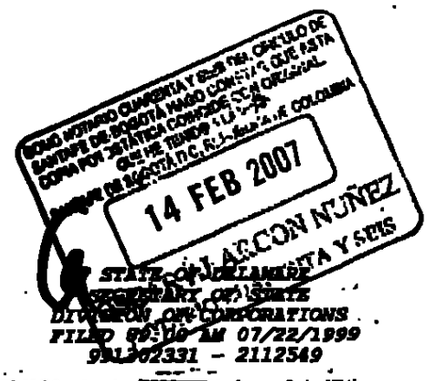
The name of the Corporation is: Genentech, Inc.

ARTICLE 2

The address of the registered office of the Corporation in the State of Delaware is Corporation Service Company, 1013 Centre Road, City of Wilmington, County of New Castle, Delaware 19805. The name of its registered agent at such address is Corporation Service Company.

ARTICLE 3

The purpose of the Corporation is to engage in any lawful act or activity for which corporations may be organized under the General Corporation Law of the State of Delaware as the same exists or may hereafter be amended ("Delaware Law").



1660
1620

ARTICLE 4

SECTION 4.01. *Capital Stock.* (a) The Corporation is authorized to issue two classes of stock to be designated, respectively, preferred stock and common stock. The total number of shares which the Corporation is authorized to issue is four hundred million (400,000,000) shares. One hundred million (100,000,000) shares shall be designated preferred stock, par value $0.02 per share ("Preferred Stock"). Three hundred million (300,000,000) shares shall be designated common stock, par value $0.02 per share ("Common Stock") The Common Stock of the Corporation shall be all of one class.

(b) The number of authorized shares of any class or classes of stock may be increased or decreased (but not below the number of shares thereof then outstanding) by the affirmative vote of a majority in voting power of the outstanding stock of the Corporation entitled to vote thereon irrespective of the provisions of Section 242(b)(2) of Delaware Law.

SECTION 4.02. *Common Stock.* (a) *Issuance and Consideration.* Any unissued or treasury shares of the Common Stock may be issued for such consideration as may be fixed in accordance with applicable law from time to time by the Board of Directors.

(b) *Dividends.* Subject to the rights of holders of the Preferred Stock, the holders of the Common Stock shall be entitled to receive, when and as declared by the Board of Directors, out of the assets of the Corporation which are by law available therefor, dividends payable either in cash, in property, or in shares of stock and the holders of the Preferred Stock shall not be entitled to participate in any such dividends (unless otherwise provided by the Board of Directors in any resolution providing for the issue of a series of Preferred Stock).

SECTION 4.03. *Preferred Stock.*

(a) *Series and Limits of Variations between Series.* Any unissued or treasury shares of the Preferred Stock may be issued from time to time in one or more series for such consideration as may be fixed from time to time by the Board of Directors. Before any shares of Preferred Stock of any particular series shall be issued, a certificate shall be filed with the Secretary of State of Delaware setting forth the designation, rights, privileges, restrictions, and conditions to be attached to the Preferred Stock of such series and such other matters as may be required, and the Board of Directors shall fix and determine, and is hereby expressly empowered to fix and determine, in the manner provided by law, the particulars of

NOTARIA 14 FEB 2007 OSCAR ALARCON NUNEZ

1621

the shares of such series (so far as not inconsistent with the provisions of this Article 4 applicable to all series of Preferred Stock), including, but not limited to, the following:

(i) the distinctive designation of such series and the number of shares which shall constitute such series, which number may be increased (except where otherwise provided by the Board of Directors in creating such series) or decreased (but not below the number of shares thereof then outstanding) from time to time by like action of the Board of Directors;

(ii) the annual rate of dividends payable on shares of such series, the conditions upon which such dividends shall be payable and the date from which dividends shall be cumulative in the event the Board of Directors determines that dividends shall be cumulative;

(iii) whether such series shall have voting rights, in addition to the voting rights provided by law and, if so, the terms of such voting rights;

(iv) whether such series shall have conversion privileges and, if so, the terms and conditions of such conversion, including, but not limited to, provision for adjustment of the conversion rate upon such events and in such manner as the Board of Directors shall determine;

(v) whether or not the shares of such series shall be redeemable and, if so, the terms and conditions of such redemption, including the date or dates upon or after which they shall be redeemable, and the amount per share payable in case of redemption, which amount may vary under different conditions and at different redemption dates;

(vi) whether such series shall have a sinking fund for the redemption or purchase of shares of that series and, if so, the terms and amount of such sinking fund;

(vii) the rights of the shares of such series in the event of voluntary or involuntary liquidation, dissolution or winding up of the Corporation, and the relative rights of priority, if any, of payment of shares of that series; and

(viii) any other relative rights, preferences and limitations of such series.

NOTARIO CUARENTA Y SEIS DEL CIRCULO DE SANTAFE DE BOGOTA HACE CONSTAR QUE ESTA COPIA FOTOSTATICA CORRESPONDE CON ORIGINAL QUE HE TENIDO A LA VISTA.
14 FEB 2007
OSCAR ALARCON NUÑEZ
NOTARIO CUARENTA Y SEIS

1640
1622

ARTICLE 5

SECTION 5.01. *Amendment of Bylaws by Directors.* (a) In furtherance and not in limitation of the powers conferred by statute, the Board of Directors is expressly authorized to make, alter, amend, repeal and rescind the bylaws of the Corporation, subject to Sections 5.01(b) and 5.01(c) hereof.

(b) Sections 2.02, 3.02, 3.03(a), 3.03(b), 3.04, 3.05, 3.14, and 5.06 of the bylaws may only be altered, amended, repealed, or rescinded by the affirmative vote of holders of capital stock entitled to vote thereon representing more than 60% of the shares entitled to be voted thereon.

(c) Sections 3.03(c) and 3.15 of the bylaws may only be altered, amended, repealed, or rescinded by the affirmative vote of holders of capital stock entitled to vote thereon representing more than 90% of the shares entitled to be voted thereon.

SECTION 5.02. *Amendment of Certificate of Incorporation.* Section 5.01(b) of this Certificate of Incorporation may only be altered, amended, repealed, or rescinded by the affirmative vote of holders of capital stock entitled to vote thereon representing more than 60% of the shares entitled to be voted thereon. Section 5.01(c) and Articles 8 and 9 of this Certificate of Incorporation may only be altered, amended, repealed, or rescinded by the affirmative vote of holders of capital stock entitled to vote thereon representing more than 90% of the shares entitled to be voted thereon.

ARTICLE 6

SECTION 6.01. *Election of Directors by Holders of Preferred Stock.* During any period when the holders of any Preferred Stock or any one or more series thereof, voting as a class, shall be entitled to elect a specified number of directors, by reason of dividend arrearages or other provisions giving them the right to do so, then and during such time as such right continues (i) the then otherwise authorized number of directors shall be increased by such specified number of directors, and the holders of such Preferred Stock or such series thereof, voting as a class, shall be entitled to elect the additional directors so provided for, pursuant to the provisions of such Preferred Stock or series; (ii) each such additional director shall serve for such term, and have such voting powers, as shall be stated in the provisions pertaining to such Preferred Stock or series; and (iii) whenever the holders of any such Preferred Stock or series thereof

14 FEB 2007

divested of such rights to elect a specified number of directors, voting as a class, pursuant to the provisions of such Preferred Stock or series, the terms of office of all directors elected by the holders of such Preferred Stock or series, voting as a class pursuant to such provisions or elected to fill any vacancies resulting from the death, resignation or removal of directors so elected by the holders of such Preferred Stock or series, shall forthwith terminate and the authorized number of directors shall be reduced accordingly.

SECTION 6.02. *Ballots.* Elections of directors at an annual or special meeting of stockholders need not be by written ballot unless the bylaws of the Corporation shall provide otherwise.

SECTION 6.03. *Elimination of Certain Personal Liability of Directors.* (a) A director of the Corporation shall not be liable to the Corporation or its stockholders for monetary damages for breach of fiduciary duty as a director to the fullest extent permitted by Delaware Law.

(b) Each person (and the heirs, executors or administrators of such person) who was or is a party or is threatened to be made a party to, or is involved in any threatened, pending or completed action, suit or proceeding, whether civil, criminal, administrative or investigative, by reason of the fact that such person is or was a director of the Corporation or is or was serving at the request of the Corporation as a director of another corporation, partnership, joint venture, trust or other enterprise, shall be indemnified and held harmless by the Corporation to the fullest extent permitted by Delaware Law. The right to indemnification conferred in this Section 6.03 shall also include the right to be paid by the Corporation the expenses incurred in connection with any such proceeding in advance of its final disposition to the fullest extent authorized by Delaware Law. The right to indemnification conferred in this Section 6.03 shall be a contract right.

