BT

Industria y Comercio

SUPERINTENDENCIA

SOLICITUD

PATENTE DE INVENCIÓN

21 EXPEDIENTE Nº 03-97535

54. TÍTULO VACUNAS DE ACIDO NUCLEICO PARA PREVENCION DE LA INFECCION DE FLAVIVIRUS

51 CLASIFICACIÓN INTERNACIONAL C12Q 1/70, A61K 39/12

71 SOLICITANTE THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES, CENTERS FOR DISEASE CONTROL AND PREVENTION

DOMICILIO GEORGIA, ESTADOS UNIDOS DE AMERICA

74. APODERADO ALVARO CORREA ORDOÑEZ
T P 20 261

22. BOGOTÁ, D C. NOVIEMBRE 4 DE 2003

AA

PETITORIO

Industria y Comercio
SUPERINTENDENCIA

03 097535 00

SUPERINTENDENCIA Y COMERCIO
Radicacion : 03097535 00000000 Folios: 422
Fecha (AMD): 2003-11-04 16:26:37
Tramite : 011 PCT-PATENT 379 FASE NACI 411 PRESENTACIÓN
Dependencia: 2020 DIVISION DE NUEVAS CREACIONES

FORMULARIO ÚNICO DE SOLICITUD DE PATENTE
PCT
ENTRADA EN FASE NACIONAL

① SOLICITUD DE :

[X]	Patente de Invención	[ ]	Patente de Modelo de Utilidad
[ ]	Capítulo I.	[X]	Capítulo II.

② SOLICITANTE (71)

Nombre: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES, CENTERS FOR DISEASE CONTROL AND PREVENTION.
Dirección: Technology Transfer Office, 1600 Clifton Road, N.E., MS E-67, Atlanta , Georgia 30333, Estados Unidos de América.
Nacionalidad o Domicilio: ESTADOUNIDENSE
Lugar de Constitución: Georgia, Estados Unidos de América.
Teléfono: Fax:
E-Mail

IDENTIFICACION
C.C. [ ] NIT [ ]
C.E. [ ] Otro [ ]
Cual
Número

③ REPRESENTANTE O APODERADO

Nombre: ALVARO CORREA ORDOÑEZ
Dirección: Avda 82 No. 10-62 Piso 6
Teléfono: 634-1500 Fax: 376-2211
E-Mail: patents.bogota@bakernet.com

IDENTIFICACION
C.C. [X] NIT [ ]
C.E. [ ] Otro [ ]
Cual
Número 79.143.366 de Usaquén TP 20.281

④ INVENTOR (ES) (72)

Nombre: CHANG, Gwong-Jen, J.
Dirección: 4237 Beaver Creek Drive, Fort Collins, Colorado 80522, Estados Unidos de América
Nacionalidad o Domicilio: ESTADOUNIDENSE

⑤ PCT (85)

Solicitud Internacional No. PCT/US02/10764 Fecha Abril 4, 2002
Publicación Internacional No. WO 02/081754 Fecha Octubre 17, 2002

⑥ Título (54) VACUNAS DE ÁCIDO NUCLÉICO PARA PREVENCIÓN DE LA INFECCIÓN DE FLAVIVIRUS.

⑦ Clasificación Internacional (51) C12 Q 1/70 ; A61K 39/12 ; C07H 21/04 ; G01N 33/53.

⑧ Prioridad	(33) País de Origen	(32) Fecha	(31) Número de Solicitud
SI [X] NO [ ]	US	Abril 4, 2001	09/826,115

⑨ Comprobante de pago No. 74,978 y 74,979 **Fecha** OCTUBRE 15 DE 2003

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

003092

10 ANEXOS

- [X] Comprobante de pago de la tasa de presentación de la solicitud
- [ ] Documento que acredita la existencia y representación legal cuando el solicitante sea persona jurídica
- [ ] Poderes, si fuere el caso
- [ ] 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 internacional efectuado por la Autoridad Internacional de Examen preliminar (IPEA)
- [ ] Traducción del examen preliminar internacional 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 12 x 12
- [ ] 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
- [ ] De ser el caso, modificaciones efectuadas a la solicitud internacional.
- [ ] Otros

11 FIGURA CARATERISTICA

12 Solicito la concesión de la patente,

NOMBRE: ALVARO CORREA ORDOÑEZ

FIRMA:

C.C. 79.143.366 de Usaquén T.P. 20.281

000003

SUPERINDUSTRIA Y COMERCIO
Radicacion : 03097535 00000000 Folios: 422
Fecha (AMD): 2003-11-04 16:26:37
Tramite : 011 PCT-PATENT 379 FASE NACI 411 PRESENTAC
Dependencia: 2020 DIVISION DE NUEVAS CREACIONES

NIT : 800176089-2

RECIBO OFICIAL DE CAJA : 03 - 74,970
FECHA : OCTUBRE 15 DE 2003

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

DEPOSITANTE	TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAG	Vr. PAGO
JORGE LUIS GOMEZ	CONSIGNACION	BANCO POPULAR	050-00110-6	11025315	15/10/2003	64,120,000.00

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

CANT.	RENTISTICO	CONCEPTO	TOTAL CONCEPTO
1	50005-01-06 REGISTRO DE CESION, CAMBIO NOM	13 PATENTES,MODELOS,DISENOS,LEMAS,NO	202,000.00
		TOTAL : $	202,000.00

SON: DOSCIENTOS DOS MIL PESOS

RESPONSABLE : _____

RECIBO DE CAJA APLICADO AL EXPEDIENTE No._____

NIT : 800174089-2

SUPERINDUSTRIA Y COMERCIO
Radicacion : 03097535 00000000
Fecha (MMM): 2003-11-04 16:26:37
Tramite : 011 PCT-PATENT 379 FASE NACI 411 PRESENTAC
Dependencia: 2020 DIVISION DE NUEVAS CREACIONES

RECIBO OFICIAL DE CAJA : 03 - 74,979
FECHA : OCTUBRE 15 DE 2003

Folios: 422

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

DEPOSITANTE	TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PSO	Vr. PAGO
JORGE LUIS GOMEZ	CONSIGNACION	BANCO POPULAR	050-00110-6	11025315	15/10/2003	64,120,000.00

C O N C E P T O

CANT.	RENTISTICO	CONCEPTO	TOTAL CONCEPTO
1	50005-01-06 REGISTRO DE CESION, CAMBIO NOM	13 PATENTES,MODELOS,DISENOS,LEMAS,NO	202,000.00
		TOTAL : $	202,000.00

SON: DOSCIENTOS DOS MIL PESOS

RESPONSABLE : _____

RECIBO DE CAJA APLICADO AL EXPEDIENTE No._____

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 October 2002 (17.10.2002)

PCT

(10) International Publication Number
WO 02/081754 A1

(51) International Patent Classification[7]: C12Q 1/70, G01N 33/53, C12P 21/06, C07K 16/00, C07H 21/04, A61K 39/12

(21) International Application Number: PCT/US02/10764

(22) International Filing Date: 4 April 2002 (04.04.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/826,115 4 April 2001 (04.04.2001) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 09/826,115 (CIP)
Filed on 4 April 2001 (04.04.2001)

(71) Applicant *(for all designated States except US)*: THE GOVERNEMENT OF THE UNITED STATES OF AMERICA, as represented by the secretary, DEPARTMENT OF HEALTH AND HUMAN SERVICES, CENTERS FOR DISEASE CONTROL AND PREVENTION [US/US]; Technology Transfer Office, 1600 Clifton Road, N.E., MS E-67, Atlanta, GA 30333 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: CHANG, Gwong-Jen, J. [US/US]; 4237 Beaver Creek Drive, Fort Collins, CO 80522 (US).

(74) Agent: MILLER, Mary, L.; Needle & Rosenberg, P.C., 127 Peachtree Street, N.E., Suite 1200, Atlanta, GA 30303-1811 (US).

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

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search 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.

WO 02/081754 A1

(54) Title: NUCLEIC ACID VACCINES FOR PREVENTION OF FLAVIVIRUS INFECTION

(57) Abstract: The present invention encompasses isolated nucleic acids containing transcriptional units which encode a signal sequence of one flavivirus and an immunogenic flavivirus antigen of a second flavivirus or of a chimeric immunogenic flavivirus antigen comprising sequence from more than one flavivirus. The invention further encompasses a nucleic acid and protein vaccine and the use of the vaccine to immunize a subject against flavivirus infection. The invention also provides antigens encoded by nucleic acids of the invention, antibodies elicited in response to the antigens and use of the antigens and/or antibodies in detecting flavivirus or diagnosing flavivirus infection.

NUCLEIC ACID VACCINES FOR PREVENTION OF FLAVIVIRUS INFECTION

This application is a continuation-in-part of, and claims the benefit of, U.S. application Serial No. 09/826,115, filed April 4, 2001, which status is pending and which is continuation-in-part of, and claims the benefit of U.S. application Serial No. 09/701,536, filed November 29, 2000, which status is pending and which is a national stage application of international application Serial No. PCT/US99/12298, filed June 3, 1999 from, and which claims the benefit of, U.S. provisional application Serial No. 60/087,908, filed June 4, 1998, which applications are hereby incorporated herein in their entirety by reference.

Field of the Invention

This invention relates to novel vaccines, diagnostics and methods of using both in the treatment and prevention of the diseases caused by flaviviruses. In particular, the vaccines are recombinant nucleic acids which contain genes for structural proteins of flaviviruses, such as Japanese encephalitis virus (JEV), West Nile virus (WNV) or related flaviviruses. These vaccines serve as a transcriptional unit for the biosynthesis of the virus protein antigens when administered *in vivo*. The diagnostics are compositions containing antigens produced from the recombinant nucleic acids that can be used to detect flavivirus infection.

Background of the Invention

Flaviviruses are members of the genus *Flavivirus*, which is classified within the family Flaviviridae. The flaviviruses are largely pathogenic to humans and other mammals. Flaviviruses that inflict disease upon humans and animals include Alfuy, Apoi, Aroa, Bagaza, Banzi, Batu Cave, Bouboui, Bukalasa bat, Bussuquara, Cacipacore, Carey Island, Cowbone Ridge, Dakar bat, Dengue (serotypes 1, 2, 3 and 4), Edge Hill, Entebbe bat, Gadgets Gully, Iguape, Ilheus, Israel turkey

meningoencephalitis, Japanese encephalitis, Jugra, Jutiapa, Kadam, Karshi, Kedougou, Kokobera, Koutango, Kunjin, Kyasanur Forest disease, Langat, Meaban, Modoc, Montana myotis leukoencephalitis, Murray Valley encephalitis, Naranjal, Negishi, Ntaya, Omsk hemorrhagic fever, Phnom Penh bat, Potiskum, Powassan, Rio Bravo, Rocio, Royal Farm, Russian spring summer encephalitis, Saboya, Sal Vieja, San Perlita, Saumarez Reef, Sepik, Sokuluk, Spondweni, St. Louis encephalitis, Stratford, Tick-borne encephalitis - central European subtype, Tick-borne encephalititis - far eastern subtype, Tembusu, THCAr, Tyuleniy, Uganda S, Usutu, West Nile, Yaounde, Yellow fever, Yokose, Ziki, Cell fusing agent and other related flaviviruses, as listed in Kuno et al. (*J. Virol.* 72: 73-83 (1998)).

The flaviviruses contain the following three structural proteins: prM/M, the premembrane and membrane protein; E, the envelope protein; and C, the capsid protein. (Monath, in Virology (Fields, ed.), Raven Press, New York, 1990, pp. 763-814; Heinz and Roehrig, in Immunochemistry of Viruses II: The Basis for Serodiagnosis and Vaccines (van Regenmortel and Neurath, eds.), Elsevier, Amsterdam, 1990, pp. 289-305). M has a molecular weight (MW) of about 7-8 kilodaltons (kDa) and E has a MW of about 55-60 kDa. M is synthesized as a larger precursor termed prM. The pr portion of prM is removed when prM is processed to form M protein in mature virions. M and E are located in the membrane of the flavivirus particle, and so have long been considered to constitute important immunogenic components of the viruses.

The flaviviruses are RNA viruses comprising single stranded RNA having a length, among the various species, of about 10 kilobases (kb). The C protein, with a MW of 12-14 kDa, complexes with the RNA to form a nucleocapsid complex. Several nonstructural proteins are also encoded by the RNA genome which are termed NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5. The genome is translated within the host cell as a polyprotein, then processed co- or post-translationally into the individual gene products by viral- or host-specific proteases (Figure 1).

The nucleotide sequences of the genomes of several flaviviruses are known, as summarized in U. S. Patent No. 5,494,671. That for JEV is provided by Sumiyoshi et al. (*Virology* 161: 497-510 (1987)) and Hashimoto et al. (*Virus Genes* 1: 305-317 (1988)). The nucleotide sequences of the virulent strain SA-14 of JEV and the attenuated strain SA-14-14-2, used as a vaccine in the People's Republic of China, are compared in the work of Nitayaphan et al. (*Virology* 177: 541-552 (1990)).

Nucleotide sequences encoding the structural proteins of other flavivirus species are also known. In many cases, the sequences for the complete genomes have been reported. The sequences available include dengue serotype 1 virus, dengue serotype 2 virus (Deubel et al., *Virology* 155: 365-377 (1986); Gruenberg et al., *J. Gen. Virol.* 69: 1391-1398 (1988); Hahn et al. *Virology* 162: 167-180 (1988)), dengue serotype 3 virus (Osatomi et al., *Virus Genes* 2: 99-108 (1988)), dengue serotype 4 virus (Mackow et al., *Virology* 159: 217-228 (1987), Zhao et al., *Virology* 155: 77-88 (1986)), West Nile virus (Lanciotti et al., *Science* 286: 2331-2333 (1999)), Powassan virus (Mandl et al., *Virology* 194: 173-184 (1993)) and yellow fever virus (YFV) (Rice et al., *Science* 229: 726-733 (1985)).

Many flaviviruses, including St. Louis encephalitis virus (SLEV), WNV and JEV, are transmitted to humans and other host animals by mosquitoes. They therefore occur over widespread areas and their transmission is not easily interrupted or prevented.

West Nile fever is a mosquito-borne flaviviral infection that is transmitted to vertebrates primarily by various species of *Culex* mosquitoes. Like other members of the Japanese encephalitis (JE) antigenic complex of flaviviruses, including JE, SLE and Murray Valley encephalitis (MVE) viruses, WNV is maintained in a natural cycle between arthropod vectors and birds. The virus was first isolated from a febrile human in the West Nile district of Uganda in 1937 (Smithburn et al., *Am. J. Trop. Med. Hyg.* 20: 471-492 (1940)). It was soon recognized as one of the most widely distributed flaviviruses, with its geographic range including Africa, the Middle East, Western Asia, Europe and Australia (Hubalek et al., *Emerg. Infect. Dis.* 5: 643-50 (1999)). Clinically,

West Nile fever in humans is a self-limited acute febrile illness accompanied by headache, myalgia, polyarthropathy, rash and lymphadenopathy (Monath and Tsai, in Clinical Virology, (Richman, Whitley and Hayden eds.), Churchill-Livingtone, New York, 1997, pp. 1133-1186). Acute hepatitis or pancreatis has been reported on occasion and cases of WNV infection in elderly patients are sometimes complicated by encephalitis or meningitis (Asnis et al., *Clin. Infect. Dis.* 30: 413-418 (2000)). Thus, infection by WNV is a serious health concern in many regions of the world.