(c) The Corporation may, by action of its Board of Directors, provide indemnification to such of the officers, employees and agents of the Corporation to such extent and to such effect as the Board of Directors shall determine to be appropriate and authorized by Delaware Law.

(d) The Corporation shall have power to purchase and maintain insurance on behalf of any person who is or was a director, officer, employee or agent of the Corporation, or is or was serving at the request of the Corporation as a director, officer, employee or agent of another corporation, partnership, joint venture, trust or other enterprise against any expense, liability or loss incurred by such person in any such capacity or arising out of his status as such, whether

NOTARIA CUARENTA Y SEIS
14 FEB 2007
OSCAR BOON NUÑEZ
NOTARIO CUARENTA Y SEIS

not the Corporation would have the power to indemnify him against such liability under Delaware Law.

(e) The rights and authority conferred in this Section 6.03 shall not be exclusive of any other right which any person may otherwise have or hereafter acquire.

(f) Neither the alteration, amendment, repeal or recission of this Section 6.03, nor the adoption of any provision of this Certificate of Incorporation or the bylaws of the Corporation, nor, to the fullest extent permitted by Delaware Law, any modification of law, shall eliminate or reduce the effect of this Section 6.03 in respect of any acts or omissions occurring prior to such alteration, amendment, repeal, recission, adoption or modification.

ARTICLE 7

SECTION 7.01. *Section 203 of Delaware Law.* The Corporation hereby elects not to be governed by Section 203 of Delaware Law.

ARTICLE 8

SECTION 8.01. *Relationship With Roche.* In anticipation that the Corporation will cease to be a wholly owned subsidiary of Roche Holdings, Inc., but that Roche (as defined in Section 8.05) will remain a stockholder of the Corporation, and in anticipation that the Corporation (as defined in Section 8.05) and Roche may engage in the same or similar activities or lines of business and have an interest in the same areas of corporate opportunities, and in recognition of (i) the benefits to be derived by the Corporation through its continued contractual, corporate and business relations with Roche (including service of officers and directors of Roche as directors of the Corporation) and (ii) the difficulties attendant to any director, who desires and endeavors fully to satisfy such director's fiduciary duties, in determining the full scope of such duties in any particular situation, the provisions of this Article 8 are set forth to regulate, define and guide the conduct of certain affairs of the Corporation as they may involve Roche and its officers and directors, and the powers, rights, duties and liabilities

14 FEB 2007

1643
1625

of the Corporation and its officers, directors and stockholders in connection therewith.

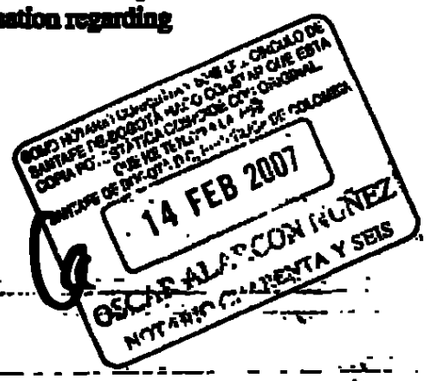
SECTION 8.02. *Similar Business Activities*. Except as Roche Holdings, Inc. may otherwise agree in writing,

(a) Roche shall not have a duty to refrain from engaging directly or indirectly in the same or similar business activities or lines of business as the Corporation, and

(b) neither Roche nor any officer or director thereof shall be liable to the Corporation or its stockholders for breach of any fiduciary duty by reason of any such activities of Roche or of such person's participation therein.

In the event that Roche acquires knowledge of a potential transaction or matter that may be a corporate opportunity for both Roche and the Corporation, Roche shall have no duty to communicate or offer such corporate opportunity to the Corporation and shall not be liable to the Corporation or its stockholders for breach of any fiduciary duty as a stockholder of the Corporation or controlling person of a stockholder by reason of the fact that Roche pursues or acquires such corporate opportunity for itself, directs such corporate opportunity to another person or entity, or does not communicate information regarding, or offer, such corporate opportunity to the Corporation.

SECTION 8.03. *Corporate Opportunities*. In the event that a director, officer or employee of the Corporation who is also a director, officer or employee of Roche acquires knowledge of a potential transaction or matter that may be a corporate opportunity for the Corporation and Roche (whether such potential transaction or matter is proposed by a third-party or is conceived of by such director, officer or employee of the Corporation), such director, officer or employee shall be entitled to offer such corporate opportunity to the Corporation or Roche as such director, officer or employee deems appropriate under the circumstances in his sole discretion, and no such director, officer or employee shall be liable to the Corporation or its stockholders for breach of any fiduciary duty or duty of loyalty or failure to act in (or not opposed to) the best interests of the Corporation or the derivation of any improper personal benefit by reason of the fact that (i) such director, officer or employee offered such corporate opportunity to Roche (rather than the Corporation) or did not communicate information regarding such corporate opportunity to the Corporation or (ii) Roche pursues or acquires such corporate opportunity for itself or directs such corporate opportunity to another person or does not communicate information regarding such corporate opportunity to the Corporation.



1626

SECTION 8.04. *Deemed Consent of Stockholders.* Any person or entity purchasing or otherwise acquiring any interest in any shares of capital stock of the Corporation shall be deemed to have notice of and to have consented to the provisions of this Article 8.

SECTION 8.05. *Definitions.* For purposes of this Article 8 and Article 9 only, (i) the term "Corporation" shall mean the Corporation and all corporations, partnerships, joint ventures, associations and other entities in which the Corporation beneficially owns (directly or indirectly) fifty percent or more of the outstanding voting stock, voting power or similar voting interests, and (ii) the term "Roche" shall mean Roche Holdings, Inc. and all corporations, partnerships, joint ventures, associations and other entities (other than the Corporation, defined in accordance with clause (i) of this Section 8.05) that are "affiliates" (as defined in Rule 12b-2 under the Securities Exchange Act of 1934, as amended) of Roche Holdings, Inc.

SECTION 8.06. *Termination; Binding Effect.* Notwithstanding anything in this Certificate of Incorporation to the contrary, the foregoing provisions of this Article 8 shall expire on the date that Roche ceases to own beneficially Common Stock representing at least 5% of the number of outstanding shares of Common Stock of the Corporation and no person who is a director or officer of the Corporation is also a director or officer of Roche. Neither such expiration, nor the alteration, amendment, change or repeal of any provision of this Article 8 nor the adoption of any provision of this Amended and Restated Certificate of Incorporation inconsistent with any provision of this Article 8 shall eliminate or reduce the effect of this Article 8 in respect of any matter occurring, or any cause of action, suit or claim that, but for this Article 8, would accrue or arise, prior to such expiration, alteration, amendment, repeal or adoption.

SECTION 8.07. *Article 9.* The provisions of this Article 8 are in addition to the provisions of Article 9.

ARTICLE 9

SECTION 9.01. *Contracts Not Void.* No contract, agreement, arrangement or transaction (or any amendment, modification or termination thereof) between the Corporation and Roche shall be void or voidable solely for the reason that Roche is a party thereto, or solely because any directors or officers of the Corporation who are affiliated with Roche are present at or participate in the

NOTARIA CINCUENTA Y SEIS D.C. CIRCULO DE BOGOTA
COPIA FIEL DEL ORIGINAL
14 FEB 2007
OSCAR ALARCON NUÑEZ
NOTARIO CINCUENTA Y SEIS

meeting of the Board of Directors or committee thereof which authorizes the contract, agreement, arrangement, transaction, amendment, modification or termination or solely because his or their votes are counted for such purpose, but any such contract, agreement, arrangement or transaction (or any amendment, modification or termination thereof) shall be governed by the provisions of this Amended and Restated Certificate of Incorporation, the Corporation's Bylaws, Delaware Law and other applicable law. For purposes of this Article 9, the terms the "Corporation" and "Roche" have the meanings set forth in Section 8.05.

SECTION 9.02. *Quorum.* Directors of the Corporation who are also directors or officers of Roche may be counted in determining the presence of a quorum at a meeting of the Board of Directors or of a committee that authorizes or approves any such contract, agreement, arrangement or transaction (or amendment, modification or termination thereof). Outstanding shares of Common Stock owned by Roche may be counted in determining the presence of a quorum at a meeting of stockholders that authorizes or approves any such contract, agreement, arrangement or transaction (or amendment, modification or termination thereof).

SECTION 9.03. *No Liability For Good Faith Actions.* Neither Roche nor any officer or director thereof shall be liable to the Corporation or its stockholders for breach of any fiduciary duty or duty of loyalty or failure to act in (or not opposed to) the best interests of the Corporation or the derivation of any improper personal benefit by reason of the fact that Roche or an officer of director thereof in good faith takes any action or exercises any rights or gives or withholds any consent in connection with any agreement or contract between Roche and the Corporation. No vote cast or other action taken by any person who is an officer, director or other representative of Roche, which vote is cast or action is taken by such person in his capacity as a director of this Corporation, shall constitute an action of or the exercise of a right by or a consent of Roche for the purpose of any such agreement or contract.

SECTION 9.04. *Deemed Consent by Stockholders.* Any person or entity purchasing or otherwise acquiring any interest in any shares of capital stock of the Corporation shall be deemed to have notice of and to have consented to the provisions of this Article 9.

SECTION 9.05. *Contracts Covered.* For purposes of this Article 9, any contract, agreement, arrangement or transaction with the Corporation or any corporation, partnership, joint venture, association or other entity in which the Corporation beneficially owns (directly or indirectly) fifty percent or more of the outstanding voting stock, voting power or similar voting interests shall be deemed to be a contract, agreement, arrangement or transaction with the Corporation.

14 FEB 2007
OSCAR ARANGO NUÑEZ

SECTION 9.06. *Binding Effect.* Neither the alteration, amendment, change or repeal of any provision of this Article 9 nor the adoption of any provision inconsistent with any provision of this Article 9 shall eliminate or reduce the effect of this Article 9 in respect of any matter occurring, or any cause of action, suit or claim that, but for this Article 9, would accrue or arise, prior to such alteration, amendment, change, repeal or adoption.

SECTION 9.07. *Article 8.* The provisions of this Article 9 are in addition to the provisions of Article 8.

14 FEB 2007
OSCAR ALARCON NUÑEZ
NOTARIO CUARENTA Y SEIS

IN WITNESS WHEREOF, this Amended and Restated Certificate of Incorporation, having been duly adopted in accordance with Sections 242 and 245 of the General Corporation Law of the State of Delaware and by the written consent of the sole stockholder of the Corporation in accordance with Section 228 of the General Corporation Law of the State of Delaware, has been executed this 22nd day of July, 1999.

GENENTECH, INC.

By: /s/ Stephen G. Juelsgaard
Name: Stephen G. Juelsgaard
Title: Senior Vice President,
General Counsel and Secretary

14 FEB 2007

TRADUCCION AL ESPAÑOL del Certificado de Incorporación de Genentech, Inc.

APOSTILLA

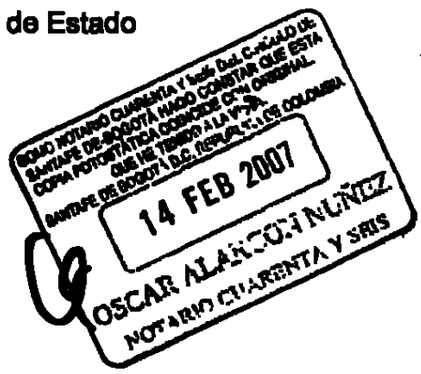
(Convención de La Haya del 5 de Octubre de 1961)

1. País: Estados Unidos de América
   Este documento público
2. ha sido firmado por Harriet Smith Windsor
3. actuando en capacidad de Secretaria de Estado de Delaware
4. lleva el sello/timbre de la Oficina del Secretario de Estado

Certificado

5. en Dover, Delaware
6. el veinticuatro de mayo, A. D. 2005
7. por el Secretario de Estado, Departamento de Estado de Delaware
8. No. 0254030
9. Sello/timbre
10. Firma

Harriet Smith Windsor (Firma)
Secretaria de Estado



preferenciales, por un valor de $0.02 por acción ("**Acción Preferencial**"). Trecientos millones (300,000,000) de acciones serán designadas acciones comunes, por un valor de $0.02 por acción ("**Acción Común**"). La acción común de la Corporación será toda de una misma clase.

(b) El número de acciones autorizadas de cualquier clase o clases de acciones puede incrementarse o disminuirse (pero no por debajo del número del número de acciones de las mismas que en su momento están vigentes) por el voto afirmativo de una mayoría con poder de votación de las acciones vigentes de la Corporación con derecho a votar por la misma independiente de las consideraciones de la Sección 242(b)(2) de la Ley de Delaware.

SECCION 4.02. *Acción Común.* (a) *Emisión y Consideración.* Cualquier acción sin emitir o del tesoro del las Acciones Comunes puede emitirse para tal consideración como pueda fijarse de acuerdo con la ley aplicable cada cierto tiempo por la Junta Directiva.

(b) *Dividendos.* Sujeto a los derechos de los accionistas de las Acciones Preferenciales, los accionistas de Acciones Comunes deberán tener el derecho a recibir, cuando y como se declare por la Junta Directiva, de los activos de la Corporación los cuales están disponibles para ello por ley, dividendos pagables ya sea en dinero, propiedad, o acciones y los accionistas de la Acción Preferencial no tendrán derecho a participar en tales dividendos (a menos que la Junta Directiva determine lo contrario en cualquier resolución que condiciona la emisión de una serie de Acciones Preferenciales).

SECCIÓN 4.03. *Acción Preferencial*

(a) *Series y Límites de Variaciones entre Series.* Cualquier acción no emitida o del tesoro de las Acciones Preferenciales puede emitirse cada cierto tiempo en una o más series para tal consideración como pueda ser fijada cada cierto tiempo por la Junta Directiva. Antes de que se emita cualquier acción de las Acciones Preferenciales de cualquier serie particular, se deberá presentar un certificado ante el Secretario de Estado de Delaware estableciendo las designaciones, derechos, privilegios, restricciones, y condiciones que irán anexas a las Acciones Preferenciales de tal serie y tales otras cuestiones como puedan requerirse, y la Junta Directiva deberá fijar y determinar, y por la presente está expresamente autorizada para fijar y determinar, en la forma provista por la ley,

14 FEB 2007
OSCAR ALARCON NUÑEZ
NOTARIO CUARENTA Y SEIS

1620
1632

CORREGIDO Y VUELTO A REDACTAR
CERTIFICADO DE INCORPORACIÓN
GENENTECH, INC.

Genentech, Inc (la **"Corporación"**), una corporación organizada y existente bajo y por virtud de la Ley General de Corporación del Estado de Delaware, por la presente corrige el Certificado de Incorporación, el fue presentado originalmente el 23 de diciembre de 1986, bajo el nombre "Genentech Delaware, Inc.", y corregido subsiguientemente.

Por la presente se vuelve a redactar el certificado de incorporación de la Corporación y se corrige aun más para que diga en su totalidad lo que sigue:

ARTICULO 1

El nombre de la Corporación es: Genentech, Inc.

ARTICULO 2

La dirección de la oficina registrada de la Corporación en el Estado de Delaware es Corporation Service Company, 1013 Centre Road, Ciudad de Wilmington, Condado de New castle, Delaware 19805. El nombre de su agente registrado en tal dirección es Corporation Service Company.

ARTICULO 3

El propósito de a Corporación es el de participar en cualquier acto o actividad legal para las cuales las corporaciones pueden ser organizadas bajo la Ley General de Corporación del Estado de Delaware como existe o puede se modificada la misma en adelante (**"Delaware Law"**).

ARTICULO 4

SECCION 4.01. *Acciones de Capital.* (a) la Corporación está autorizada a emitir dos clases de acciones que deben designarse, respectivamente, acciones preferenciales y acciones comunes. El número total de acciones que la Corporación esta autorizada a emitir es de cuatrocientos millones (400,000,000) de acciones. Cien millones (100,000,000) de acciones serán designadas acciones



1621
1633

particulares de las acciones de tal serie (siempre y cuando no sean inconsistentes con las condiciones de este Artículo 4 aplicables a todas las series de Acciones Preferenciales), incluyendo, pero sin limitarse a, las siguientes:

(i) la designación distintiva de tal serie y el número de las acciones que han de constituir tal serie, numero que puede incrementarse (excepto cuando lo determine de otro modo la Junta Directiva al crear tal serie) o reducida (pero no por debajo del número de acciones de las mismas que están vigentes) cada cierto tiempo por una acción semejante de la Junta Directiva;

(ii) la tasa anual de dividendos a pagar sobre las acciones de tal serie, las condiciones sobre las cuales se pagarán tales dividendos y le fecha a partir de la cual los dividendos serán acumulativos en el evento que la Junta Directiva determine que los dividendos serán acumulativos;

(iii) si la serie tendrá o no derechos de voto, además de los derechos de voto provistos por la ley, y de ser así, los términos de tales derechos de voto;

(iv) si tal serie tendrá o no privilegios de conversión y, de ser así, los términos y condiciones de tal conversión, incluyendo, pero sin estar limitado a, la condición para ajuste de la tasa de conversión en tales eventos y de tal forma como la Junta Directiva lo determine;

(v) si las acciones de tal serie serán o no redimibles y, de ser así, los términos y condiciones de tal redención, incluyendo la fecha o fechas en las cuales o después de las cuales pueden ser redimibles, y la cantidad a pagar por acción en caso de redención, cantidad que variará bajo diferentes condiciones y én diferentes fechas de redención;

(vi) si tal serie tendrá o no un fondo de inversión para la redención o compra de acciones de esa serie y, de ser así, los términos y la cantidad de tal fondo de inversión;

(vii) los derechos de las acciones de tal serien caso de liquidación, disolución o cierre voluntaria o involuntaria de



los derechos relativos de prioridad, de haber, de pago de acciones de esa serie; y

(viii) cualquier otro derecho relativo, preferencia y limitación de tal serie.