The geographical spread of the disease, particularly the introduction of WNV into the U.S. in 1999, has greatly increased awareness of the human and animal health concerns of this disease. Between late August and early September 1999, New York City and surrounding areas experienced an outbreak of viral encephalitis, with 62 confirmed cases, resulting in seven deaths. Concurrent with this outbreak, local health officials observed increased mortality among birds (especially crows) and horses. The outbreak was subsequently shown to be caused by WNV, based on monoclonal antibody (Mab) mapping and detection of genomic sequences in human, avian and mosquito specimens (Anderson et al., *Science* 286: 2331-2333 (1999); Jia et al., *Lancet* 354: 1971-1972 (1999); Lanciotti et al., *Science* 286: 2333-2337 (1999)). Virus activity detected during the ensuing winter months indicated that the virus had established itself in North America (*Morb. Mortal. Wkly. Rep.* 49: 178-179 (2000); Asnis et al., *Clin. Infect. Dis.* 30: 413-418 (2000); Garmendia et al., *J. Clin. Micro.* 38: 3110-3111 (2000)). Surveillance data reported from the northeastern and mid-Atlantic states during the year 2000 confirmed an intensified epizootic/epidemic transmission and a geographic expansion of the virus with documentation of numerous cases of infection in birds, mosquitoes and horses, as well as cases in humans (*Morb. Mortal. Wkly. Rep.* 49: 820-822 (2000)).

Currently, no human or veterinary vaccine is available to prevent WNV infection and mosquito control is the only practical strategy to combat the spread of the disease.

Japanese encephalitis virus (JEV) infects adults and children and there is a high mortality rate among infants, children and the elderly in areas of tropical and subtropical Asia (Tsai et al., in Vaccines (Plotkin, ed.) W.B. Saunders, Philadelphia, Pa, 1999, pp. 672-710). Among survivors, there are serious neurological consequences, related to the symptoms of encephalitis, that persist after infection. In more developed countries of this region, such as Japan, the Republic of China (Taiwan) and Korea, JEV has been largely controlled by use of a vaccine of inactivated JEV. Nevertheless, it is still prevalent in other countries of the region.

Vaccines available for use against JEV infection include live virus inactivated by such methods as formalin treatment, as well as attenuated virus (Tsai et al., in Vaccines (Plotkin, ed.) W.B. Saunders, Philadelphia, Pa, 1994, pp. 671-713). Whole virus vaccines, although effective, do have certain problems and/or disadvantages. The viruses are cultivated in mouse brain or in cell culture using mammalian cells as the host. Such culture methods are cumbersome and expensive. Furthermore, there is the attendant risk of incorporating antigens from the host cells, i.e., the brain or other host, into the final vaccine product, potentially leading to unintended and undesired allergic responses in the vaccine recipients. There is also the risk of inadvertent infection among workers involved in vaccine production. Finally, there is the risk that the virus may not be fully or completely inactivated or attenuated and thus, the vaccine may actually cause disease.

Dengue fever and dengue hemorrhagic fever (DF/DHF) are caused by dengue virus, which is also a mosquito-borne flavivirus. There are four antigenically related, but distinct, dengue virus serotypes, (DEN-1, DEN-2, DEN-3 and DEN-4), all of which can cause DF/DHF. Symptoms of DF, the mild form of dengue-related disease, include fever, rash, severe headache and joint pain. Mortality among those subjects suffering from DF is low; however, among those subjects suffering from DHF, mortality can be as high as 5%. From available evidence, more than 3 million cases of DHF and 58,000 deaths have been attributed to DHF over the past 40 years, making DHF a major emerging disease (Halstead, in Dengue and Dengue Hemorrhagic Fever (Gubler and Kuno, eds.) CAB International, New York, NY, (1997) pp 23-44). Nevertheless,

despite decades of effort, safe and effective vaccines to protect against dengue virus infection are not yet available.

Yellow fever is prevalent in tropical regions of South America and sub-Saharan Africa and is transmitted by mosquitos. Infection leads to fever, chills, severe headache and other pains, anorexia, nausea and vomiting, with the emergence of jaundice. A live virus vaccine, 17D, grown in infected chicken embryos, is considered safe and effective. Nevertheless, there remains a need for a vaccine that is stable under adverse conditions, such as are commonly encountered in the tropical regions of Africa and the Americas where the vaccine is most needed.

A recombinant flavivirus which is a chimera between two flaviviruses is disclosed in PCT publication WO 93/06214. The chimera is a construct fusing non-structural proteins from one "type," or serotype, of dengue virus or a flavivirus, with structural proteins from a different "type," or serotype, of dengue virus or other flavivirus.

Several recombinant subunit and viral vaccines have been devised in recent years. U.S. Patent No. 4,810,492 describes the production of the E glycoprotein of JEV for use as the antigen in a vaccine. The corresponding DNA is cloned into an expression system in order to express the antigen protein in a suitable host cell such as *E. coli*, yeast, or a higher organism cell culture. U. S. Patent No. 5,229,293 discloses recombinant baculovirus harboring the gene for JEV E protein. The virus is used to infect insect cells in culture such that the E protein is produced and recovered for use as a vaccine.

U.S. Patent No. 5,021,347 discloses a recombinant vaccinia virus genome into which the gene for JEV E protein has been incorporated. The live recombinant vaccinia virus is used as the vaccine to immunize against JEV. Recombinant vaccinia viruses and baculoviruses in which the viruses incorporate a gene for a C-terminal truncation of the E protein of dengue serotype 2, dengue serotype 4 and JEV are disclosed in U.S. Patent 5,494,671. U.S. Patent 5,514,375 discloses various

recombinant vaccinia viruses which express portions of the JEV open reading frame extending from prM to NS2B. These pox viruses induced formation of extracellular particles that contain the processed M protein and the E protein. Two recombinant viruses encoding these JEV proteins produced high titers of neutralizing and hemagglutinin-inhibiting antibodies, and protective immunity, in mice. The extent of these effects was greater after two immunization treatments than after only one. Recombinant vaccinia virus containing genes for the prM/M and E proteins of JEV conferred protective immunity when administered to mice (Konishi et al., *Virology* 180: 401-410 (1991)). HeLa cells infected with recombinant vaccinia virus bearing genes for prM and E from JEV were shown to produce subviral particles (Konishi et al., *Virology* 188: 714-720 (1992)). Dmitriev et al. reported immunization of mice with a recombinant vaccinia virus encoding structural and certain nonstructural proteins from tick-borne encephalitis virus (*J. Biotechnology* 44: 97-103 (1996)).

Recombinant virus vectors have also been prepared to serve as virus vaccines for dengue fever. Zhao et al. (*J. Virol.* 61: 4019-4022 (1987)) prepared recombinant vaccinia virus bearing structural proteins and NS1 from dengue serotype 4 and achieved expression after infecting mammalian cells with the recombinant virus. Similar expression was obtained using recombinant baculovirus to infect target insect cells (Zhang et al., *J. Virol.* 62: 3027-3031(1988)). Bray et al. (*J. Virol.* 63: 2853-2856 (1989)) also reported a recombinant vaccinia dengue vaccine based on the E protein gene that confers protective immunity to mice against dengue encephalitis when challenged. Falgout et al. (*J. Virol* 63: 1852-1860 (1989)) and Falgout et al. (*J. Virol.* 64: 4356-4363 (1990)) reported similar results. Zhang et al. (*J. Virol* 62: 3027-3031 (1988)) showed that recombinant baculovirus encoding dengue E and NS1 proteins likewise protected mice against dengue encephalitis when challenged. Other combinations in which structural and nonstructural genes were incorporated into recombinant virus vaccines failed to produce significant immunity (Bray et al., *J. Virol.* 63: 2853-2856 (1989)). Also, monkeys failed to develop fully protective immunity to dengue virus challenge when immunized with recombinant baculovirus expressing the E protein (Lai et al. (1990) pp. 119-124 in F. Brown, R. M. Chancock, H. S. Ginsberg

and R. Lerner (eds.) Vaccines 90: Modern approaches to new vaccines including prevention of AIDS, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).

Immunization using recombinant DNA preparations has been reported for SLEV and dengue-2 virus, using weanling mice as the model (Phillpotts et al., *Arch. Virol.* 141: 743-749 (1996); Kochel et al., *Vaccine* 15: 547-552 (1997)). Plasmid DNA encoding the prM and E genes of SLEV provided partial protection against SLEV challenge with a single or double dose of DNA immunization. In these experiments, control mice exhibited about 25% survival and no protective antibody was detected in the DNA-immunized mice (Phillpotts et al., *Arch. Virol.* 141: 743-749 (1996)). In mice that received three intradermal injections of recombinant dengue-2 plasmid DNA containing prM, 100% developed anti-dengue-2 neutralizing antibodies and 92% of those receiving the corresponding E gene likewise developed neutralizing antibodies (Kochel et al., *Vaccine* 15: 547-552 (1997)). Challenge experiments using a two-dose schedule, however, failed to protect mice against lethal dengue-2 virus challenge.

The vaccines developed to date for immunizing against infection by JEV, SLEV, dengue virus and other flaviviruses have a number of disadvantages and problems attending their use. Inactivated vaccine is costly and inconvenient to prepare. In addition, any such vaccine entails the risk of allergic reaction originating from proteins of the host cell used in preparing the virus. Furthermore, such vaccines present considerable risk to the workers employed in their production. Candidate attenuated JEV vaccines are undergoing clinical trials, but as of 1996 have not found wide acceptance outside of the People's Republic of China (Hennessy et al., *Lancet* 347: 1583-1586 (1996)).

Recombinant vaccines based on the use of only certain proteins of flaviviruses, such as JEV, produced by biosynthetic expression in cell culture with subsequent purification or treatment of antigens, do not induce high antibody titers. Also, like the whole virus preparations, these vaccines carry the risk of adverse allergic reaction to antigens from the host or to the vector. Vaccine development against dengue virus and

WNV is less advanced and such virus-based or recombinant protein-based vaccines face problems similar to those alluded to above.

There is therefore a need for vaccines or improved vaccines directed against flaviviruses such as yellow fever virus, dengue virus, JEV, SLEV and WNV which are inexpensive to prepare, present little risk to workers involved in their manufacture, carry minimal risk of adverse immunological reactions due to impurities or adventitious immunogenic components and are highly effective in eliciting neutralizing antibodies and protective immunity. There is furthermore a need for a vaccine against JEV, WNV and related flaviviruses that minimizes the number of immunizing doses required.

Many of the shortcomings of the current art as described in detail for the production of vaccines also apply to the production of antigens and antibodies to be used for the production of immunodiagnostics. Particularly, the concurrent risks and costs involved in the production of antigens from viruses and the failure of most currently available recombinantly expressed antigens to elicit effective immune responses are paralleled in the field of immunodiagnostics by the same risks, high costs and a corresponding lack of sensitivity. Thus, because of the high costs, risk of accidental infection with live virus and the lower than desired levels of sensitivity of the previously available tests, there exists a need for rapid, simple and highly sensitive diagnostic tests for detecting flavivirus infection and/or contamination.

The present invention meets these needs by providing highly immunogenic recombinant antigens for use in diagnostic assays for the detection of antibodies to selected flaviviruses. The present invention further provides for the use of recombinant antigens derived from flaviviruses, flavivirus genes or mimetics thereof in immunodiagnostic assays for the detection of antibodies to flavivirus proteins.

Summary of the Invention

The present invention provides a nucleic acid molecule which contains a transcriptional unit (TU) for an immunogenic flavivirus antigen. The TU directs a host

cell, after being incorporated within the cell, to synthesize the antigen. In an important aspect of the invention, the flavivirus can be yellow fever virus (YFV), dengue serotype 1 virus (DEN-1), dengue serotype 2 virus (DEN-2), dengue serotype 3 virus (DEN-3), dengue serotype 4 virus (DEN-4), St. Louis encephalitis virus (SLEV), Japanese encephalitis virus (JEV), West Nile virus (WNV), Powassan virus or any other flavivirus. In important embodiments of the present invention, the antigen can be the flavivirus prM/M protein, the E protein, or both. In important embodiments of the present invention, the antigen can be a chimeric flavivirus protein. In particular, when the TU includes both the prM/M and E proteins, the host cell secretes subviral particles containing the prM/M and E antigens. In a further important aspect of the invention, the nucleic acid is a DNA molecule. In additional significant embodiments, the nucleic acid TU includes a control sequence disposed appropriately such that it operably controls the expression of the prM/M and E antigens and this control sequence can be the cytomegalovirus immediate early promoter. In an additional embodiment, the nucleotide sequence of the TU is engineered to optimize eukaryotic translation by minimizing large hairpin structures in the 5'-end untranslated region of an mRNA produced by the TU and/or the inclusion of a Kozak consensus sequence at the translational start site of an mRNA produced by the TU. In an additional embodiment, the transcriptional unit also includes a poly-A terminator.

The present invention further provides a host cell comprising a nucleic acid molecule which includes a transcriptional unit for an immunogenic flavivirus antigen that directs the host cell to synthesize the immunogenic antigen. The flavivirus may be YFV, DEN-1, DEN-2, DEN-3, DEN-4, SLEV, JEV, WNV, Powassan virus or other flavivirus. In important embodiments, the antigen may be the prM/M protein, the E protein, or both the prM/M and the E proteins. In the latter case, the cell secretes subviral particles containing the prM/M and E antigens.

Additionally, the invention provides a composition for vaccinating a subject against a flavivirus containing a nucleic acid molecule that includes a transcriptional unit for an immunogenic flaviviral antigen. The transcriptional unit directs a cell within the body of the subject, after being incorporated therein, to synthesize the

immunogenic antigen. The composition further includes a pharmaceutically acceptable carrier. In significant embodiments, the flavivirus may be YFV, DEN-1, DEN-2, DEN-3, DEN-4, SLEV, JEV, WNV, Powassan virus or other flavivirus. Furthermore, the antigen may be the prM/M protein, the E protein, or both the prM/M and the E proteins. In the latter instance, the cell secretes subviral particles comprising the flavivirus prM/M and E antigens. These subviral particles are also referred to as noninfectious recombinant antigen (NRA). In important embodiments, the nucleic acid molecule is a DNA molecule. In further significant embodiments, the transcriptional unit additionally contains a control sequence disposed appropriately such that it operably controls the synthesis of the prM/M and E antigens when the nucleic acid is introduced into the cell of the subject. This control sequence can be the cytomegalovirus immediate early promoter. In a still further embodiment, the transcriptional unit can also include a poly-A terminator.

The compositions provided by the present invention for vaccinating a subject against a flavivirus can include a nucleic acid molecule, or molecules, that include transcriptional units for more than one immunogenic flaviviral antigen. The more than one immunogenic flaviviral antigen can be from different flavivirus species, strains or isolates in any combination. In significant embodiments, the flaviviruses included can be two or more, three or more, four or more, five or more, or seven or more flaviviruses. Examples of such flaviviruses include, but are not limited to YFV, DEN-1, DEN-2, DEN-3, DEN-4, SLEV, JEV, WNV, Powassan virus or other flaviviruses. Combination vaccines can be formulated to confer immunity to flavivirus disease common to particular geographical regions. In a particular embodiment directed toward tropical and subtropical Asia, DEN-1, DEN-2, DEN-3, DEN-4, WN, and JE viruses can be selected. In a particular embodiment directed to Africa, DEN-1, DEN-2, DEN-3, DEN-4, WN and YF can be selected. In a particular embodiment directed to Latin America, DEN-1, DEN-2, DEN-3, DEN-4, Rocio and YF viruses can be selected.