ARTICULO 5

SECCION 5.01. *Enmiendas a los estatutos por los Directores.* (a) En adición y no limitando los poderes conferidos por estatutos, la Junta Directiva esta expresamente autorizada para hacer, alterar, enmendar, rechazar y rescindir los estatutos de la Corporación, sujetos a las secciones 5.01(b) y 5.01(c) a continuación.

(b) Las secciones 2.02, 3.02, 3.03(a), 3.04, 3.05, 3.14 y 5.06 de los estatutos solo pueden alterarse, enmendarse, repelerse, o rescindirse por el voto afirmativo de los accionistas de capital con derecho a votar las mismas que representen más del 60% de las acciones con derecho a votar las mismas.

(c) Las secciones 3.03(c) y 3.15 de los estatutos solo pueden ser alteradas, enmendadas, repelidas, o rescindidas por le voto afirmativo de los accionistas de capital con derecho a votar las mismas que representen más del 90% de las acciones con derecho a votar las mismas.

o

SECCION 5.02. *Enmienda del Certificado de Incorporación.* La sección 5.01 (b) de este Certificado de Incorporación solo puede ser alterada, enmendada, repelida, o rescindida por el voto afirmativo de los accionistas de capital con derecho a voto sobre la misma que representen a más del 60% de las acciones con derecho a votar la misma. La sección 5.01(c) y los Artículos 8 y 9 de este Certificado de Incorporación solo pueden alterarse, enmendarse, repelerse, o rescindirse por el voto afirmativo de los accionistas de capital con derecho a votar la misma que representen más del 90% de las acciones con derecho a votar la misma.

ARTICULO 6

SECCION 6.01: *Elección del Directorio por los Accionistas; Voto Preferencial.* Durante cualquier periodo cuando los acci

14 FEB 2007
OSCAR ALARCON NUÑEZ
NOTARIO CUARENTA Y SEIS

acción preferencial cualquiera de una o más series de las mismas, que voten como una clase, tengan derecho a elegir un número específico de directores, por razón a atrasos en los dividendos u otras condiciones que les dan el derecho a hacerlo, entonces y durante el tiempo que continúe el derecho (i) el número de directores autorizado de otro modo será incrementado por tal número de directores especificado, y los accionistas de tales Acciones Preferenciales o tales series de las mismas, que votan como una clase, tendrán derecho a elegir los directores adicionales para lo que se provee, de acuerdo con las condiciones de tales Acciones Preferenciales o series; (ii) cada uno de tales directores adicionales servirá durante tal termino, y tendrá tales poderes de voto, como se han de otorgar en las provisiones pertinentes a tales Acciones Preferenciales o series; y (iii) siempre que los accionistas de cualquiera de tales Acciones Preferenciales o series de las mismas están desprovistos de tales derechos a elegir un número específico de directores, votando como clase, de acuerdo con las condiciones de tales Acciones preferenciales o series, los términos de oficio de todos los directores elegidos por los accionistas de tales Acciones preferenciales o series, que votan como clase, de acuerdo con las condiciones o elegidos para llenar cualquier vacante resultante de la muerte, renuncia o remoción de directores así elegidos por los accionistas de tales Acciones Preferenciales o series, terminarán en adelante y el número autorizado de directores será reducido en concordancia.

SECCION 6.02. *Votaciones*. Las elecciones de directores en una reunión anual o especial de accionistas necesaria no serán por votos escritos a menos que los estatutos de la Corporación digan lo contrario.

SECCION 6.03. *Eliminación de cierta responsabilidad personal de los directores*. (a) Un director de la Corporación no será responsable ante la Corporación o sus accionistas por daños monetarios por violación del deber fiduciario como director hasta donde lo permita la Ley de Delaware.

(b) Cada persona (y sus herederos, ejecutores o administradores de tal persona) quien fue o es una parte o esta amenazado de ser parte de, o esta involucrado en cualquier acción, demanda o proceso amenazado, pendiente, o completado, ya sea civil, criminal, administrativo o investigativo, por razón del hecho de que esa persona es o fue un director de la Corporación o es o estaba sirviendo por solicitud de la Corporación como un director de otra corporación, sociedad, sociedad de riesgo compartido, fideicomiso u otra [illegible] será indemnizado y completamente protegido por la Corporación



permita la Ley de Delaware. El derecho a la indemnización concedido en esta sección 6.03 también debe incluir el derecho a que la Corporación le pague los gastos incurridos en conexión cuna cualquiera de tales procesos por adelantado de su decisión final hasta donde los autoriza la Ley de Delaware. El derecho a indemnización conferido en esta sección 6.03 será un derecho contractual.

(c) La Corporación puede, por acción de su Junta Directiva, proporcionar indemnización a tales funcionarios, empleados y agentes de la Corporación hasta tal punto y a tal efecto como lo determine apropiado la Junta Directiva y lo autorice la Ley de Delaware.

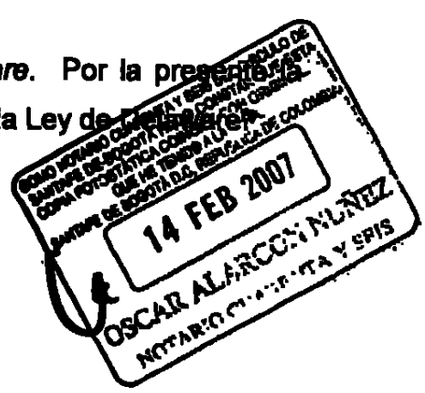
(d) La Corporación tendrá el poder de comprar y mantener un seguro a nombre de cualquier persona quien es o fue director, funcionario, empleado o agente de la corporación, o esta o estuvo en servicio por pedido de la Corporación como director, funcionario, empleado o agente de otra corporación, sociedad, sociedad de riesgo compartido, fideicomiso u otra empresa contra cualquier gasto, responsabilidad o pérdida incurrida por tal persona en cualquier capacidad o que surge de su estado como tal, ya sea si la Corporación tiene o no el poder de indemnizarlo contra tal responsabilidad bajo la ley de Delaware.

(e) Los derechos y autoridad conferidos en esta Sección 6.03 no serán exclusivos de ningún otro derecho el cual cualquier otra persona pueda tener o adquirir de aquí en adelante.

(f) Ni la alteración, enmienda, rechazo o rescisión de esta Sección 6.03, ni la adopción de ninguna condición de este Certificado de Incorporación o de los Estatutos de la Corporación, ni, hasta donde lo permite la Ley de Delaware, ninguna modificación de la ley, debe eliminar o reducir el efecto de esta Sección 6.03 respecto a ninguno de los actos u omisiones que ocurran antes de tal alteración, enmienda, rechazo, rescisión, adopción o modificación.

ARTICULO 7

SECCION 7.01, *Sección 203 de la Ley de Delaware.* Por la presente la Corporación elige no ser gobernada por la Sección 203 de la Ley de [illegible]



1605
1637

ARTICULO 8

SECCION 8.01. *Relación con Roche.* Anticipándose a que la Corporación cese de ser un subsidiaria completamente de propiedad de Roche Holdings, Inc., pero que Roche (como se define en la sección 8.05) permanezca siendo un accionista de la Corporación, y anticipándose a que la Corporación (como se define en la Sección 8.05) y Roche puede entrar en actividades o líneas de negocios semejantes o similares y tener un interés en las mismas áreas de oportunidades corporativas, y en reconocimiento de (i) los beneficios a ser derivados por la Corporación a través de sus relaciones contractuales, corporativas y de negocios con Roche (incluyendo el servicio de funcionarios y directores de Roche como directores de la Corporación) y (ii) las dificultades concomitantes a cualquier director, quien desea y se propone satisfacer completamente tales deberes fiduciarios del director, al determinar el alcance completo de tales deberes en cualquier situación particular, las condiciones de este artículos 8 se establecen para regular, definir y guiar la conducta de ciertas cuestiones de la Corporación ya que pueden involucrar a Roche y sus funcionaros y directores, y los poderes, derechos, deberes y responsabilidades de la Corporación y sus funcionarios, directores y accionistas en conexión con éste.