The invention provides still further a method of immunizing a subject against infection by a flavivirus. The method involves administering to the subject an effective amount of a vaccinating composition that contains a nucleic acid molecule which

includes a transcriptional unit for an immunogenic flavivirus antigen. The transcriptional unit directs a cell within the body of the subject, after being taken up by the cell, to synthesize the immunogenic antigen. The composition additionally includes a pharmaceutically acceptable carrier. In significant embodiments of the method, the flavivirus may be YFV, DEN-1, DEN-2, DEN-3, DEN-4, SLEV, JEV, WNV, Powassan virus or other flavivirus. In yet other important aspects of the method, the antigen may be the prM/M protein, the E protein, or both the prM/M and the E proteins. When the antigen is both the prM/M and the E proteins, the cell within the body of the subject, after incorporating the nucleic acid within it, secretes subviral particles comprising the flaviviral prM/M and E antigens. Additionally, in significant embodiments of the method, the vaccinating composition is administered to the subject in a single dose, via a parenteral route. In yet a further aspect of the method, the nucleic acid is a DNA molecule. In yet additional embodiments of the method, the transcriptional unit further includes a control sequence disposed appropriately such that it operably controls the synthesis of the prM/M and E antigens and in a significant aspect of this embodiment, the control sequence is the cytomegalovirus immediate early promoter. Furthermore, the transcriptional unit may include a poly-A terminator.

These aspects and embodiments of the invention are the basis for its distinct attributes and advantages. Being a nucleic acid construct involving only portions of the flavivirus genome rather than the sequence encompassing the complete genome, the nucleic acid TU-containing vaccine is completely nonviable. It therefore poses no danger of infection by the flavivirus to those involved in its manufacture or to subjects receiving the vaccine. The nucleic acid vaccine is easy to prepare and easy to administer and is stable in storage prior to use. Unexpectedly it has been found that the nucleic acid vaccine of the invention is essentially 100% successful in conferring protective immunity in mammals after administering only a single dose. A further unexpected result is that the nucleic acid TU is able to engender immunity to a flavivirus in a female mammal which can be transmitted to its progeny through the milk. Without wishing to be limited by theory, the inventor believes that a possible mechanism for the success of the nucleic acid in conferring protective immunity is that a host cell harboring the nucleic acid, such as the cell of a subject to whom the vaccine

is administered, produces subviral particles containing the flaviviral prM/M and E antigens. These particles mimic the immunogenic attributes of native flavivirus virions.

The present invention also provides noninfectious antigenic polypeptides, antigenic polypeptide fragments and NRA comprising the prM/M and/or E proteins of flaviviruses, wherein the transmembrane signal sequence is derived from a first flavivirus and the M and/or E proteins are derived from a second flavivirus. Further, the prM/M protein can comprise amino acid sequences from both the first and the second flaviviruses. Further, the E protein can comprise amino acid sequences from both the first and second flaviviruses. "Chimeric" as used herein means any protein or nucleic acid comprising sequence from more than one flavivirus. As used herein, "non-virulent" means the antigen or vaccine of this invention is incapable of causing disease. More particularly, the recombinant protein antigens are free of contaminating genomic material from flaviviruses that is necessary for flavivirus infection, replication and pathogenesis.

The polypeptides of the present invention can comprise the amino acid sequences defined herein, or that are known in the art, of the prM, M and/or E proteins of selected flaviviruses. The nucleic acids of this invention can comprise nucleotide sequence that encodes the prM, M and/or E proteins of selected flaviviruses.

The antigens of the present invention can be unconjugated, or they can be conjugated to a carrier molecule that facilitates placement of the antigen on a solid phase. A carrier molecule is one to which antigens can be conjugated and which will not react with antibodies in human serum. An example of such a carrier is bovine serum albumin (BSA).

The antigens of the present invention can also be recombinant proteins obtained by expressing nucleic acids encoding the antigen in an expression system capable of producing the antigen.

The amino acid sequences of the present antigens can contain an immunoreactive portion of the prM, M and/or E antigen. These antigens may further be attached to sequences designed to provide for some additional property, such as to remove/add amino acids capable of disulfide bonding to increase the reactivity of an epitope by providing a more rigid secondary structure, to increase its bio-longevity or to alter its cytotoxicity or to prevent infection. In any case, the antigen must possess immunoreactivity and/or immunogenicity.

Brief Description of the Drawings

Figure 1 is a schematic representation of flaviviral polyprotein processing. The central horizontal region provides a schematic representation of the viral genome. The lines denote the 5' and 3' non-translated regions and the boxed regions represent the open reading frame for structural (left and top) and non-structural (right and bottom) proteins. Cleavage by host cell signalase occurs simultaneously with translation at the E protein C-terminus, separating structural and non-structural regions. A subtilase-like cellular enzyme, furin, may be responsible for prM cleavage. Potential transmembrane domains of viral polyprotein are indicated by shaded areas.

Figure 2 is a map of the JEV genome (top) and the DNA sequence of oligonucleotides used in a reverse transcriptase-polymerase chain reaction (RT-PCR) (center) to construct the transcription unit for the expression of prM-E protein coding regions (bottom). Potential transmembrane domains of viral polyprotein are indicated by shaded areas.

Figure 3 shows a schematic representation of the plasmid vectors, pCDNA3, pCBamp, and pCIBamp, and the relationship between them. These plasmids include the CMV (cytomegalovirus) promoter/enhancer element, BGHp(A) (bovine growth hormone polyadenylation signal and transcription termination sequence), ampicillin resistance gene and ColE1 origin of replication for selection and maintenance in *E. coli*. The f1 origin of replication for single-stranded rescue in *E. coli* cells, SV40 origin of replication (SV40 ORI), neomycin resistance coding region and SV40p(A) sequences

were deleted from pCDNA3 to generate pCBamp. An intron sequence was inserted in the NcoI-KpnI site of pCBamp to generate plasmid pCIBamp.

Figure 4 shows SDS-PAGE-immunoblot analyses of the sucrose gradient purified subviral particles from JE-4B COS-1 culture fluid (4B, right lane of each pair). The density gradient purified JE virion from JEV infected C6/36 cell culture was used as a positive control (JEV, left lane of each pair). JE HIAF (hyperimmune ascitic fluid); 4G2, anti-E monoclonal antibody; JM01, anti-M monoclonal antibody; NMAF (normal mouse ascitic fluid).

Figure 5 shows a profile of the E antigen in a rate zonal sucrose gradient analysis prepared from the PEG precipitate of JE-4B cell culture medium with or without Triton X-100 treatment.

Figure 6 shows signal peptide probability of the pCBJE1-14 (pCBJE) predicted by the SignalP-HMM program (A). The signal peptide probability is improved by altering the c-region sequence at -4 and -2 positions (C-4G and G-2S) (panel B, JE-LSS-M), by shortening the n-region (panel C, JE-SS-ORI), or by a combination of both modifications (panel D, JE-SS-M).

Figure 7 shows schematic representations of plasmid vectors pCBD2-14-16 (100% DEN-2 E), pCBD2-1J-4-3 (90% DEN-2 E : 10% JEV E), and pCB8D2-2J-2-9-1 (80% DEN-2 E : 20% JEV E). These plasmids include the human cytomegalovirus (CMV) early gene promoter; JE virus signal sequence; DEN-2 virus prM and E gene region (amino terminal 100%, 90%, or 80%, respectively); JE virus E gene region (none, 10% or 20%, respectively); and bovine growth hormone poly A signal (BGH).

Figure 8 shows a comparison of secreted and membrane-bound recombinant protein by western blot. (A) Analysis of secreted recombinant antigen following PEG-precipitation and ethanol extraction of culture fluid for DEN-2 plasmids pCB8D2-2J-2-9-1, pCB9D2-1J-4-3, pCBD2-14-16, and control plasmid pEGFP. Lane 1(V), purified DEN-2 virus stained by Gold Blot (Owl Separation Systems, Portsmouth, NH).

Reactivity of secreted, recombinant antigen from each plasmid with a, anti-envelope (E) specific Mab 1A6A-8; b, a mixture of MAB 1A6A-8, anti-capsid (C) specific Mab 1A2A-1, anti-serum specific for DEN-2 virus premembrane (prM) protein; and c, normal mouse ascites. (B) Analysis of recombinant plasmid-transformed cell hydrophobic membrane proteins. Lane 1 (V), purified DEN-2 virus stained by Gold Blot; lane 2 (V), reactivity of purified DEN-2 virus with a mixture of Mab 1A6A-8, Mab 1A2A-1, anti-serum specific for DEN-2 virus M protein, and anti-serum for DEN-2 virus prM protein. Reactivity of isolated hydrophobic membrane proteins from each plasmid-transformed cell line with a, Mab 1A6A-8; b, a mixture of Mab 1A6A-8, Mab 1A2A-1, anti-serum specific for DEN-2 virus M protein, and anti-serum for DEN-2 virus prM protein; and c, normal mouse ascites.

Detailed Description of the Invention

The invention encompasses nucleic acid transcriptional units which encode flaviviral antigenic proteins, such as the prM/M and E protein antigens. The nucleic acids function to express the prM/M and E protein antigens when the nucleic acid is taken up by an appropriate cell, especially when the cell is the cell of a subject. The invention also encompasses a vaccine whose active agent is the nucleic acid transcriptional unit (TU). The invention further encompasses cells containing a TU. The invention in addition encompasses a method of immunizing a subject against flaviviral infection by administering to the subject an effective amount of a vaccine containing the nucleic acid TU molecules.

The invention provides an isolated nucleic acid comprising a transcriptional unit encoding a signal sequence of a structural protein of a first flavivirus and an immunogenic flavivirus antigen of a second flavivirus, wherein the transcriptional unit directs the synthesis of the antigen. The invention further encompasses the use of the nucleic acid transcriptional unit (TU) to generate flaviviral antigens and the flaviviral antigens produced by the nucleic acid TU. The flaviviral antigens encompassed by the present invention include chimeric flavivirus antigens incorporating amino acid

sequence from a first flavivirus and at least one further flavivirus. The invention still further encompasses the use of the flaviviral antigens encoded by the TU of the invention to produce flavivirus-specific antibodies and to detect the presence of flavivirus-specific antibodies.

In one embodiment, the isolated nucleic acid of this invention can comprise a transcriptional unit encoding a Japanese encephalitis virus signal sequence.

In another embodiment, the transcriptional unit of this invention can encode an immunogenic flavivirus antigen which can be from one or more of the following flaviviruses: yellow fever virus, dengue serotype 1 virus, dengue serotype 2 virus, dengue serotype 3 virus, dengue serotype 4 virus, Japanese encephalitis virus, Powassan virus and West Nile virus.

In another embodiment, the transcriptional unit of this invention can encode an immunogenic chimeric flavivirus antigen which can include sequence from more than one of the following flaviviruses: yellow fever virus, dengue serotype 1 virus, dengue serotype 2 virus, dengue serotype 3 virus, dengue serotype 4 virus, Japanese encephalitis virus, Powassan virus and West Nile virus.

In a particular embodiment, the nucleic acid of this invention can encode a signal sequence of Japanese encephalitis virus and an M protein and an E protein of West Nile virus, SLEV, YFV and/or Powassan virus. The nucleic acid can also encode an immunogenic antigen which can be an M protein of a flavivirus, an E protein of a flavivirus, both an M protein and an E protein of a flavivirus, a portion of an M protein of a flavivirus, a portion of an E protein of a flavivirus and/or both a portion of an M protein of a flavivirus and a portion of an E protein of a flavivirus. In a preferred embodiment, the isolated nucleic acid encodes both the M protein and the E protein of the flavivirus. Further, the nucleic acid of the invention can be DNA and can comprise nucleotide sequence SEQ ID NO:15, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23 or SEQ ID NO: 42.

In another particular embodiment, the nucleic acid of this invention can encode a signal sequence of Japanese encephalitis virus, an M protein of a second virus and a chimeric E protein formed by from substitution of a portion of the nucleic acid encoding the second virus' E protein with nucleic acid encoding the corresponding portion of the JEV E protein. Alternatively, the portion of sequence corresponding to the deleted portion of the second virus' E protein can be substituted by other sequence selected from a third virus or it can be a non-viral sequence. The second protein can be West Nile virus, SLEV, YFV, Powassan and/or a serotype of Dengue virus . Chimeric E proteins can include those where the carboxy terminal portion can be from one flavivirus and the remainder of the chimeric E protein is from another flavivirus. The carboxy terminal portion can be, for example, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, or 75% of the chimeric E protein. The nucleic acid of the invention can be DNA and can comprise the protein-encoding sequence from nucleotide sequence SEQ ID NO:44 or SEQ ID NO:46. The nucleic acid of the invention can comprise nucleotide sequence SEQ ID NO:44 or SEQ ID NO:46.

The transcriptional unit of this invention can also comprise a control sequence disposed appropriately so that it operably controls the syntheais of the antigen. The control sequence can be, for example, the cytomegalovirus immediate early promoter. The nucleic acid of this invention can also comprise a Kozak consensus sequence located at a translational start site for a polypeptide comprising the antigen encoded by the transcriptional unit. The transcriptional unit of this invention can also comprise a poly-A terminator.

The present invention further provides a cell comprising the nucleic acid of this invention.

Also provided is a composition comprising a pharmaceutically acceptable carrier and nucleic acid or cell or antigen of this invention. The present invention additionally provides a method of immunizing a subject against infection by a flavivirus, comprising administering to the subject an effective amount of a composition of this invention. In a particular embodiment, the composition used to

immunize a subject directs the synthesis of both the M protein and the E protein of a flavivirus and a cell within the body of the subject, after incorporating the nucleic acid within it, secretes subviral particles comprising the M protein and the E protein. Alternatively, the composition can comprise an M protein and/or E protein of a flavivirus or subviral particles comprising the M protein and E protein. In the methods of this invention, the immunizing composition can be administered to the subject in a single dose and can be administered via a parenteral route.

This invention further provides the antigens produced from the isolated nucleic acids of this invention. As an example, the antigen from the second flavivirus encoded by the nucleotide sequence of TU can be the M protein which can be, for example, from West Nile virus. The antigen can also be protein from dengue virus, St. Louis encephalitis virus, Japanese encephalitis virus, Powassan virus and/or yellow fever virus. In a further embodiment, the antigen comprises a prM/M protein comprising the transmembrane signal sequence from a first flavivirus and further amino acid sequence comprising the remainder of the prM/M protein from a second flavivirus, which can be from SLEV, JEV, YFV, WNV and/or Powassan virus. The transmembrane signal sequence from a first flavivirus can be an improved or modified signal sequence wherein the signal sequence imparts desired characteristics such as a high signal sequence probability. Accomplishing these goals by design or selection can be with the use of machine-learning computer programs including, but not limited to, those using a hidden Markov model.

The antigen encoded by the nucleotide sequence of the TU can be West Nile virus antigen, dengue virus antigen, St. Louis encephalitis virus antigen, Japanese encephalitis virus antigen, Powassan virus antigen and/or yellow fever virus antigen.

The antigen encoded by the nucleotide sequence of the TU can also be the E protein, which can be the E protein from West Nile virus, dengue virus, St. Louis encephalitis virus, Japanese encephalitis virus, Powassan virus and/or yellow fever virus. The antigen encoded can also be a chimeric E protein comprising amino acid sequence selected from more than one flavivirus.

Additionally, the antigen encoded by the nucleotide sequence of the TU can be the M protein and the E protein, which can be from West Nile virus, dengue virus, St. Louis encephalitis virus, Japanese encephalitis virus, Powassan virus and/or yellow fever virus.

As used herein, "M protein" or "prM protein" or "prM/M protein" means a flavivirus M protein or flavivirus prM protein. Examples include, but are not limited to, prM proteins comprising amino acid sequence from one or more flavivirus prM proteins, M proteins comprising no additional amino acid sequence and proteins comprising additional amino acid sequences which are processed *in vitro* or *in vivo* to generate the mature M protein.

As used herein, "nucleic acid transcriptional unit" or "nucleic acid transcriptional unit molecule" means a nucleic acid encoding one or more specified proteins. The TU has biological activity such that, after having been introduced into a suitable cell, the nucleic acid induces the synthesis of one or more specified gene products encoded by the nucleic acid. The gene product(s) is(are) other biological macromolecules, such as proteins, not chemically related to the TU. The nucleic acid TU induces the cell to employ its cellular components to produce the specific gene product or products encoded by the nucleic acid of the TU. Although any nucleic acid may serve as a TU, in a preferred embodiment, the TU is the DNA of a plasmid or similar vector, wherein the plasmid or vector comprises coding sequences of marker genes or other sequence constructions that facilitate use of the TU for experimentation and biosynthesis.

As used herein, a "control sequence" is a regulatory nucleotide sequence incorporated within a TU which interacts with appropriate cellular components of the cell and leads to enhanced or activated biosynthesis of the gene products encoded by the TU. Thus a suitable control sequence is one with which the components of the cell have the capability to interact, resulting in synthesis of the gene product. When operably disposed in a nucleic acid with respect to a specified coding sequence, a

control sequence effectively controls expression of the specified nucleic acid to produce the gene product.

As used herein, a "promoter" is a nucleotide sequence in a TU which serves as a control sequence.

As used herein, a "Kozak sequence" or "Kozak consensus sequence" is a nucleotide sequence at the translational start site which optimizes translation of eukaryotic mRNAs (Kozak, *Mol. Cell. Biology* 9: 5134-5142 (1989)).

As used herein, a "terminator" is an extended nucleotide sequence which acts to induce polyadenylation at the 3' end of a mature mRNA. A terminator sequence is found after, or downstream from, a particular coding sequence.

As used herein, a "cell" is a prokaryotic or eukaryotic cell comprising a TU coding for one or more gene products, or into which such a TU has been introduced. Thus, a cell harbors a foreign or heterologous substance, the TU, which is not naturally or endogenously found in the cell as a component. A suitable cell is one which has the capability for the biosynthesis of the gene products as a consequence of the introduction of the TU. In particular, a suitable cell is one which responds to a control sequence and to a terminator sequence, if any, that may be included within the TU. In important embodiments of the present invention, the cell is a mammalian cell. In particularly important embodiments of this invention, the cell is a naturally occurring cell in the body of a human or nonhuman subject to whom (which) the TU has been administered as a component of a vaccine. Alternatively, in analytical, or diagnostic applications, including preparation of antigen for use as a vaccine or in immunodiagnostic assays, or for demonstrative purposes, the cell may be a human or nonhuman cell cultured *in vitro*.

As used herein, a "vaccine" or a "composition for vaccinating a subject" specific for a particular pathogen means a preparation, which, when administered to a subject, leads to an immunogenic response in a subject. As used herein, an "immunogenic"

response is one that confers upon the subject protective immunity against the pathogen. Without wishing to be bound by theory, it is believed that an immunogenic response may arise from the generation of neutralizing antibodies (i.e., a humoral immune response) or from cytotoxic cells of the immune system (i.e., a cellular immune response) or both. As used herein, an "immunogenic antigen" is an antigen which induces an immunogenic response when it is introduced into a subject, or when it is synthesized within the cells of a host or a subject. As used herein, an "effective amount" of a vaccine or vaccinating composition is an amount which, when administered to a subject, is sufficient to confer protective immunity upon the subject. Historically, a vaccine has been understood to contain as an active principle one or more specific molecular components or structures which comprise the pathogen, especially its surface. Such structures may include surface components such as proteins, complex carbohydrates, and/or complex lipids which commonly are found in pathogenic organisms.

As used herein, however, it is to be stressed that the terms "vaccine" or "composition for vaccinating a subject" extend the conventional meaning summarized in the preceding paragraph. As used herein, these terms also relate to the TU of the instant invention or to compositions containing the TU. The TU induces the biosynthesis of one or more specified gene products encoded by the TU within the cells of the subject, wherein the gene products are specified antigens of a pathogen. The biosynthetic antigens then serve as an immunogen. As already noted, the TU, and hence the vaccine, may be any nucleic acid that encodes the specified immunogenic antigens. In a preferred embodiment of this invention, the TU of the vaccine is DNA. The TU can include a plasmid or vector incorporating additional genes or particular sequences for the convenience of the skilled worker in the fields of molecular biology, cell biology and viral immunology (See Molecular Cloning: A Laboratory Manual, 2nd Ed., Sambrook, Fritsch and Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; and Current Protocols in Molecular Biology, Ausubel et al., John Wiley and Sons, New York 1987 (updated quarterly), which are incorporated herein by reference).

The TU molecules of the instant invention comprise nucleic acids, or derivatives of nucleic acids, having nucleotide sequences that encode specific gene products related to antigens of flaviviruses such as, but not limited to, WNV, JEV, dengue virus, yellow fever virus and SLEV. Although any nucleic acid may serve as a TU, in an important embodiment, the TU is DNA. Alternatively, the nucleic acids may be RNA molecules. They may also be any one of several derivatives of DNA or RNA having a backbone of phosphodiester bonds that have been chemically modified to increase the stability of the TU as a pharmaceutical agent. Modifications so envisioned include, but are not limited to, phosphorothioate derivatives or phosphonate derivatives. These and other examples of derivatives are well known to persons skilled in the field of nucleic acid chemistry.

The genome of JEV has been characterized and sequenced (Figures 1 and 2). The M structural protein is expressed as a portion of the polyprotein which includes a pre-M sequence (pr). This pr sequence, immediately amino terminal to the M protein sequence, prevents conformational problems in the processing of the polyprotein. In particular, the presence of the pr sequence is important in preventing misfolding of the E protein. Thus, the presence of prM allows for assembly of JEV particles. Once the virion or particle is formed, the pr sequence can be cleaved from the prM protein to yield mature virus particles containing M proteins, although cleavage of the prM protein to yield M protein is not necessary to produce infectious particles. The prM sequences from many different, related flaviviruses are cleaved to but a low extent, but the flaviviruses themselves are nonetheless, infectious. Examples of such related flaviviruses with similar genomic structures and functions include, but are not limited to WNV, YFV, dengue virus and SLEV.

In one embodiment, the TU encoding flaviviral M and E proteins in the instant invention is DNA. In accord with the discussion in the preceding paragraph, this DNA comprises a nucleotide sequence which encodes the M protein, comprising the pre-M sequence, and a nucleotide sequence encoding the E protein. In this way, the intended gene products are enabled to form subviral particles within the cell. The pre-M

sequence can then be cleaved in a fashion analogous to that which occurs with respect to replete virions.

In order to function effectively *in vivo* as a vaccine, it is advantageous to include within the TU a control sequence that has the effect of enhancing or promoting the transcription of the nucleotide sequences encoding the antigens. Use of such promoters is well known to those of skill in the fields of molecular biology, cell biology and viral immunology (See Molecular Cloning: A Laboratory Manual, 2nd Ed., Sambrook, Fritsch and Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; and Current Protocols in Molecular Biology, Ausubel et al., John Wiley and Sons, New York 1987 (updated quarterly)). When the TU is used as a vaccine in a mammalian host, the promoter to be employed is preferably one which operates effectively in mammalian cells. Such a promoter is disposed with respect to the coding sequences from which transcription is to be promoted, at a position at which it may operably promote such transcription. In a significant embodiment of the instant invention, this promoter is the cytomegalovirus early promoter. In addition, in a further preferred embodiment of the invention, the coding sequences are followed, in the TU nucleic acid, by a terminator sequence (Sambrook et al.). Particular embodiments of the invention relate to both prokaryotic and eukaryotic cells. Many promoter sequences are known that are useful in either prokaryotic or eukaryotic cells. (See Sambrook et al.)

The nucleic acids of the invention may further include DNA sequences known to those of skill in the art to act as immunostimulatory elements. Examples of such elements include, but are not limited to, certain CpG motifs in bacterial DNA (Sato et al., *Science* 273: 352-354 (1996); Klinman et al., *Vaccine* 17: 19-25 (1998)).

Preparation of the TU of the invention is readily accomplished by methods well known to workers of skill in the field of molecular biology. Procedures involved are set forth, for example, in Molecular Cloning: A Laboratory Manual, 2nd Ed., Sambrook, Fritsch and Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989 and Current Protocols in Molecular Biology, Ausubel et al., John Wiley and Sons, New York 1987 (updated quarterly). The flaviviral RNA molecule may be

isolated from a sample of live virus by methods widely known among virologists familiar with flaviviruses, for example, and with other groups of viruses as well. Methods used with JEV are summarized in Kuno et al. (*J. Virol.* 72: 73-83 (1998)). The RNA is used as a template for the synthesis of cDNA using reverse transcriptase. From the cDNA, a fragment containing the pre-M through E coding region (Figure 2) is obtained by digestion with restriction nucleases known to cleave the cDNA appropriately to provide such fragments. Examples of restriction digestion of JEV are provided in Nitayaphan et al. (1990) and Konishi et al. (1991). Incorporation of promoters, such as the cytomegalovirus promoter, sequences to promote efficient translation, such as the Kozak sequence, and of the polyadenylation signal, is likewise well known to skilled practitioners in molecular biology and recombinant DNA engineering (Kozak, *Mol. Cell. Biology* 9: 5134-5142 (1989); Azevedo et al., *Braz. J. Med. Biol. Res.* 32: 147-153 (1999)). When a nucleic acid compriaing a TU containing the desired coding sequences and control sequences is prepared, it may be obtained in larger quantities by methods that amplify nucleic acids. Such methods are widely known to workers skilled in molecular biology and recombinant DNA engineering. Examples of these methods include incorporation of the nucleic acid into a plasmid for replication by culturing in a cell such as a prokaryotic cell and harvesting the plasmid after completing the culture, as well as amplification of the nucleic acid by methods such as PCR and other amplification protocols, as are well known in the art. These examples are not intended to limit the ways in which the nucleic acid containing the TU may be obtained.

The TU-containing nucleic acid molecules of the instant invention may be introduced into appropriate cells in many ways well known to skilled workers in the fields of molecular biology and viral immunology. By way of example, these include, but are not limited to, incorporation into a plasmid or similar nucleic acid vector which is taken up by the cells, or encapsulation within vesicular lipid structures such as liposomes, especially liposomes comprising cationic lipids, or adsorption to particles that are incorporated into the cell by endocytosis.

In general, a cell of this invention is a prokaryotic or eukaryotic cell comprising a TU, or into which a TU has been introduced. The TU of the present invention induces the intracellular biosynthesis of the encoded prM/M and E antigens. A suitable cell is one which has the capability for the biosynthesis of the gene products as a consequence of the introduction of the nucleic acid. In particular embodiments of the invention, a suitable cell is one which responds to a control sequence and to a terminator sequence, if any, which may be included within the TU. In order to respond in this fashion, such a cell contains within it components which interact with a control sequence and with a terminator and act to carry out the respective promoting and terminating functions. When the cell is cultured *in vitro*, it may be a prokaryote, a single-cell eukaryote or a multicellular eukaryote cell. In particular embodiments of the present invention, the cell is a mammalian cell. In these cases, the synthesized prM/M and E protein gene products are available for use in analytical, or diagnostic applications, including preparation of antigen for use as a vaccine or in immunodiagnostic assays, or for demonstrative purposes.

In some circumstances, such as when the cell is a cultured mammalian cell, the prM/M and E antigens are secreted in the form of subviral particles. These are aggregates of prM/M and E proteins resembling live virus in surface ultrastructural morphology and immunogenic properties. Since the TU of the invention does not include the remainder of the flaviviral genome, however, there is no capsid incorporated, and most importantly, no infectious viral RNA.

In another important embodiment of this invention, the cell is a natural cellular component of the subject to whom the TU has been administered as a vaccine. The TU, when administered to the subject, is taken up by the cells of the subject. The subject's cells have the capability of responding to any promoter sequences, and terminator, if present. In any case, the TU induces the subject's cells to synthesize flaviviral prM/M and E gene products. Without wishing to be constrained by theoretical considerations, it is believed that the subject's cells produce subviral particles *in vivo* consisting of the prM/M and E antigens, just as has been found to occur with cultured mammalian cells *in vitro*. Such subviral particles, it is believed,

then serve as the *in vivo* immunogen, stimulating the immune system of the subject to generate immunological responses which confer protective immunity on the subject. Again without wishing to be limited by theory, the resulting protective immunity may arise via either humoral or cellular immunity, i.e., via either an MHC class II- or class I-restricted mechanism, respectively, or by both mechanisms.

According to the invention, subjects are immunized against infection by flaviviruses, such as JEV, YFV, dengue virus, SLEV, WNV or other flaviviruses by administering to them an effective amount of a TU comprising nucleic acid which encodes the prM and/or E antigens. The nucleic acid, after being incorporated into the cells of the subject, leads to the synthesis of the flaviviral prM/M and/or E antigens.

In order to administer the TU to the subject, it is incorporated into a composition which comprises a pharmaceutically acceptable carrier. The term "pharmaceutically acceptable" means a material that is not biologically or otherwise undesirable, i.e., the material may be administered to an subject along with the immunogenic material (i.e., recombinant flavivirus protein antigens or portions thereof) without causing any undesirable biological effects or interacting in a deleterious manner with any of the other components of the vaccine in which it is contained. Examples of pharmaceutically acceptable carriers, or components thereof, include water, physiological saline and common physiological buffers (for further examples, see Arnon, R. (Ed.) Synthetic Vaccines I: pp. 83-92, CRC Press, Inc., Boca Raton, Florida, 1987).

It is understood by those skilled in the art that the critical value in describing a vaccination dose is the total amount of immunogen needed to elicit a protective response in a host which is subject to infectious disease caused by virulent or wild-type flavivirus infection. The number and volume of doses used can be varied and are determined by the practitioner based on such parameters as, age, weight, gender, species, type of vaccine to be administered, mode of administration, overall condition of the subject, et cetera, as well as other important factors recognized by those of skill in the art.

The TU may be administered to a subject orally, parenterally (e.g., intravenously), by intramuscular injection, by intraperitoneal injection, transdermally, extracorporeally, intranasally, topically or the like. Delivery can also be directly to any area of the respiratory system (e.g., lungs) via intubation. The exact amount of the TU required will vary from subject to subject, depending on the species, age, weight and general condition of the subject, the immunogenicity of the vaccine used, the strain or species of flavivirus against which the subject is being immunized, the mode of administration and the like. Thus, it is not possible to specify an exact amount for every embodiment of the present invention. However, an appropriate amount can be determined by one of ordinary skill in the art using only routine experimentation given the teachings herein and what is available in the art.