SECCION 8.02. *Actividades en Negocios Similares.* Excepto que Roche Holdings, Inc. pueda de otra forma acordarlo por escrito,

(a) Roche no tendrá el deber de abstenerse de entrar directamente o indirectamente en actividades de negocios o líneas de negocios semejantes o similares a las de la Corporación, y

(b) ni Roche ni ninguno de sus funcionarios o director de la misma será responsable a la Corporación o sus accionistas por violar cualquier deber fiduciario por razón de alguna de tales actividades de Roche o de la participación de tal persona en las mismas.

En el evento que Roche adquiera conocimientos de una transacción o materia potencial que pueda ser una oportunidad corporativa tanto para Roche como para la Corporación, Roche no tendrá el deber de comunicar u ofrecer tal oportunidad corporativa a la Corporación y no será responsable ante la Corporación o sus accionistas por violar ningún deber fiduciario por el accionista de la Corporación o la persona que controla a un accionista por razón del hecho



que Roche prosigue o adquiere tal oportunidad corporativa para si misma, dirige tal oportunidad corporativa a otra persona o entidad, o no comunica información respecto, u ofrece, tal oportunidad corporativa a la Corporación.

SECCION 8.03. *Oportunidades Corporativas.* En el evento que un director, funcionario o empleado de la Corporación quien también es director, funcionario o empleado de Roche adquiera conocimiento de una transacción o materia potencial que pueda ser una oportunidad corporativa para la Corporación y Roche (ya se si esta transacción o materia potencial es propuesta por un tercero o es concebida por tal director, funcionario o empleado de la Corporación o no), tal director, funcionario o empleado debe tener el derecho a ofrecer tal oportunidad corporativa a la Corporación o Roche como tal director, funcionario o empleado los considere apropiado bajo las circunstancias de su discreción, y tal director, funcionario o empleado no será responsable a la Corporación o sus accionistas por violar ningún deber fiduciario o deber de lealtad o faltar en actuar en (o no oponerse a) los mejores intereses de la Corporación o la derivación de ningún beneficio personal inapropiado por razón del hecho que (i) tal director, funcionario o empleado ofreció tal oportunidad corporativa a Roche (en lugar de a la Corporación) o no comunicó información respecto a tal oportunidad corporativa a la Corporación o (ii) Roche busca o adquiere tal oportunidad corporativa para si misma o dirige tal oportunidad corporativa a otra persona o no comunica información respecto a tal oportunidad corporativa a la Corporación.

SECCION 8.04. *Consentimiento estimado de los accionistas.* Se considerará que cualquier persona o entidad que compre o de otro modo adquiera algún interés en cualquiera de las acciones de capital de la Corporación ha tenido noticia o ha estado de acuerdo con las condiciones de este Artículo 8.

SECCION 8.05, *Definiciones.* Para propósitos de este Artículo 8 y el Artículo 9 solamente, (i) el término ***"Corporación"*** debe significar la Corporación y todas las corporaciones, sociedades, sociedades de riesgo compartido y otras entidades en las cuales la Corporación posee con beneficio (directa o indirectamente) cincuenta por ciento o más de las acciones con voto, poder de voto o intereses de voto similares, y (ii) el término ***"Roche"*** debe significar Roche Holdings, Inc y todas las corporaciones, sociedades, sociedades de riesgo compartido, asociaciones y otras entidades (diferentes a la Corporación definida de acuerdo con la cláusula (i) de esta sección 8.05) que son [illegible] como se

14 FEB 2007
OSCAR ALARCON NUÑEZ
NOTARIO

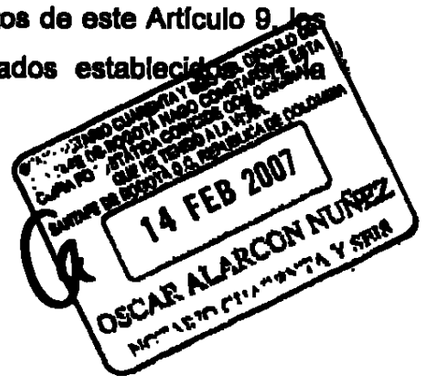
define en la Regla 12b-2 bajo el Acta de Intercambio de Valores de 1934, como se enmendó) de Roche Holdings, Inc.

SECCION 8.06, *Finalización: Efecto Obligante.* Son consideración con nada en este Certificado de Incorporación que lo contradiga, las condiciones anteriores de este Artículo 8 deberán expirar en la fecha que Roche cese de poseer con beneficios Acciones Comunes que representen al menos 5% del número de acciones vigentes de las Acciones Comunes de la Corporación y nadie que se director o funcionario de la Corporación sea también director o funcionario de Roche. Ni tal vencimiento, ni la alteración, enmienda, cambio o rechazo de alguna de las condiciones de este Artículo 8 así como tampoco ninguna condición de este Certificado de Incorporación Enmendado y Vuelto a Redactar inconsistente con alguna de las condiciones de este Artículo 8 eliminará o reducirá el efecto de este Artículo 8 respecto a cualquier materia, o cualquier causa de acción, demanda o reclamo que, excepto por este Artículo 8, pueda acaecer o surgir, antes de tal vencimiento, alteración, enmienda o adopción.

SECCION 8.07. *Artículo 9.* Las condiciones de este Artículo son adicionales a las condiciones del Artículo 9.

ARTICULO 9

SECCION 9.01. *Contratos no nulos.* Ningún contrato, acuerdo, arreglo o transacción (o cualquier enmienda, modificación o finalización de los mismos) entre la Corporación y Roche será nulo u anulable solamente por razón de que Roche es parte del mismo, o solamente porque cualquiera de los directores o funcionarios de la Corporación quienes están afiliados a Roche están presentes en o participan en la reunión de la Junta Directiva o comité de la misma que autoriza el contrato, acuerdo, arreglo, transacción, enmienda, modificación o finalización o solamente porque su o sus votos se cuentan para tal propósito, pero cualquiera de tal contrato, acuerdo, arreglo o transacción (o cualquier enmienda, modificación o finalización del mismo) estará gobernado por las condiciones de este Cerificado de Incorporación Enmendado y vuelto a Redactar, los Estatutos de la Corporación, la Ley de Delaware y otras leyes aplicables. Para propósitos de este Artículo 9, los términos "***Corporación***" y "***Roche***" tienen los significados establecidos en la Sección 8.05.



SECCION 9.02. *Quórum.* Los Directores de la Corporación que son también directores o funcionarios de Roche pueden contarse al determinar la presencia de un quórum en una reunión de la Junta Directiva o de un comité que autoriza o aprueba cualquiera de tal contrato, acuerdo, arreglo o transacción (o cualquier enmienda, modificación o finalización del mismo). Las acciones vigentes de las Acciones Comunes de propiedad de Roche pueden contarse al determinar la presencia de un quórum en una reunión de accionistas que autorice o apruebe tal contrato, acuerdo, arreglo o transacción (o cualquier enmienda, modificación o finalización del mismo).

SECCION 9.03. *No responsabilidad por acciones de buena fe.* Ni Roche ni ningún funcionario o director de la misma será responsable ante la Corporación o sus accionistas por violar cualquiera de los deberes fiduciarios o deberes de lealtad o por fallar en actuar en (o no oponerse a) los mejores intereses de la Corporación o por derivación de cualquier beneficio personal por razón del hecho que Roche o un funcionario o director de la misma de buena fe tome cualquier acción o ejerza cualquier derecho o de o retenga cualquier consentimiento en conexión con cualquier acuerdo o contrato entre Roche y la Corporación. Ningún voto dado u otra acción tomada por cualquier persona que es un funcionario, director o representante de Roche, donde el voto o acción es tomado por tal persona en su capacidad como director de esta Corporación, constituirá una acción de o el ejercicio de un derecho o un consentimiento de Roche para los propósitos de cualquiera de tal acuerdo o contrato.

SECCION 9.04. *Consentimiento estimado de los accionistas.* Se considerará que cualquier persona o entidad que compre o de otro modo adquiera algún interés en cualquiera de las acciones de capital de la Corporación ha tenido noticia o ha estado de acuerdo con las condiciones de este Artículo 9.

SECCION 9.05. *Contratos cubiertos.* Para propósitos de este Artículo 9, cualquier contrato, acuerdo, arreglo o transacción con la Corporación o cualquier corporación, sociedad, sociedad de riesgo conjunto, asociación u otra entidad en la cual la Corporación posee con beneficio (directa o indirectamente) el cincuenta por ciento o más de las acciones con voto, poderes de voto o intereses de voto similares vigentes se considerará un contrato, acuerdo o transacción con la Corporación.