Parenteral administration of the vaccine of the present invention, if used, is generally characterized by injection. Injectables can be prepared in conventional forms, either as liquid solutions or suspensions, solid forms suitable for solution or suspension in liquid prior to injection, or as emulsions. A more recently revised approach for parenteral administration involves use of a slow release or sustained release system such that a constant dosage is maintained. See, e.g., U.S. Patent No. 3,610,795, which is incorporated by reference herein.

For solid compositions, conventional nontoxic solid carriers include, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin, talc, cellulose, glucose, sucrose, magnesium carbonate, and the like. Liquid pharmaceutically administrable compositions can, for example, be prepared by dissolving, dispersing, etc. an active compound as described herein and optional pharmaceutical adjuvants in an excipient, such as, for example, water, saline, aqueous dextrose, glycerol, ethanol, and the like, to thereby form a solution or suspension. If desired, the pharmaceutical composition to be administered may also contain minor amounts of nontoxic auxiliary substances such as wetting or emulsifying agents, pH buffering agents and the like, for example, sodium acetate, sorbitan monolaurate, triethanolamine sodium acetate, triethanolamine oleate, etc. Actual methods of preparing such dosage forms are known, or will be apparent, to those skilled in this art;

for example, see Remington's Pharmaceutical Sciences (Martin, E.W. (ed.), latest edition, Mack Publishing Co., Easton, PA).

In one embodiment, the TU of this invention can be administered to the subject by the use of electrotransfer mediated *in vivo* gene delivery, wherein immediately following administration of the TU to the subject, transcutaneous electric pulses are applied to the subject, providing greater efficiency and reproducibility of *in vivo* nucleic acid transfer to tissue in the subject (Mir et al., *Proc. Nat. Acad. Sci USA* 96: 4262-4267 (1999)).

In the methods of the present invention which describe the immunization of a subject by administering a vaccine of this invention to a subject, the efficacy of the immunization can be monitored according the clinical protocols well known in the art for monitoring the immune status of a subject.

An effective amount of a vaccinating composition is readily determined by those of skill in the art to be an amount which, when administered to a subject, confers protective immunity upon the subject. In order to undertake such a determination, the skilled artisan can assess the ability to induce flaviviral prM/M- and E-specific antibodies and/or flaviviral prM/M- and E-specific cytotoxic T lymphocytes present in the blood of a subject to whom the vaccine has been administered. One can also determine the level of protective immunity conferred upon an experimental subject by challenge with live flavivirus corresponding to the antigenic composition used to immunize the experimental subject. Such challenge experiments are well known to those of skill in the art.

In general, in order to immunize a subject against infection by WNV, JEV, YFV, dengue virus, SLEV, or other flaviviruses according to the present invention, and recognizing that the TUs employed in such methods may have differing overall sizes, doses ranging from about 0.1 μg/kg body weight to about 50 μg/kg body weight can be used.

It has unexpectedly been found that a TU of the present invention which is a DNA confers protective immunity at a level of effectiveness approximating 100% after administration of only a single effective dose of the TU by i.m. injection or by electrotransfer. This is in contrast to many immunization methods carried out using conventional vaccines (as described above), which require one or more booster vaccinations and which may not confer protective immunity to an effectiveness near 100%.

It has further been found unexpectedly that protective immunity may be transmitted from a vaccinated female subject to the offspring of the subject. A significant proportion of neonatal mice was shown to be protected against viral challenge after the mothers were vaccinated using the TU DNA of the invention. Without wishing to be limited by theory, it is known that passive immunity may be conferred on neonatal mammals due to the presence in maternal milk of neutralizing antibodies specific for various pathogens. It is possible that the protective immunity against JEV found within the neonates was transmitted to them in this way.

In another embodiment of the invention, the TU encodes a signal sequence of a structural protein of a first flavivirus and an immunogenic flavivirus antigen of a second flavivirus. Thus, in one embodiment, for example, the signal sequence of structural protein of a first flavivirus is replaced by a signal sequence of structural protein of a second flavivirus, which results in proper folding of the nascent polypeptide, proper processing in a host, and/or proper folding of the processed protein.

In another embodiment of the invention, the TU may encode an immunogenic flavivirus antigen wherein the antigen comprises sequence from one or more than one flavivirus. The signal sequence can be an improved signal peptide. Improvement of signal sequences, or selection of more optimal signal sequences, can be accomplished by application of the principles and techniques taught in Example 18 and references cited therein, each of which are incorporated herein by reference for the express teachings in each related to the selection, identification and design of signal sequences

with desired properties and functions. Generally, these desired properties and functions will include a high signal sequence probability.

In another embodiment of the invention, more than one TU or a TU encoding an immunogenic flavivirus antigen from more than one flavivirus are included in a single composition. Thus, in one embodiment, for example, a TU can encode a nascent polypeptide or polypeptides that are processed into proteins from more than one flavivirus. Preferably, the processed proteins form subviral particles that elicit an immunological response against the proteins. The subviral particles can be formed from processed proteins derived from the sequence of the same flavivirus, a combination of flaviviruses, or chimeric flavivirus proteins. Combination vaccines, comprising more than one TU or a TU encoding an immunogenic flavivirus antigen from more than one flavivirus can be tailored for use in particular geographical regions by inclusion of proteins from flaviviruses endemic to the region or otherwise likely to be encountered. For example, a vaccine for tropical and subtropical Asia can include TU(s) that encode proteins from the four serotypes of DEN, WN and JE virus vaccines. Similarly useful vaccines for Africa and Latin America could include TU(s) that encode proteins from the four serotypes of DEN, WN and YF viruses and the four serotypes of DEN, Rocio and YF viruses, respectively.

In another embodiment, the TU encodes a signal sequence of a structural protein of a first flavivirus and an immunogenic chimeric flavivirus antigen that includes amino acid sequence from more than one flavivirus. The signal sequence can be a Japanese encephalitis virus signal sequence. The chimeric flavivirus antigen can include sequence from a Japanese encephalitis virus antigen. In certain embodiments, the chimeric antigen is an E protein. The carboxy terminal portion of the E protein can be E protein sequence from the Japanese encephalitis virus. The carboxy terminal portion can be, for example, 5, 10, 15, 20, 25, 30, 40, 50 or 75% of the chimeric E protein. In a preferred embodiment, the TU encodes a signal sequence of a structural protein of Japanese encephalitis virus, a prM protein of a Dengue virus and a chimeric E protein containing sequence from both the Japanese encephalitis virus and the Dengue virus. The chimeric protein can be an E protein wherein the carboxy terminal portion

comprises Japancse encephalitis virus sequence. Examples of TUs include nucleic acid sequences shown in SEQ ID NO:44 and SEQ ID NO:46 that can direct the synthesis of flavivirus antigens such as those shown in SEQ ID NO:45 and SEQ ID NO:47.

The present invention further provides immunogenic compositions comprising the polypeptides of this invention in a pharmaceutical acceptable carrier for use as a protein vaccine. Antigens produced from the transcriptional units of the present invention can be used to elicit effective immune responses in a subject. Antigens for this purpose can comprise flavivirus prM protein, flavivirus M protein, flavivirus E protein or any combination thereof, including immunogenic fragments of the proteins. A particularly preferred embodiment is the use of the NRA described herein. A further preferred embodiment is a chimeric protein comprising the signal sequence of one flavivirus and the structural protein(s) of one or more different flaviviruses. In a particularly preferred embodiment, the signal sequence of the antigen is the Japanese encephalitis virus signal sequence. In other preferred embodiments, the signal sequence is a an improved signal peptide. Improvement of signal sequences, or selection of more optimal signal sequences, can be accomplished by application of the principles and techniques taught in Example 18 and references cited therein each of which are incorporated herein by reference for the express teachings in each related to the selection, identification and design of signal sequences with desired properties and function. Generally, these desired properties and function will include a high signal sequence probability.

In other embodiments, the protein vaccine of this invention further comprises a suitable adjuvant. As used herein, an "adjuvant" is a potentiator or enhancer of the immune response. The term "suitable " is meant to include any substance which can be used in combination with the vaccine immunogen (i.e., flavivirus prM protein, flavivirus M protein, flavivirus E protein, or any combination thereof) to augment the immune response, without producing adverse reactions in the vaccinated subject. Effective amounts of a specific adjuvant may be readily determined so as to optimize the potentiation effect of the adjuvant on the immune response of a vaccinated subject. In a preferred embodiment, adjuvanting of the vaccines of this invention is a 2 – stage

process, utilizing first a 2% aluminum hydroxide solution and then a mineral oil. In specific embodiments, suitable adjuvants can be chosen from the following group: mineral, vegetable or fish oil with water emulsions, incomplete Freund's adjuvant, *E. coli* J5, dextran sulfate, iron oxide, sodium alginate, Bacto-Adjuvant, certain synthetic polymers such as Carbopol (BF Goodrich Company, Cleveland, Ohio), poly-amino acids and co-polymers of amino acids, saponin, carrageenan, REGRESSIN (Vetrepharm, Athens, GA), AVRIDINE (N, N-dioctadecyl-N',N'-bis(2-hydroxyethyl)-propanediamine), long chain polydispersed ß (1,4) linked mannan polymers interspersed with O-acetylated groups (e.g. ACEMANNAN), deproteinized highly purified cell wall extracts derived from non-pathogenic strain of Mycobacterium species (e.g. EQUIMUNE, Vetrepharm Research Inc., Athens GA), Mannite monooleate, paraffin oil and muramyl dipeptide.

In another aspect, this invention provides a method for immunizing subjects with immunogenic amounts of the protein vaccine of the invention to elicit an effective immune response in the subject. Immunization can be carried out orally, parenterally, intranasally, intratracheally, intramuscularly, intramammarily, subcutaneously, intravenously and/or intradermally. The vaccine containing the flavivirus prM protein, flavivirus M protein and/or the flavivirus E protein can be administered by injection, by inhalation, by ingestion, or by infusion. A single dose can be given and/or repeated doses of the vaccine preparations, i.e. "boosters," can be administered at periodic time intervals to enhance the initial immune response or after a long period of time since the last dose. The time interval between vaccinations can vary, depending on the age and condition of the subject.

The term " immunogenic amount" means an amount of an immunogen, or a portion thereof, which is sufficient to induce an immune response in a vaccinated subject and which protects the subject against disease caused by wild-type or virulent flavivirus infections upon exposure thereto or which has a therapeutic or commercially beneficial effect that lessens the effect of flavivirus infection on the vaccinated subject.

The invention further provides an antibody produced in response to immunization by the antigen of this invention. The antibodies of the present invention can include polyclonal and monoclonal antibodies which can be intact immunoglobulin molecules, chimeric immunoglobulin molecules, "humanized antibodies," or Fab or $F(ab')_2$ fragments. Such antibodies and antibody fragments can be produced by techniques well known in the art which include those described in Harlow and Lane (Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989) and Kohler et al. (*Nature* 256:495-97, 1975) and U.S. Patents 5,545,806, 5,569,825 and 5,625,126, incorporated herein by reference. The antibodies can be of any isotype IgG, IgA, IgD, IgE and IgM.

The present invention can also include single chain antibodies (ScFv), comprising linked V_H and V_L domains and which retain the conformation and specific binding activity of the native idiotype of the antibody. Such single chain antibodies are well known in the art and can be produced by standard methods. (see, e.g., Alvarez et al., *Hum. Gene Ther.* 8: 229-242 (1997)).

Antibodies can be produced against the antigens of this invention which are synthesized from nucleic acid sequences encoding immunogenic amino acid sequences of the prM, M and/or E antigens of one or more flaviviruses and the signal sequence of a different flavivirus (e.g., JEV). Immunogenic peptides synthesized from the use of these chimeric constructs can easily be identified by use of methods well known in the art for identifying immunogenic regions in an amino acid sequence and used to produce the antibodies of this invention.

Conditions whereby an antigen/antibody complex can form, as well as assays for the detection of the formation of an antigen/antibody complex and quantitation of the detected protein, are standard in the art. Such assays can include, but are not limited to, Western blotting, immunoprecipitation, immunofluorescence, immunocytochemistry, immunohistochemistry, fluorescence activated cell sorting (FACS), fluorescence *in situ* hybridization (FISH), immunomagnetic assays, ELISA, ELISPOT (Coligan et al., eds. 1995. Current Protocols in Immunology. Wiley, New

York.), agglutination assays, flocculation assays, cell panning, etc., as are well known to the artisan.

As used herein, the term "bind" means the well characterized binding of antibody to antigen as well as other nonrandom association with an antigen. "Specifically bind" as used herein describes an antibody or other ligand that does not cross react substantially with any antigen other than the one specified, which in this case, is an antigen of this invention.

The antibody or ligand of this invention can be bound to a substrate (e.g., beads, tubes, slides, plates, nitrocellulose sheets, etc.) or conjugated with a detectable moiety or both bound and conjugated. The detectable moieties contemplated for the present invention can include, but are not limited to, an immunofluorescent moiety (e.g., fluorescein, rhodamine), a radioactive moiety (e.g., ^{32}P, ^{125}I, ^{35}S), an enzyme moiety (e.g., horseradish peroxidase, alkaline phosphatase), a colloidal gold moiety and a biotin moiety. Such conjugation techniques are standard in the art (for example, Harlow and Lane, Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989); Yang et al., *Nature* 382: 319-324 (1996)).

The present invention further provides a method of detecting flavivirus antibody in a sample, comprising contacting the sample with the flavivirus antigen of the present invention, under conditions whereby an antigen/antibody complex can form; and detecting formation of the complex, thereby detecting flavivirus antibody in the sample.

The present invention further provides a method of detecting flavivirus antigen in a sample, comprising contacting the sample with an antibody of this invention under conditions whereby an antigen/antibody complex can form; and detecting formation of the complex, thereby detecting flavivirus antigen in the sample.

The method of detecting flavivirus antigen in a sample can be performed, for example, by contacting a fluid or tissue sample from a subject with an antibody of this invention and detecting binding of the antibody to the antigen. It is contemplated that

the antigen will be on an intact flavivirus virion, will be a flavivirus-encoded protein displayed on the surface of a flavivirus-infected cell expressing the antigen, or will be a fragment of the antigen. A fluid sample of this method can comprise any biological fluid which could contain the antigen or a cell containing the antigen, such as cerebrospinal fluid, blood, bile, plasma, serum, saliva and urine. Other possible examples of body fluids include sputum, mucus and the like.

The method of detecting flavivirus antibody in a sample can be performed, for example, by contacting a fluid or tissue sample from a subject with an antigen of this invention and detecting the binding of the antigen to the antibody. A fluid sample of this method can comprise any biological fluid which could contain the antibody, such as cerebrospinal fluid, blood, bile, plasma, serum, saliva and urine. Other possible examples of body fluids include sputum, mucus and the like.

Enzyme immunoassays such as immunofluorescence assays (IFA), enzyme linked immunosorbent assays (ELISA) and immunoblotting can be readily adapted to accomplish the detection of flavivirus antibodies according to the methods of this invention. An ELISA method effective for the detection of the antibodies can, for example, be as follows: (1) bind the antigen to a substrate; (2) contact the bound antigen with a fluid or tissue sample containing the antibody; (3) contact the above with a secondary antibody bound to a detectable moiety which is reactive with the bound antibody (e.g., horseradish (rábano rusticano) peroxidase enzyme or alkaline phosphatase enzyme); (4) contact the above with the substrate for the enzyme; (5) contact the above with a color reagent; and (6) observe/measure color change or development.

Another immunologic technique that can be useful in the detection of flavivirus antibodies uses monoclonal antibodies (MAbs) for detection of antibodies specifically reactive with flavivirus antigens in a competitive inhibition assay. Briefly, sample is contacted with an antigen of this invention which is bound to a substrate (e.g., an ELISA 96-well plate). Excess sample is thoroughly washed away. A labeled (e.g., enzyme-linked, fluorescent, radioactive, etc.) monoclonal antibody is then contacted with any previously formed antigen-antibody complexes and the amount of monoclonal

antibody binding is measured. The amount of inhibition of monoclonal antibody binding is measured relative to a control (no antibody), allowing for detection and measurement of antibody in the sample. The degree of monoclonal antibody inhibition can be a very specific assay for detecting a particular flavivirus variety or strain, when based on monoclonal antibody binding specificity for a particular variety or strain of flavivirus. MAbs can also be used for direct detection of flavivirus antigens in cells by, for example, immunofluorescence assay (IFA) according to standard methods.

As a further example, a micro-agglutination test can be used to detect the presence of flavivirus antibodies in a sample. Briefly, latex beads, red blood cells or other agglutinable particles are coated with the antigen of this invention and mixed with a sample, such that antibodies in the sample that are specifically reactive with the antigen crosslink with the antigen, causing agglutination. The agglutinated antigen-antibody complexes form a precipitate, visible with the naked eye or measurable by spectrophotometer. In a modification of the above test, antibodies of this invention can be bound to the agglutinable particles and antigen in the sample thereby detected.

The present invention further provides a method of diagnosing a flavivirus infection in a subject, comprising contacting a sample from the subject with the antigen of this invention under conditions whereby an antigen/antibody complex can form; and detecting antigen/antibody complex formation, thereby diagnosing a flavivirus infection in a subject.

The present invention further provides a method of diagnosing a flavivirus infection in a subject, comprising contacting a sample from the subject with the antibody of this invention under conditions whereby an antigen/antibody complex can form; and detecting antigen/antibody complex formation, thereby diagnosing a flavivirus infection in a subject.

In the diagnostic methods taught herein, the antigen of this invention can be bound to a substrate and contacted with a fluid sample such as blood, serum, urine or saliva. This sample can be taken directly from the patient or in a partially purified

form. In this manner, antibodies specific for the antigen (the primary antibody) will specifically react with the bound antigen. Thereafter, a secondary antibody bound to, or labeled with, a detectable moiety can be added to enhance the detection of the primary antibody. Generally, the secondary antibody or other ligand, which is reactive, either specifically with a different epitope of the antigen or nonspecifically with the ligand or reacted antibody, will be selected for its ability to react with multiple sites on the primary antibody. Thus, for example, several molecules of the secondary antibody can react with each primary antibody, making the primary antibody more detectable.

The detectable moiety allows for visual detection of a precipitate or a color change, visual detection by microscopy, or automated detection by spectrometry, radiometric measurement or the like. Examples of detectable moieties include fluorescein and rhodamine (for fluorescence microscopy), horseradish peroxidase (for either light or electron microscopy and biochemical detection), biotin-streptavidin (for light or electron microscopy) and alkaline phosphatase (for biochemical detection by color change).

Particular embodiments of the present invention are set forth in the examples which follow. These examples are not intended to limit the scope of the invention as disclosed in this specification.

Examples

General methods utilizing molecular biology and recombinant DNA techniques related to preparing and expressing the nucleic acid TU molecules of the invention are set forth in, for example, Current Protocols in Molecular Biology, Ausubel et al., John Wiley and Sons, New York 1987 (updated quarterly), and Molecular Cloning: A Laboratory Manual 2nd Ed., Sambrook, Fritsch and Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989.

Example 1. Preparation of recombinant plasmids containing the transcriptional unit encoding JEV prM and E antigens. Genomic RNA was extracted from 150 μL of JEV strain SA 14 virus seed grown from mouse brain using a QIAamp™ Viral RNA

Kit (Qiagen, Santa Clarita, CA). RNA, adsorbed on a silica membrane, was eluted in 80 μL of nuclease-free water, and used as a template for the amplification of JEV prM and E gene coding sequences. Primer sequences were obtained from the work of Nitayaphan et al. (*Virology* 177: 541-552 (1990)). A single cDNA fragment containing the genomic nucleotide region 389-2478 was amplified by the reverse transcriptase-polymerase chain reaction (RT-PCR). Restriction sites KpnI and XbaI, the consensus Kozak ribosomal binding sequence, and the translation initiation site were engineered at the 5' terminus of the cDNA by amplimer 14DV389 (nucleotide sequence, SEQ ID NO:1; amino acid sequence, SEQ ID NO:2). An in-frame translation termination codon, followed by a NotI restriction site, was introduced at the 3' terminus of the cDNA by amplimer c14DV2453 (SEQ ID NO:3) (Figure 2). One-tube RT-PCR was performed using a Titan RT-PCR Kit (Boehringer Mannheim, Indianapolis, IN). 10 μL of viral RNA was mixed with 1 μL each of 14DV389 (50 μM) and c14DV2453 (50 μM) and 18 μL of nuclease-free water and the mixture was heated at 85°C for 5 min and then cooled to 4°C. 75 μL of reaction mix [20 μL 5x buffer, 2 μL of dNTP mixture (10 mM each), 5 μL of dithiothreitol (0.1 mM), 0.5 μL of RNasin™ (40 U/μL, Boehringer Mannheim), 2 μL of polymerase mixture, and 45.5 μL of nuclease-free water] was added and RT-PCR performed as follows: 1 cycle (50°C for 30 min, 94°C for 3 min, 50°C for 30 s, 68°C for 2.5 min), 9 cycles (94°C for 30 s, 50°C for 30 s, 68°C for 2.5 min), 20 cycles (94°C for 30 s, 50°C for 30 s, 68°C for 2.5 min in the first cycle, with an increment of 5 s per cycle thereafter), and a final extension at 68°C for 15 min. The RT-PCR product was purified by a QIAquick™ PCR Purification Kit (Qiagen) and eluted with 50 μL of 1 mM Tris-HCl, pH 7.5.

All vector constructions and analyses were carried out by using standard techniques (Sambrook et al., 1989). RT-PCR amplified cDNA, digested with KpnI and NotI nucleases, was inserted into the KpnI-NotI site of eukaryotic expression plasmid vector (pCDNA3, Invitrogen, Carlsbad, CA). Electroporation-competent *Escherichia coli* XL1-Blue cells (Stratagene, La Jolla, CA) were transformed by electroporation (Gene Pulser™, Bio-Rad, Hercules, CA) and plated onto LB agar plates containing 100 μg/mL carbenicillin (Sigma Chemical Co., St. Louis, MO). Clones were picked and inoculated into 3 mL LB broth containing 100 μg/mL carbenicillin. Plasmid DNA was

extracted from a 14 h culture using a QIAprep™ Spin Miniprep Kit (Qiagen). Automated DNA sequencing was performed as recommended (Applied Biosystems/Perkin Elmer, Foster City, CA). Both strands of the cDNA were sequenced and shown to be identical to the sequence for the original SA14 strain (Nitayaphan et al., 1990).

The fragment of plasmid pCDNA3 (Invitrogen, Carlsbad, CA) from nucleotide (nt) 1289 to nt 3455, containing f1 ori, SV40 ori, the neomycin resistance gene, and SV40 poly(A) elements was deleted by PvuII digestion and then ligated to generate the pCBamp plasmid. The vector pCIBamp, containing a chimeric intron insertion at the NcoI/KpnI site of the pCBamp was constructed by excising the intron sequence from pCI (Promega, Madison, WI) by digestion with NcoI and KpnI. The resulting 566-bp fragment was cloned into pCBamp by digesting with NcoI-KpnI to replace its 289-bp fragment. Figure 3 presents the relationships between the plasmids pCDA3, pCBamp, and pCIBamp.

Plasmids containing the transcriptional unit encoding JEV prM and E proteins were prepared from these plasmids. The cDNA fragment containing the JEV prM and E coding regions in the recombinant plasmid pCDJE2-7 (nucleotide sequence, SEQ ID NO:10; amino acid sequence, SEQ ID NO:11), derived from the pCDNA3 vector, was excised by digestion with NotI and KpnI or XbaI and cloned into the KpnI-NotI site of pCBamp, pCIBamp, pCEP4 (Invitrogen, Carlsbad, CA), or pREP4 (Invitrogen, Carlsbad, CA), or into the SpeI-NotI site of pRc/RSV (Invitrogen, Carlsbad, CA) expression vector to create pCBJE1-14 (nucleotide sequence, SEQ ID NO:17; amino acid sequence, SEQ ID NO:18), pCIBJES14, pCEJE, pREPE, and pRCJE, respectively. Both strands of the cDNA from clones of each plasmid were sequenced and recombinant clones with the correct nucleotide sequence were identified. Plasmid DNA for use in the in vitro transformation of mammalian cells or mouse immunization experiments was purified by anion exchange chromatography using an EndoFree™ Plasmid Maxi Kit (Qiagen).

Example 2. Evaluation of JEV prM and E proteins expressed by various recombinant plasmids using an indirect immunofluorescent antibody assay. The expression of JEV specific gene products by the various recombinant expression plasmids was evaluated in transiently transfected cell lines of COS-1, COS-7 and SV-T2 (ATCC, Rockville MD; 1650-CRL, 1651-CRL, and 163.1-CCL, respectively) by indirect immunofluorescent antibody assay (IFA). The SV-T2 cell line was excluded from further testing since a preliminary result showed only 1-2% of transformed SV-T2 cells were JEV antigen positive. For transformation, cells were grown to 75% confluence in 150 cm^2 culture flasks, trypsinized, and resuspended at 4°C in phosphate buffered saline (PBS) to a final cell count 5 x 10^6 per mL. 10 μg of plasmid DNA was electroporated into 300 μL of cell suspension using a BioRad Gene Pulse™ (Bio-Rad) set at 150 V, 960 μF and 100 Ω resistance. Five minutes after electroporation, cells were diluted with 25 mL fresh medium and seeded into a 75 cm^2 flask. 48 h after transformation the medium was removed from the cells, and the cells were trypsinized and resuspended in 5 mL PBS with 3% normal goat serum. 10 μL aliquots were spotted onto slides, air dried and fixed with acetone at -20°C for 20 min. IFA was performed with acetone-fixed plasmid-transformed cells using fluorescein isothiocyanate-conjugated goat anti-mouse immunoglobulin G (Sigma Chemical Co.) and JEV HIAF.

To determine the influence of various promoter and poly(A) elements on the JEV prM and E protein expression, COS-1 and COS-7 cell lines were transiently transformed by an equal amount of pCDJE2-7 (SEQ ID NO:10), pCEJE, pREJE, or pRCJE plasmid DNA. JEV antigens were expressed in both cell lines transformed by all four recombinant plasmids, thus confirming that the CMV or RSV (rous sarcoma virus) promoter and BGH or SV40 poly(A) elements were functionally active. However, the percentage of transformed cells and the level of JEV antigens expressed, as determined by the number of IFA positive cells and IFA intensity, respectively, differed greatly among the various plasmids (Table 1). A significantly high percentage of COS-1 cells transformed by pCDJE2-7 (SEQ ID NO:10), pCBJE1-14 (SEQ ID NO:17) and pCIBJES14 expressed the JEV antigens, and the level of the expressed proteins was compatible with JEV-infected cells. Cells transfected with pCEJE,

pREJE, or pRCJE vectors, on the other hand, had a low percentage of antigen-expressing cells, as well as a low intensity of fluorescence, indicating weak expression of the antigens.

In order to ascertain whether the enhanced expression of JEV proteins by pCDJE2-7 (SEQ ID NO:10) was influenced by the SV40-encoded eukaryotic origin of replication, the plasmid pCBJE1-14 (SEQ ID NO:17) was constructed so that a 2166-bp fragment, containing f1 ori, SV40 ori, the neomycin resistance gene and SV40 poly(a) elements from pCDJE2-7, was deleted. A chimeric intron was then inserted into pCBJE1-14 to generate pCIBJES14. The pCIBJES14 plasmid was used to determine if the expression of JEV proteins could be enhanced by the intron sequence. Following transformation, cells harboring both pCBJE1-14 and pCIBJES14 vectors expressed a level of JEV antigens similar to that observed with pCDJE2-7 (Table 1). This result indicates that expression of JEV prM and E antigens by recombinant vectors is influenced only by the transcriptional regulatory elements. Neither the eukaryotic origin of replication nor the intron sequence enhanced JEV antigen expression in the cells used. Vectors containing the CMV promoter and BGH poly(A) (Figure 3) were selected for further analysis.

Example 3. Selection of an in vitro transformed, stable cell line constitutively expressing JEV specific gene products. COS-1 cells were transformed with 10 μg of pCDJE2-7 DNA by electroporation as described in the previous example. After a 24 hr incubation in non-selective culture medium, cells were treated with neomycin (0.5 mg/mL, Sigma Chemical Co.). Neomycin-resistant colonies, which became visible after 2-3 weeks, were cloned by limited dilution in neomycin-containing medium. Expression of vector-encoded JEV gene products was initially screened by IFA using JEV HIAF. One JEV-IFA positive clone (JE-4B) and one negative clone (JE-5A) were selected for further analysis and maintained in medium containing 200 μg/mL neomycin.

Authenticity of the JEV E protein expressed by the JE-4B clone was demonstrated by epitope mapping by IFA using a panel of JEV E-specific murine

monoclonal antibodies (Mab) (Kimura-Kuroda et al., *J. Virol.* 45: 124-132 (1983); Kimura-Kuroda et al., *J. Gen. Virol.* 67: 2663-2672 (1986); Zhang et al., *J. Med. Virol.* 29: 133-138 (1989); and Roehrig et al., *Virol.* 128: 118-126 (1983)). JEV HIAF and normal mouse serum were used as positive and negative antibody controls, respectively. Four JEV-specific, six flavivirus-subgroup specific, and two flavivirus-group reactive Mabs reacted similarly with the 4B clone or JEV-infected COS-1 cells (Table 2).