14 FEB 2007

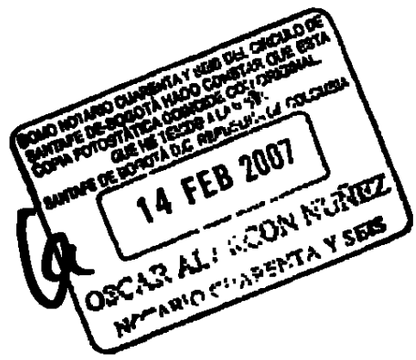
1641

SECCION 9.06. *Efecto obligante.* Ni la alteración, enmienda, cambio o rechazo de alguna de las condiciones de este Artículo 9 ni la adopción de alguna condición inconsistente con cualquier condición de este Artículo 9 eliminará o reducirá el efecto de este Artículo 9 respecto a cualquier materia que ocurra, o cualquier causa de acción, demanda o reclamo que, a no ser por este Artículo 9, pueda acaecer o surgir, antes de tal alteración, enmienda, cambio, rechazo o adopción.

SECCION 9.07. *Artículo 8.* Las condiciones de este Artículo 9 son adicionales a las condiciones del Artículo 8.

EN TESTIMONIO DE LO CUAL, este Certificado de Incorporación Enmendado y vuelto a Redactar, el cual ha sido adoptado de acuerdo con las secciones 242 y 245 de la Ley General de Corporación del Estado de Delaware y por consentimiento escrito de el único accionista de la Corporación de acuerdo con la sección 228 de Ley General de Corporación del Estado de Delaware, ha sido ejecutado este 22 de Julio de 1999.

GENENTECH, INC.
Por: /s/ Stephen G. Juelsgaard
Nombre: Stephen G. Juelsgaard
Título: Vicepresidente
Consejo General y Secretaría



Apostille

(Convention de La Haye du 5 Octobre 1961)

1. Country: United States of America

This public document:

2. has been signed by Harriet Smith Windsor

3. acting in the capacity of Secretary of State of Delaware

4. bears the seal/stamp of Office of Secretary of State

Certified

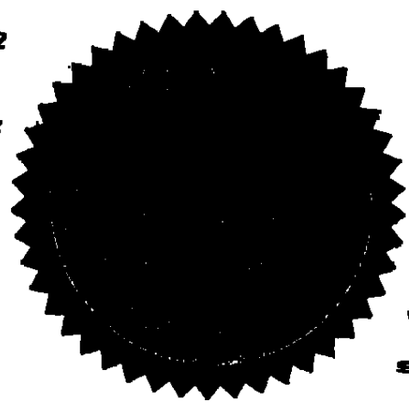
5. at Dover, Delaware

6. the twenty-fourth day of May, A.D. 2005

7. by Secretary of State, Delaware Department of State

8. No. 0254032

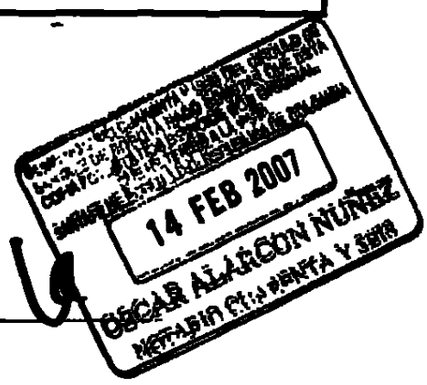
9. Seal/Stamp:



10. Signature:

Harriet Smith Windsor

Secretary of State



Delaware

1643
1631

The First State

I, HARRIET SMITH WINDSOR, SECRETARY OF STATE OF THE STATE OF DELAWARE, DO HEREBY CERTIFY THE ATTACHED IS A TRUE AND CORRECT COPY OF THE CERTIFICATE OF AMENDMENT OF "GENENTECH, INC.", FILED IN THIS OFFICE ON THE TENTH DAY OF MAY, A.D. 2001, AT 2:05 O'CLOCK P.M.

2112549

050426506

Harriet Smith Windsor

Harriet Smith Windsor, Secretary of State

AUTHENTICATION: 3900386

DATE: 05-24-05

NOTARIA CUARENTA Y SEIS DEL CIRCULO DE SANTAFE DE BOGOTA D.C. REPUBLICA DE COLOMBIA
COMPARE: ATENTAMENTE LA FIRMA CONSIGNADA EN ESTA COPIA CON LA QUE APARECE EN EL DOCUMENTO ORIGINAL QUE HA TENIDO A LA VISTA.
SANTAFE DE BOGOTA D.C.
14 FEB 2007
OSCAR ALARCON NUÑEZ
NOTARIO CUARENTA Y SEIS

STATE OF DELAWARE
SECRETARY OF STATE
DIVISION OF CORPORATIONS
FILED 02:05 PM 05/10/2001
010228224 – 2112549

BY Sandy Miller

1644
632

CERTIFICATE OF AMENDMENT OF AMENDED AND RESTATED CERTIFICATE OF INCORPORATION

Genentech, Inc. a corporation organized and existing under and by virtue of the General Corporation Law of the State of Delaware.

DOES HEREBY CERTIFY:

FIRST: That at a meeting of the Board of Directors of Genentech, Inc. resolutions were duly adopted setting forth a proposed amendment of the Amended and Restated Certificate of Incorporation of said corporation, declaring said amendment to be advisable and calling a meeting of the stockholders of said corporation for consideration thereof. The resolution setting forth the proposed amendment is as follows:

RESOLVED, that the Amended and Restated Certificate of Incorporation of this corporation be amended by changing "Section 4.01(a)" of the Article thereof numbered "4" so that, as amended, said Section of said Article shall be and read as follows:

> "Section 4.01. *Capital Stock.* (a) The Corporation is authorized to issue two classes of stock to be designated, respectively, preferred stock and common stock. The total number of shares which the Corporation is authorized to issue is one billion three hundred million (1,300,000,000) shares. One hundred million (100,000,000) shares shall be designated preferred stock, par value $0.02 per share ("Preferred Stock"). One billion two hundred million (1,200,000,000) shares shall be designated common stock, par value $0.02 per share ("Common Stock"). The Common Stock of the Corporation shall be all of one class."

SECOND: That thereafter, pursuant to resolution of its Board of Directors, the regular meeting of the stockholders of said corporation was duly called and held upon notice in accordance with Section 222 of the General Corporation Law of the State of Delaware at which meeting the necessary number of shares as required by statute were voted in favor of the amendment.

THIRD: That said amendment was duly adopted in accordance with the provisions of Section 242 of the General Corporation Law of the State of Delaware.

FOURTH: That the capital of said corporation shall not be reduced under or by reason of said amendment.

IN WITNESS WHEREOF, said Genentech, Inc. has caused this Certificate of Amendment to be signed by Stephen G. Juelsgaard, an Authorized Officer, this 10th day of May, 2001.

GENENTECH, INC.

By: _____
Name: Stephen G. Juelsgaard
Title: Senior Vice President, General Counsel and Secretary

88767

NOTARIA QUINCUAGESIMA Y SIETE DEL CIRCULO DE BOGOTA D.C. ESTADO DE PRESENTACION PERSONAL
COPIA FOTOSTATICA QUE COINCIDE CON EL ORIGINAL QUE TUVE A LA VISTA
14 FEB 2007
OSCAR J. LARCON NUÑEZ
NOTARIO QUINCUAGESIMA Y SIETE

1645

TRADUCCION AL ESPAÑOL del Certificado de Enmienda de Genentech, Inc.

APOSTILLA

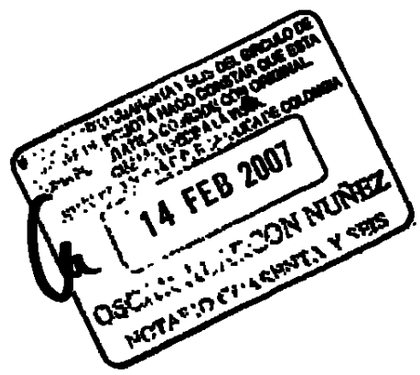
(Convención de La Haya del 5 de Octubre de 1961)

1. País: Estados Unidos de América
   Este documento público
2. ha sido firmado por Harriet Smith Windsor
3. actuando en capacidad de Secretaria de Estado de Delaware
4. lleva el sello/timbre de la Oficina del Secretario de Estado

Certificado

5. en Dover, Delaware
6. el veinticuatro de mayo, A. D. 2005
7. por el Secretario de Estado, Departamento de Estado de Delaware
8. No. 0254032
9. Sello/timbre
10. Firma

Harriet Smith Windsor (Firma)
Secretaria de Estado

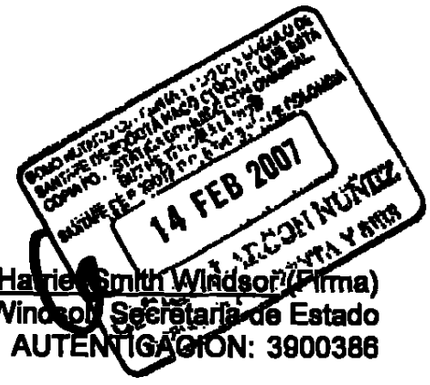


1646
1634

DELAWARE

EL PRIMER ESTADO

YO, HARRIET SMITH WINDSOR, SECRETARIA DE ESTADO DEL ESTADO DE DELAWARE, CERTIFICO POR LA PRESENTE QUE EL ANEXO ES UNA COPIA AUTÉNTICA Y CORRECTA DEL CERTIFICADO DE ENMIENDA DE "GENENTECH, INC", PRESENTADO EN ESTA OFICINA EL DECIMO DIA DE MAYO, A. D. 2001, A LAS 2:05 EN PUNTO DE LA TARDE.



Harriet Smith Windsor (Firma)
Harriet Smith Windsor, Secretaria de Estado
AUTENTIGACION: 3900386

FECHA: 05 – 24 – 05

1647
635

CERTIFICADO DE TERCERA ENMIENDA DEL
CERTIFICADO DE INCORPORACIÓN
CORREGIDO Y VUELTO A REDACTAR

Genentech, Inc. una corporación organizada y existente bajo y por virtud de la Ley General de Corporación del Estado de Delaware.

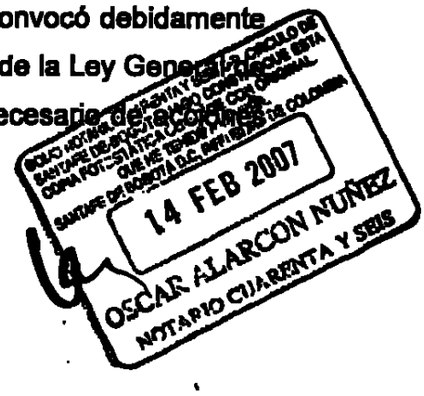
CERTIFICA POR LA PRESENTE:

PRIMERO: Que en una reunión de la Junta Directiva de Genentech, Inc. se adoptaron debidamente resoluciones que establecen una enmienda propuesta del certificado de Incorporación Corregido y vuelto a Redactar de dicha corporación, declarando dicha enmienda como aconsejable y llamando a una reunión de accionistas de dicha corporación para consideración de la misma. La resolución establece la enmienda propuesta de la siguiente manera:

RESUELVE, que el Certificado de Incorporación Corregido y vuelto a Redactar de esta corporación se enmiende cambiando "la Sección 4.01(a)" del Artículo allí numerado "4" tal que, como se enmienda, dicha Sección de dicho Artículo se lea de la siguiente manera:

> SECCION 4.01. *Acciones de Capital.* (a) la Corporación está autorizada a emitir dos clases de acciones que deben designarse, respectivamente, acciones preferenciales y acciones comunes. El número total de acciones que la Corporación esta autorizada a emitir es de mil trescientos millones (1,300,000,000) de acciones. Cien millones (100,000,000) de acciones serán designadas acciones preferenciales, por un valor de $0.02 por acción (**"Acción Preferencial"**). Mil doscientos millones (1,200,000,000) de acciones serán designadas acciones comunes, por un valor de $0.02 por acción (**"Acción Común"**). La acción común de la Corporación será toda de una misma clase.

SEGUNDO: Que por lo tanto, de acuerdo con la resolución de su Junta Directiva, la reunión ordinaria de accionistas de dicha corporación se convocó debidamente y se sostuvo al ser anunciada de acuerdo con la sección 222 de la Ley General de Corporación del Estado de Delaware, en donde el número necesario de acciones requerido por los estatutos votaron a favor de la enmienda.



TERCERO: Que dicha enmienda se adoptó debidamente de acuerdo con lo previsto en la Sección 242 de la Ley General de Corporación del Estado de Delaware.

CUARTO: Que el capital de dicha corporación no será reducido bajo o por razón de dicha enmienda

EN TESTIMONIO DE LO CUAL, dicha Genentech, Inc. ha causado este Certificado de Enmienda para que sea firmado por Stephen G. Juelsgaard, un Funcionario Autorizado, este 10 de Mayo de 2001.

GENENTECH, INC.

Por: (Firma)
Nombre: Stephen G. Juelsgaard
Título: Vicepresidente, Consejo
y Secretaría General.

14 FEB 2007
OSCAR ALARCON NUÑEZ
NOTARIO CUARENTA Y OCHO

97.200

1649

Apostille

(Convention de La Haye du 5 Octobre 1961)

1. *Country: United States of America*

This public document:

2. *has been signed by Harriet Smith Windsor*

3. *acting in the capacity of Secretary of State of Delaware*

4. *bears the seal/stamp of Office of Secretary of State*

Certified

5. *at Dover, Delaware*

6. *the twenty-fourth day of May, A.D. 2005*

7. *by Secretary of State, Delaware Department of State*

8. *No. 0254031*

9. *Seal/Stamp:*

10. *Signature:*

Harriet Smith Windsor

Secretary of State



Delaware

1650 + 35

The First State

I, HARRIET SMITH WINDSOR, SECRETARY OF STATE OF THE STATE OF DELANARE, DO HEREBY CERTIFY THE ATTACHED IS A TRUE AND CORRECT COPY OF THE CERTIFICATE OF AMENDMENT OF "GENENTECH, INC.", FILED IN THIS OFFICE ON THE FIFTEENTH DAY OF MAY, A.D. 2000, AT 5 O'CLOCK P.M.

2112549

050426506

Harriet Smith Windsor

Harriet Smith Windsor, Secretary of State

AUTHENTICATION: 3900385

DATE: 05-2[illegible]

14 FEB 2007

OSCAR ALARCON NUNEZ

NOTARIO CUARENTA Y SEIS

STATE OF DELAWARE
SECRETARY OF STATE
DIVISION OF CORPORATIONS
FILED 05:00 PM 05/15/2000
001245251 – 2112549

651
634

CERTIFICATE OF AMENDMENT OF AMENDED AND RESTATED CERTIFICATE OF INCORPORATION

Genentech, Inc. a corporation organized and existing under and by virtue of the General Corporation Law of the State of Delaware.

DOES HEREBY CERTIFY:

FIRST: That at a meeting of the Board of Directors of Genentech, Inc. resolutions were duly adopted setting forth a proposed amendment of the Amended and Restated Certificate of Incorporation of said corporation, declaring said amendment to be advisable and calling a meeting of the stockholders of said corporation for consideration thereof. The resolution setting forth the proposed amendment is as follows:

RESOLVED, that the Amended and Restated Certificate of Incorporation of this corporation be amended by changing "Section 4.01(a)" of the Article thereof numbered "4" so that, as amended, said Section of said Article shall be and read as follows:

> "Section 4.01. *Capital Stock.* (a) The Corporation is authorized to issue two classes of stock to be designated, respectively, preferred stock and common stock. The total number of shares which the Corporation is authorized to issue is seven hundred million (700,000,000) shares. One hundred million (100,000,000) shares shall be designated preferred stock, par value $0.02 per share ("Preferred Stock"). Six hundred million (600,000,000) shares shall be designated common stock, par value $0.02 per share ("Common Stock"). The Common Stock of the Corporation shall be all of one class."

SECOND: That thereafter, pursuant to resolution of its Board of Directors, the regular meeting of the stockholders of said corporation was duly called and held upon notice in accordance with Section 222 of the General Corporation Law of the State of Delaware at which meeting the necessary number of shares as required by statute were voted in favor of the amendment.

THIRD: That said amendment was duly adopted in accordance with the provisions of Section 242 of the General Corporation Law of the State of Delaware.

FOURTH: That the capital of said corporation shall not be reduced under or by reason of said amendment.

IN WITNESS WHEREOF, said Genentech, Inc. has caused this Certificate of Amendment to be signed by Stephen G. Juelsgaard, an Authorized Officer, this 15th day of May, 2000.

GENENTECH, INC.

By: Stephen Juelsgaard
Name: Stephen G. Juelsgaard
Title: Senior Vice President, General Counsel and Secretary

75841

14 FEB 2007
OSCAR ALARCON NUÑEZ

TRADUCCION AL ESPAÑOL del Certificado de Incorporación de Genentech, Inc.