Example 4. Antigenic properties and immunological detection of subviral particles secreted by the JE-4B COS-1 cell line.

a. Preparation of subviral particles. JE-4B COS-1 cells were grown and maintained in medium containing 200 μg/mL of neomycin. The cultured medium was routinely harvested and stored at 4°C, and replenished twice weekly, and the cells were split 1:5 every 7-10 days. Culture medium was clarified by centrifugation at 10,000 rpm for 30 min in a Sorvall F16/250 rotor at 4°C, and centrifuged further for 4 hr at 39,000 rpm in a Sorvall TH641 rotor at 4°C through a 5% sucrose cushion (w/w, prepared with 10 mM Tris HCl, pH 7.5, 100 mM NaCl (TN buffer)). The pellet containing subviral particles was resuspended in TN buffer and stored at 4°C. Alternatively, 7% or 10% PEG-8000 (w/v) was added to the clarified culture medium. The mixture was stirred at 4°C for at least 2 hr, and the precipitated particles were collected by centrifugation at 10,000 rpm for 30 min. The precipitate was resuspended in TN buffer and stored at 4°C. The subviral particles were purified from both pelleted and PEG-precipitated preparations by rate zonal centrifugation in a 5-25% continuous sucrose gradient in TN at 38,000 rpm at 4°C for 90 min. 1-mL fractions were collected from the top of the gradient, tested by antigen capture ELISA (see below), and the positive fractions loaded onto a 25-50% sucrose gradient in TN. This was centrifuged overnight in an equilibrium density centrifugation at 35,000 rpm at 4°C. 0.9-mL fractions from the equilibrium gradients were collected from the bottom. They were tested by antigen-capture ELISA and assessed for hemagglutination (HA) activity at pH 6.6. An aliquot of 100 μL of each fraction was weighed precisely to determine its density. The ELISA-positive fractions were pooled and pelleted at 39,000 rpm at 4°C for 3-4 hr and the pellet resuspended in TN buffer. Antigen-capture ELISA and HA

titers were determined on the pelleted samples. JEV-infected COS-1 cell supernatant was also subjected to similar purification protocols as detailed above and used as a positive control for the gradient analysis. JE virions were also purified from infected C6/36 cells 5-6 days postinfection by sedimentation in a glycerol/tartrate equilibrium gradient.

b. Western blots of subviral particles. Gradient-purified samples of the subviral particles were mixed with electrophoresis sample buffer and run on 10 or 12.5% sodium dodecyl sulfate-containing polyacrylamide gels (SDS-PAGE) as described by Laemmli (*Nature* 277: 680-685 (1970)). Proteins were transferred to a nitrocellulose membrane and immunochemically detected with polyclonal JEV HIAF, flavivirus cross-reactive anti-E Mab 4G2 (Henchal et al., *Amer. J. Trop. Med. Hyg.* 31: 830-836 (1982)), or mouse anti-prM peptide hyperimmune serum (JM01). Figure 4 shows a comparison of the M and E proteins produced by JEV infected C6/36 and JE-4B COS-1 cells. Some nonspecific reactivity to E protein was observed in the normal mouse ascitic fluid and Jmo1 anti-peptide serum. Proteins identical in size to M and E were secreted in the subviral particles and could be detected by E-specific Mab 4G2 and prM-specific JM01 antiserum, respectively.

c. Density gradient detection of JEV subviral particles in culture medium. For ELISA, antigen-capture antibody (4G2) was diluted in 0.1 M sodium carbonate buffer, pH 9.6, and used to coat 96-well microtiter plates (Immulon II, Dynatech. Chantilly, VA) by overnight incubation at 4° C. After blocking with 3% normal goat serum in PBS, two-fold serially-diluted samples were added to the 4G2-coated plate and incubated 1.5 hours at 37°C. Captured antigen was detected by horseradish peroxidase-conjugated 6B6C-1 Mag, and incubated for 1 hour at 37°C. The enzyme activity on the solid phase was then detected with TMB (3,3',5,5'-tetramethylbenzidine)-ELISA (Life Technologies, Grand Island, NY).

Approximately 500 mL of cell culture medium from 15 X 150 cm^2 flasks of JE-4B cells was collected four days after cells were seeded. PEG-precipitated subviral particles were resuspended in 2 mL of TN buffer, pH 7.5; a 0.7 mL aliquot of this

resuspended pellet was loaded onto a 5-25% sucrose gradient. Triton X-100, which disrupts subviral particles, was added to another 0.7 mL aliquot to a final concentration of 0.1% and this was loaded onto a 5-25% sucrose gradient prepared in TN buffer containing 0.1% Triton X-100. A definite opaque band was observed approximately 2.5 cm from the top of the gradient containing Triton X-100, but not in the gradient without detergent. Fractions (1 mL) were collected from top to bottom for each gradient (Figure 5). Each collected fraction was analyzed by antigen capture ELISA. Antigen was detected in fractions 4-6, indicating relatively rapid sedimentation characteristic of subviral particles. Treatment of the PEG precipitate from JE-4B culture medium with Triton X-100 shifted the position of ELISA-reactive material to the top of the gradient. Thus treatment with Triton X-100 produces only slow-sedimenting molecules. A similar finding was reported by Konishi et al. (*Virol.* 188: 714-720 (1992)). These results show that rapidly sedimenting subviral particles containing prM/M and E could be disrupted by detergent treatment.

Hemagglutination (HA) activity was determined in the pH range from 6.1 to 7.0 by the method of Clarke and Casals (*Amer. J. Trop. Med. Hyg.* 7: 561-573 (1958)). The subviral particle secreted by JE-4B cells and the virion particle produced by JEV infected COS-1 cells had a similar HA profile with the optimum pH determined to be 6.6.

Example 5. Comparison of the immune response in mice vaccinated with pCDJE2-7 nucleic acid vaccine of the invention and commercial JEV vaccine. Groups of five 3-week-old female, ICR outbred mice were injected intramuscularly in the left and right quadriceps with 100 μg of pCDJE2-7 plasmid in 100 μL of dH_2O or were given doses of JE-VAX (manufactured by the Research Foundation for Microbial Disease of Osaka University and distributed by Connaught Laboratories, Swiftwater, PA.) subcutaneously that are one-fifth the dose given to humans. The plasmid pCDNA3/CAT (Invitrogen), which encodes and expresses an unrelated protein, was used as the negative vaccination control. Except for one group of pCDJE2-7-vaccinated mice, all animals were boosted 3 weeks later with an additional dose of plasmid or JE-VAX. Mice were bled from the retroorbital sinus at 3, 6, 9, 23, 40 and

60 weeks after inoculation. JEV antibody titers were determined by enzyme-linked immunosorbent assay (ELISA) against purified JEV or by plaque reduction neutralization tests (PRNT) (Roehrig et al., *Virol.* 171: 49-60 (1989); and Hunt and Calisher, *Amer. J. Trop. Med. Hyg.* 28: 740-749 (1979)).

The pCDJE2-7 nucleic acid vaccine and JE-VAX provided 100% seroconversion within three weeks after the first vaccination in all three groups of mice (Table 3). The JEV ELISA and PRNT antibody titers reached the highest level at week 6 and week 9, respectively, after immunization. Mice receiving 1 dose of DNA vaccine had similar antibody responses as those receiving 2 doses. Comparable ELISA antibody titers were maintained in DNA-vaccinated groups up to 60 weeks, after which the experiment was terminated. However, only one of four mice in the JE-VAX group was JEV antibody positive at 60 weeks post-inoculation. The pCDNA3/CAT control group did not have any measurable JEV antibody. These results demonstrate that a single dose of JEV-specific nucleic acid vaccine is more effective in maintaining JEV antibody in mice than the commercial, FDA-approved JE-VAX vaccine.

Example 6. Comparison of various nucleic acid vaccine constructs of the invention and commercial JEV vaccine for effectiveness of vaccination at different ages. A similar level of JEV protein was expressed by COS-1 cells transformed by either pCDJE2-7, pCBJE1-14, or pCIBJES14. JEV antibody induction by these nucleic acid constructs was compared to JE-VAX commercial vaccine at two different ages at vaccination. Three-day (mixed sex) or 3-week-old (female) ICR outbred mice, 10 per group, were vaccinated intramuscularly with 50 or 100 μg of plasmid DNA, or subcutaneously with doses of JE-VAX that are one-tenth or one-fifth the dose given to humans. Serum specimens were collected at 3 and 7 weeks after immunization and tested at a 1:1600 dilution by ELISA using purified JEV as an antigen. Results are shown in Table 4.

Plasmid pCBJE1-14 provided the highest extent of seroconversion, i.e., antibody titer greater than 1:1600, achieving 80-100% at both ages of vaccination. Administration of pCDJE2-7 or pCIBJES14 provided moderate seroconversion by 7

weeks when 3-day old mice were vaccinated (60% for each), but weaker seroconversion (40% and 10%, respectively) when measured 3 weeks after vaccination. When these plasmids were administered at the age of 3 weeks, however, seroconversions of 90% or 100% were attained at both 3 weeks and 7 weeks after vaccination. In contrast, the commercial vaccine, JE-VAX, conferred no seroconversion when administered at 3 days of age, and 100% when given at 3 weeks of age. Thus the nucleic acid TU's for JEV prM and E provided an extent of seroconversion better than a very high dose of the commercial vaccine, and unexpectedly high seroconversion in both young and more mature animals.

Example 7. Protective immunity conferred by the nucleic acid vaccine of the invention. Three-day old vaccinated groups from Example 6 were challenged 7 weeks after vaccination by intraperitoneal injection of 50,000 pfu/100 μL of the mouse-adapted JEV strain SA14 and observed for 3 weeks. 100% protection was achieved in groups that received various nucleic acid TU-containing vaccine constructs for up to 21 days (Table 5). In contrast, 60% of the JE-VAX-vaccinated mice, as well as 70% of the pCDNA3/CAT-vaccinated negative controls, did not survive virus challenge by 21 days. These results indicate that the nucleic acid TU's of the invention confer unexpectedly effective protection on vaccinated mice. This suggests the possibility of employing the nucleic acid vaccine of the invention as an early childhood vaccine for humans. In contrast, JE-VAX, the inactivated human vaccine currently used, does not appear to be effective in young animals.

Example 8. Passive protection of neonatal mice correlated with the maternal antibody titer. Female ICR mice at the age of 3 weeks were vaccinated with either one dose or two doses spaced two days apart of pCDJE2-7 plasmid DNA, at 100 μg/100 μL, or with two doses of JE-VAX that were one-fifth the dose given to humans. The negative control group received two doses of 100 μg/100 μL of pCDNA-3/CAT plasmid. Passive protection by maternal antibody was evaluated in pups resulting from matings of experimental females with non-immunized male mice that occurred nine weeks following the first vaccination or 6 weeks following the second vaccination. Pups were challenged between 3-15 days after birth by intraperitoneal administration of

5,000 pfu/100 μL of mouse-adapted SA14 virus and observed daily for 3 weeks (Table 6). The survival rates correlated with the maternal neutralizing antibody titers. 100% of pups nursed by mothers with a PRNT of 1:80 survived viral infection, whereas none of the pups from the control mother survived (Table 6). Partial protection of 45% and 75% was observed in older pups that were nursed by mothers with a PRNT titer of 1:20 and 1:40, respectively. The survival rates also correlated with the length of time that pups were nursed by the immune mother. As just indicated, 13-15 day old pups had high survival rates. None of the 3-4 day old pups, however, survived virus challenge when the mother had a PRNT titer of 1:20 or 1:40. Thus maternal antibody provides partial to complete protective immunity to the offspring. In addition, JEV antibody was detected by ELISA in the sera of 97% (29/30) of the post-challenge pups.

Mice were inoculated intramuscularly with 1 or 2, 100 μg doses of plasmid DNA, or subcutaneously with two, 1/5 human doses of JE-VAX vaccine. Sera were collected 9 weeks post-vaccination for PRNT testing prior to mating with non-immune male.

Example 9. Preparation of recombinant plasmids containing the transcriptional unit encoding WNV prM and E antigens. Genomic RNA was extracted from 150 μL of Vero cell culture medium infected with NY 99-6480 strain, an strain isolated from the outbreak in New York 1999, using the QIAamp ™ Viral RNA Kit (Qiagen, Santa Clarita, CA). Extracted RNA was eluted and suspended in 80 μl of nuclease-free water, and used as a template for the amplification of WNV prM and E gene coding sequences. Primer sequences were obtained from the work of Lanciotti et al. (*Science* 286: 2333-2337 (1999)). A cDNA fragment containing the genomic nucleotide region was amplified by the reverse transcriptase-polymerase chain reaction (RT-PCR). Restriction sites BamBI and KasI were engineered at the 5' terminus of the cDNA by using amplimer WN466 (nucleotide sequence, SEQ ID NO:12). An in-frame translation termination codon , followed by a NotI restriction site was introduced at the 3' terminus of the cDNA by using amplimer cWN2444 (SEQ ID NO:13). The RT-PCR product was purified by a QIAquick™ PCR Purification Kit (Qiagen).

The double-stranded amplicon produced by use of the two amplimers above (SEQ ID NO:12 and SEQ ID NO:13) was digested with KasI and NotI enzymes to generate a 998 bp (nt-1470 to 2468) fragment of DNA was inserted into the KasI and NotI sites of a pCBJESS vector to form an intermediate plasmid, pCBINT. The pCBJESS was derived from the pCBamp plasmid, that contained the cytomegalovirus early gene promoter and translational control element and an engineered JE signal sequence element (Chang et al., *J. Virol.* 74: 4244-4252 (2000)). The JE signal sequence element comprises the JE signal sequence (SEQ ID NO:14).

The cDNA amplicon was subsequently digested with BamBI and Kas I enzymes and the remaining 1003 bp fragment (nt-466 to 1470) was inserted in to the KasI site of pCBINT to form pCBWN (nucleic acid sequence, SEQ ID NO:15; amino acid sequence, SEQ ID NO:16). Automated DNA sequencing using an ABI priam 377 Sequencer (Applied Biosystems/Perkin Elmer, Foster City, CA) was used to confirm that the recombinant plasmid had a correct prM and E sequence as defined by Lanciotti et al. (*Science* 286: 2333-2337 (1999)).

Plasmid DNA for use in the in vitro transformation of mammalian cells or mouse immunization experiments was purified by anion exchange chromatography as described in Example 1.

Example 10. Immunochemical characterization and evaluation of WNV prM and E proteins expressed by pCBWN. WNV specific gene products encoded by the pCBWN plasmid were expressed in COS-1 cells. Cells were electroporated and transformed with pCBWN plasmid according to Chang et al. (*J. Virol.* 74: 4244-4252 (2000)). Electroporated cells were seeded onto 75 cm^2 culture flasks or a 12-well tissue culture dish containing one sterile coverslip/well. All flasks and 12-well plates were kept at 37°C, 5% CO_2 incubator. Forty hours following electroporation, coverslips containing adherent cells were removed from the wells, washed briefly with PBS, fixed with acetone for 2 minutes at room temperature, and allowed to air dry.

Protein expression was detected using indirect immunofluorescence antibody

assay (IFA), as described in Example 2. Flavivirus E-protein specific monoclonal antibody (Mab) 4G2, WNV mouse hyperimmune ascitic fluid (HIAF) and normal mouse serum (NMS) at 1:200 dilution in PBS were used as the primary antibody to detect protein expression (Henchal et al., *Am. J. Trop. Med. Hyg.* 31: 830-836 (1982)).

Tissue culture medium was harvested 40 and 80 hours following electroporation. Antigen-capture (Ag-capture) ELISA was used to detect secreted WN virus antigen in the culture medium of transiently transformed COS-1 cells. The Mab 4G2 and horseradish peroxidase-conjugated Mab 6B6C-1 were used to capture the WN virus antigens and detect captured antigen, respectively (Chang et al., *J. Virol.* 74: 4244-4452 (2000); Henchal et al., *Am. J. Trop. Med. Hyg.* 31: 830-836 (1983); Roehrig et al., *Virology* 128: 118-126 (1983)).

WN virus antigen in the medium was concentrated by precipitation with 10 % polyethylene glycol (PEG)-8000. The precipitant was resuspended in TNE buffer (50 mM Tris, 100 mM NaCl, 10 mM EDTA, pH 7.5), clarified by centrifugation, and stored at 4°C. Alternatively, the precipitant was resuspended in a lyophilization buffer (0.1 M TRIZMA and 0.4% bovine serum albumin in borate saline buffer, pH 9.0), lyophilized and stored at 4°C. Lyophilized preparations were used as antigen for the evaluation in MAC- and indirect IgG ELISAs.