1652

APOSTILLA

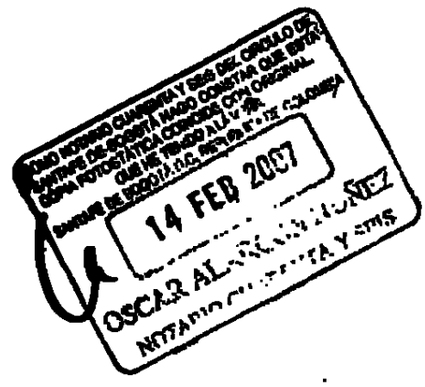
(Convención de La Haya del 5 de Octubre de 1961)

1. País: Estados Unidos de América
   Este documento público
2. ha sido firmado por Harriet Smith Windsor
3. actuando en capacidad de Secretaria de Estado de Delaware
4. lleva el sello/timbre de la Oficina del Secretario de Estado

Certificado

5. en Dover, Delaware
6. el veinticuatro de mayo, A. D. 2005
7. por el Secretario de Estado, Departamento de Estado de Delaware
8. No. 0254031
9. Sello/timbre
10. Firma

Harriet Smith Windsor (Firma)
Secretaria de Estado



1653

DELAWARE

EL PRIMER ESTADO

YO, HARRIET SMITH WINDSOR, SECRETARIA DE ESTADO DEL ESTADO DE DELAWARE, CERTIFICO POR LA PRESENTE QUE EL ANEXO ES UNA COPIA AUTÉNTICA Y CORRECTA DEL CERTIFICADO VUELTO A REDACTAR DE "GENENTECH, INC", PRESENTADO EN ESTA OFICINA EL DECIMOQUINTO DIA DE MAYO, A. D. 2000, A LAS 5 EN PUNTO DE LA TARDE.



Harriet Smith Windsor (Firma)
Harriet Smith Windsor, Secretaria de Estado
AUTENTICACIÓN: 3900385

FECHA: 05 – 24 – 05

1654

DELAWARE

EL PRIMER ESTADO

YO, HARRIET SMITH WINDSOR, SECRETARIA DE ESTADO DEL ESTADO DE DELAWARE, CERTIFICO POR LA PRESENTE QUE EL ANEXO ES UNA COPIA AUTÉNTICA Y CORRECTA DEL CERTIFICADO VUELTO A REDACTAR DE "GENENTECH, INC", PRESENTADO EN ESTA OFICINA EL VIGÉSIMO SEGUNDO DÍA DE JULIO, A. D. 1999, A LAS 9 EN PUNTO DE LA MAÑANA.

14 FEB 2007

Harriet Smith Windsor (Firma)
Harriet Smith Windsor, Secretaria de Estado
AUTENTICACIÓN: 3900384

FECHA: 05 – 24 – 05

1655

CERTIFICADO DE ENMIENDA DEL CERTIFICADO DE INCORPORACIÓN CORREGIDO Y VUELTO A REDACTAR

Genentech, Inc. una corporación organizada y existente bajo y por virtud de la Ley General de Corporación del Estado de Delaware.

CERTIFICA POR LA PRESENTE:

PRIMERO: Que en una reunión de la Junta Directiva de Genentech, Inc. se adoptaron debidamente resoluciones que establecen una enmienda propuesta del certificado de Incorporación Corregido y vuelto a Redactar de dicha corporación, declarando dicha enmienda como aconsejable y llamando a una reunión de accionistas de dicha corporación para consideración de la misma. La resolución establece la enmienda propuesta de la siguiente manera:

RESUELVE, que el Certificado de Incorporación Corregido y vuelto a Redactar de esta corporación se enmiende cambiando "la Sección 4.01(a)" del Artículo allí numerado "4" tal que, como se enmienda, dicha Sección de dicho Artículo se lea de la siguiente manera:

> SECCION 4.01. *Acciones de Capital.* (a) la Corporación está autorizada a emitir dos clases de acciones que deben designarse, respectivamente, acciones preferenciales y acciones comunes. El número total de acciones que la Corporación esta autorizada a emitir es de setecientos millones (700,000,000) de acciones. Cien millones (100,000,000) de acciones serán designadas acciones preferenciales, por un valor de $0.02 por acción ("**Acción Preferencial**"). Seiscientos millones (600,000,000) de acciones serán designadas acciones comunes, por un valor de $0.02 por acción ("**Acción Común**"). La acción común de la Corporación será toda de una misma clase.

SEGUNDO: Que por lo tanto, de acuerdo con la resolución de su Junta Directiva, la reunión ordinaria de accionistas de dicha corporación se convocó debidamente y se sostuvo al ser anunciada de acuerdo con la sección 222 de la Ley General de Corporación del Estado de Delaware, en donde el número necesario de acciones requerido por los estatutos votaron a favor de la enmienda.

14 FEB 2007

OSCAR ALARCON NUÑEZ

NOTARIO CUARENTA Y SEIS

1656
+643

TERCERO: Que dicha enmienda se adoptó debidamente de acuerdo con lo previsto en la Sección 242 de la Ley General de Corporación del Estado de Delaware.

CUARTO: Que el capital de dicha corporación no será reducido bajo o por razón de dicha enmienda

EN TESTIMONIO DE LO CUAL, dicha Genentech, Inc. ha causado este Certificado de Enmienda para que sea firmado por Stephen G. Juelsgaard, un Funcionario Autorizado, este 15 de Mayo de 2000.

GENENTECH, INC.

Por: (Firma)
Nombre: Stephen G. Juelsgaard
Título: Vicepresidente, Consejo
y Secretaría General.

14 FEB 2007

OSCAR ALARCON NUÑEZ

1657

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
Rad: 07-033379- -00000-0000
Fec: 2007-04-02 17:56:24 Dep: 2020 NUEVECREACIONES
Tra: 11 PATENTEIYE
Cve: 378 FASENACIONALI
Act: 411 PRESENTACION Folios: 1543

Industria y Comercio
SUPERINTENDENCIA

REQUISITOS MINIMOS PARA ADMISION A TRAMITE
PCT

PATENTE DE INVENCION [X] MODELO DE UTILIDAD [ ]

- Indicación de que se solicita la concesión de una patente [X]
- Datos de identificación del solicitante o de la persona que presenta la solicitud [X]
- Descripción de la invención [X]
- Dibujos de ser estos pertinentes [X]
- Comprobante de pago de las tasas establecidas [X]

COMPLETA [ ] INCOMPLETA [ ]

DISEÑO INDUSTRIAL [ ]

- Indicación de que se solicita el registro de Diseño Industrial [ ]
- Datos de identificación del solicitante o de la persona que presenta la solicitud [ ]
- Representación gráfica y fotográfica del Diseño Industrial o muestra del material que incorpora el diseño [ ]
- Comprobante de pago de las tasas establecidas [ ]

COMPLETA [ ] INCOMPLETA [ ]

ESQUEMA DE TRAZADO [ ]

- Indicación de que se solicita el registro de un Esquema de Trazado [ ]
- Datos de identificación del solicitante o de la persona que presenta la solicitud [ ]
- Representación gráfica del esquema trazado [ ]
- Comprobante de pago de las tasas establecidas [ ]

COMPLETA [ ] INCOMPLETA [ ]

Firma Responsable

Fecha 02 ABR. 2007

Carrera 13 No. 27-00 Piso 5, 7 y 10
Conmutador: 3 82 08 40
Fax: 350 52 20
E-mail: info@sic.gov.co
www.sic.gov.co
Bogotá, D.C., Colombia

1658

REPUBLICA DE COLOMBIA

SUPERINTENDENCIA DE INDUSTRIA Y COMERCI

Bogotá,
2020

Doctor(a)
CADENA SARMIENTO JUAN PABLO
Apoderado(a) y/o Representante de
GENENTECH, INC.

REFERENCIA	Radicación No.	7 33379
	Solicitud internacional	PCT/US2005/031798
	Fecha inicio fase nacional	08/04/2007
	Trámite	11
	Evento	378
	Actuación	430
	Oficio No.	2020-5783

Como resultado del examen de forma, realizado con fundamento en el artículo 38 de la Decisión 486 de la Comisión de la Comunidad Andina, artículo 27 del Tratado de Cooperación en Materia de Patentes (PCT), regla 51 bis del Reglamento y Circular Única de la Superintendencia de Industria y Comercio, se le comunica que:

1. La solicitud no contiene la copia del documento en que consta la cesión del derecho a la patente del inventor al solicitante o a su causante. (Art. 26-k) Decisión 486.

En cumplimiento de lo dispuesto por el artículo 39 de la Decisión 486, se le requiere para que dentro del término de dos meses siguientes a la fecha de notificación del presente oficio, complete los requisitos anteriormente enunciados. En caso de no dar cumplimiento al presente requerimiento, la solicitud se considerará abandonada y perderá su prelación.

Notifíquese el presente oficio conforme a lo dispuesto en el numeral 5.2, literal e), del capítulo quinto del título primero de la Circular Unica No.10 de 2001.

ALIX CARMENZA CESPEDES DE VERGEL
Jefe de la división de Nuevas Creaciones

El anterior requerimiento fue notificado por fijación en lista de fecha 11 MAR 2020
adpa11