WN virus-specific protein was detected by IFA on the transiently transformed COS-1 cells. E, prM and M proteins expressed in these cells were secreted into the culture medium. WN virus antigen concentrated by PEG precipitation was extracted with 7.0 % ethanol to remove residual PEG (Aizawa et al., *Appl. Enviro. Micro.* 39: 54-57 (1980)). Ethanol extracted antigens and gradient-purified WN virions were analyzed on a NuPAGE, 4-12% gradient Bis-Tris Gel in a Excel Plus Electrophoresis Apparatus (Invitrogen Corp., Carlsbad,CA) and followed by electroblotting onto nitrocellulose membranes using a Excel Plus Blot Unit (Invitrogen Corp.). WN virus-specific proteins produced by the transiently transformed COS-1 cells were detected by WN virus specific mouse HIAF or flavivirus E protein reactive Mab 4G2 in a Western blot analysis, using NMS as a negative serum control. The proteins displayed similar

reactivity and identical molecular weights to the corresponding gradient purified virion E, prM and M protein derived from WN virus infected suckling mouse brain (SMB).

In analysis of the NRA as an antigen for diagnostic ELISA, one vial of lyophilized NRA, representing antigen harvested from 40 ml of tissue culture fluid, was reconstituted in 1.0 ml of distilled water and compared with the reconstituted WN virus infected suckling mouse brain (SMB) antigen provided as lyophilized as β-propiolactone-inactivated sucrose-acetone extracts (Clarke et al., *Am. J. Trop. Med. Hyg.* 7: 561-573 (1958)). All recombinant proteins, prM, M and E , had a similar reactivity to that of the gradient-purified virion E, prM and M proteins.

Coded human specimens were tested concurrently with antigens in the same test at the developmental stage. The MAC- and IgG ELISA protocols employed were identical to the published methods (Johnson et al., *J. Clin. Microbiol.* 38: 1827-1831 (2000); Martin et al., *J. Clin. Microbiol.* 38: 1823-1826. (2000)). Human serum specimens were obtained from the serum bank in our facility, which consists of specimens sent to the DVBID for WN virus confirmation testing during the 1999 outbreak In these tests, a screening MAC- and IgG ELISA were performed on a 1:400 specimen dilution. Specimens yielding positive/negative (P/N) OD ratios between 2 and 3 were considered suspect positives. Suspect serum specimens were subject to confirmation as positives by both ELISA end-point titration and plaque-reduction neutralization test (PRNT). All specimens yielding P/N OD ratios greater than 3.0 were considered positives without further confirmatory testing.

An Ag-capture ELISA employing flavivirus-group reactive, anti-E Mab, 4G2 and 6B6C-1, was used to detect NRA secreted into culture fluid of pCBWN transformed COS-1 cells. The antigen could be detected in the medium one day following transformation; and the maximum ELISA titer (1:32-1:64) in the culture fluid without further concentration was observed between day two and day four. NRA was concentrated by PEG precipitation, resuspended in a lyophilization buffer, and lyophilized for preservation. For diagnostic test development, one vial of lyophilized NRA was reconstituted with 1.0 ml distilled water and titrated in the MAC- or indirect

IgG ELISA using WN virus positive and negative reference human sera (Johnson et al., *J. Clin. Microbiol.* 38: 1827-1831 (2000); Martin et al., *J. Clin. Microbiol.* 38: 1823-1826 (2000)). Dilutions 1:320 and 1:160 of the NRA were found to be the optimal concentrations for use in MAC- and IgG ELISA, respectively. These dilutions resulted in a P/N OD_{450} ratio of 4.19 and 4.54, respectively, for MAC- and IgG test. The WN virus SMB antigens produced by NY-6480 and Eg101 strains were used at 1:320 and 1:640 dilution for MAC-ELISA, and 1:120 and 1:320 for IgG ELISA, respectively. The negative control antigens, PEG precipitates of the culture medium of normal COS-1 cells and normal SMB antigen, were used at the same dilutions as for the respective NRA and SMB antigen. Human serum specimens, diluted at 1:400, were tested concurrently in triplicate with virus-specific and negative control antigens. For the positive test result to be valid, the OD_{450} for the test serum reacted with viral antigen (P) had to be at least two-fold greater than the corresponding optical density value of the same serum reacted with negative control antigen (N).

The reactivity of NRA and NY-06480, Eg101 and SLE virus SMBs were compared by the MAC- and IgG ELISAs using 21 coded human serum specimens. Of the 21 specimens, 19 had similar results on all three antigens (8 negatives and 11 suspect positives or positives). Eighteen specimens were also tested separately using SLE SMB antigen. Only three of 13 Eg-101-SMB positive specimens were positive in the SLE MAC-ELISA (Table 1). None of WN antigen negative specimens was positive by SLE MAC-ELISA. This result confirmed a previous observation that anti-WN virus IgM did not cross-react significantly with other flaviviruses (Tardei et al., *J. Clin. Microbiol.* 38: 2232-2239 (1940)) and was specific to diagnose acute WN virus infection regardless of whether NRA or SMB antigen was used in the test. All of the specimens were also tested concurrently by indirect IgG ELISA. Ten of 21 specimens were positive using any of the three antigens.

The two discrepant serum specimens (7 and 9) both from the same patient, collected on day-4 and 44 after onset of disease, respectively, were IgM-negative with NRA and SMB NY antigen and IgM-positive using Eg-101 SMB antigen in the initial test. To investigate these two discordant specimens further, six sequentially collected

specimens from this patient were retested by end-point MAC- and IgG ELISAs. A greater than 32-fold serial increase shown in the MAC-ELISA titer between day-3 and day-15 could be demonstrated with all antigens used. Cerebrospinal fluid collected on day-9 after onset of disease also confirmed that this patient indeed was infected by WN shortly prior to taking the sample. The cerebrospinal fluid had IgM P/N reading of 13.71 and 2.04 against Eg-101- and SLE-SMB antigens, respectively. Day-31 and day-44 specimens were negative (<1:400) by using NY-SMB antigen but positive by using NRA and Eg101-SMB. Compatible IgG titers were observed with all three antigens used in the test.

Example 11. Evaluation of the immune response in animals vaccinated with pCBWN. Groups of ten, three-wk-old female ICR mice were used in the study. Mice were injected intramuscularly (i.m.) with a single dose of pCBWN or a green fluorescent protein expressing plasmid (pEGFP) DNA (Clonetech, San Francisco, CA.). The pCBWN plasmid DNA was purified from XL-1 blue cells with EndoFree Plasmid Giga Kits (Qiagen) and resuspended in PBS, pH 7.5, at a concentration of 1.0 μg/μl. Mice that received 100 μg of pEGFP were used as unvaccinated controls. Mice were injected with the pCBWN plasmid at a dose of 100, 10, 1.0, or 0.1 μg in a volume of 100 μl. Groups that received 10, 1.0, or 0.1 μg of pCBWN were vaccinated by the electrotransfer mediated *in vivo* gene delivery protocol using the EMC-830 square wave electroporator (Genetronics Inc. San Diego, CA.). The electrotransfer protocol was based on the method of Mir et al., (*Proc. Natl. Acad. Sci. USA* 96: 4262-4267.(1999)). Immediately following DNA injection, transcutaneous electric pulses were applied by two stainless steel plate electrodes, placed 4.5-5.5 mm apart, at each side of the leg. Electrical contact with the leg skin was ensured by completely wetting the leg with PBS. Two sets of four pulses of 40 volts/mm of 25 msec duration with a 200 msec interval between pulses were applied. The polarity of the electrode was reversed between the set of pulses to enhance electrotransfer efficiency.

Mice were bled every 3 wks following injection. The WN virus specific antibody response was evaluated by Ag-capture ELISA and plaque reduction neutralization test (PRNT). Individual sera were tested by IgG-ELISA, and pooled sera

from 10 mice of each group were assayed by PRNT. All the mice vaccinated with pCBWN had IgG ELISA titers ranging from 1:640 to 1:1280 three wks after vaccination. The pooled sera collected at three and six wks had a Nt antibody titer of 1:80. None of the serum specimens from pEGFP control mice displayed any ELISA or Nt titer to WN virus.

To determine if the single i.m. vaccination of pCBWN could protect mice from WN virus infection, mice were challenged with NY-6480 virus either by intraperitoneal injection or by exposure to the bite of virus-infected *Culex* mosquitoes. Half of the mouse groups were challenged intraperitoneally (ip) at 6 wks post vaccination with 1,000 LD_{50} (1,025 PFU/100 μl) of NY99-6480 virus. The remaining mice were each exposed to the bites of three Culex tritaeniorhynchus mosquitoes that has been infected with NY99-6480 virus 7 days prior to the challenge experiment. Mosquitoes were allowed to feed on mice until they were fully engorged. Mice were observed twice daily for three wks after challenge.

It was evident that the presence of Nt antibodies correlated with protective immunity, since all mice immunized with WN virus DNA remained healthy after virus challenge while all control mice developed symptoms of CNS infection 4-6 days following virus challenge and died on an average of 6.9 and 7.4 days after intraperitoneal or infective mosquito challenge, respectively. In the vaccinated group, the pooled sera collected three wks after virus challenge (9-wk post immunization) had Nt antibody titers of 1:640 or 1:320. Pooled vaccinated mouse sera reacted only with E protein in the Western blot analysis.

Groups of ten mice were immunized with 10.0 to 0.1 μg of pCBWN per animal by use of electrotransfer. All groups that received pCBWN were completely protected from virus challenge. At 6 wks after immunization all groups of electrotransfer mice had Nt titer less than four-fold different than animals receiving 100 μg of pCBWN by conventional i.m. injection without electrotransfer. Both these results evidencing effective immunization suggest that the electrotransfer protocol enhances the immunogenicity and protective efficacy of the DNA vaccine of the invention (when

60

carried out as described in (Mir et al., *Proc. Natl. Acad. Sci. USA.* 96: 4262-4267.(1999))).

Mixed-bred mares and geldings of various ages used in this study were shown to be WN virus and SLE virus antibody-negative by ELISA and PRNT. Four horses were injected i.m. with a single dose (1,000 μg/1,000 μl in PBS, pH 7.5) of pCBWN plasmid. Serum specimens were collected every other day for 38 days prior to virus challenge, and the WN virus specific antibody response was evaluated by MAC- or IgG ELISA and PRNT.

Two days prior to virus challenge, 12 horses (4 vaccinated and 8 control) were relocated into a bio-safety level (BSL)-3 containment building at the Colorado State University. The eight unvaccinated control horses were the subset of a study that was designed to investigate WN virus induced pathogenesis in horses and the potential of horses to serve as amplifying hosts. Horses were each challenged by the bite of 14 or 15 *Aedes albopictus* mosquitoes that had been infected by NY99-6425 or BC787 virus 12 days prior to horse challenge. Mosquitoes were allowed to feed on horses for a period of 10 min. Horses were examined for signs of disease twice daily. Body temperature was recorded, and serum specimens collected twice daily from days 0 (day of infection) to 10, then once daily through day 14. Pulse and respiration were recorded daily after challenge. The collected serum samples were tested by plaque titration for detection of viremia, and by MAC- or IgG ELISA and PRNT for antibody response.

No systemic or local reaction was observed in any vaccinated horse. Individual horse sera were tested by PRNT. Vaccinated horses developed Nt antibody greater than or equal to 1:5 between days 14 and 31. End point titers for vaccinated horses, #5, #6, #7, and #8, on day-37 (two days prior to mosquito challenge) were 1:40, 1:5, 1:20, and 1:20, respectively. Horses vaccinated with the pCBWN plasmid remained healthy after virus challenge. None of them developed a detectible viremia or fever from days 1 to 14. All unvaccinated control horses became infected with WN virus after exposure to infected mosquito bites. Seven of the eight unvaccinated horses developed viremia that appeared during the first 6 days after virus challenge. Viremic horses developed Nt

antibody between day-7 and day-9 after virus challenge. The only horse from the entire study to display clinical signs of disease was horse #11, which became febrile and showed neurologic signs beginning 8 days after infection. This horse progressed to severe clinical disease within 24 hours and was euthanized on day 9. Four representing horses, #9, 10, 14 and 15, presenting viremia for 0, 2, 4, or 6 days, were selected and used as examples in this example. Virus titers ranged from $10^{1.0}$ PFU/ml of serum (in horse #10), the lowest level detectable in our assay, to $10^{2.4}$/ml (in horse #9). Horse #14 did not develop a detectible viremia during the test period. However, this horse was infected by the virus, as evidenced by Nt antibody detected after day 12.

Anamnestic Nt antibody response was not observed in vaccinated horses as evidenced by the gradual increase in Nt titer during the experiment. Pre-existing Nt antibody in the vaccinated horse prior to mosquito challenge could suppress initial virus infection and replication. Without virus replication, the challenge virus antigen provided by infected mosquitoes may not contain a sufficient antigen mass to stimulate anamnestic immune response in the vaccinated horse. All vaccinated horses were euthanized at 14 days after virus challenge. Gross pathological and histopathological lesions indicative of WN viral infection were not observed.

Example 12. Preparation of recombinant plasmids containing coding sequences for yellow fever virus (YFV) or St. Louis encephalitis virus (SLEV) prM and E proteins. A strategy similar to constructing the pCDJE2-7 recombinant plasmid was used to prepare YFV and SLEV recombinant plasmids. Genomic RNA was extracted from 150 μL of YFV strain TRI-788379 or SLE strain 78V-6507 virus seeds using Q1Aamp™ Viral RNA Kit (Qiagen, Santa Clarita, CA.). The viral RNA was used as a template for amplification of YFV or SLEV prM and E gene coding regions. Primers YFDV389 (nucleotide sequence, SEQ ID NO:4; amino acid sequence, SEQ ID NO:5), cYFDV2452 (SEQ ID NO:6), SLEDV410 (nucleotide sequence, SEQ ID NO:7; amino acid sequence, SEQ ID NO:8) and cSLEDV2449 (SEQ ID NO:9) were used to generate the corresponding recombinant nucleic acids as described above for the preparation of the JEV and WNV recombinant plasmids. RT-PCR amplified cDNA, digested with KpnI and NotI enzymes, was inserted into the KpnI-NotI site of a eukaryotic expression

plasmid vector, pCDNA3 (Invitrogen). Both strands of the cDNA were sequenced and verified for identity to sequences from YFV strain TRI-788379 or SLEV strain 78V-6507. Recombinant plasmids pCDYF2 and pCDSLE4-3, which contained the nucleotide sequences of the prM and E coding regions for YFV or SLEV, respectively, were purified using an EndoFree™ Plasmid Maxi Kit (Qiagen), and used for *in vitro* transformation or mouse immunization.

YFV or SLEV specific antigens were expressed in COS-1 cells transformed by pCDYF2 or pCDSLE4-3, respectively. The level of expressed proteins was similar to a YFV- or SLEV-infected COS-1 cell control. As in the JEV model, COS-1 cell lines transformed by vectors bearing genes for the viral antigens were obtained which constitutively express YFV or SLEV antigenic proteins. Epitope mapping by IFA using a panel of YFV or SLEV E-specific Mabs indicated that the authentic E protein was expressed by the pCDYF2- or pCDSLE4-3-transformed COS-1 cells. A preliminary study indicated that 100% of three week-old female, ICR mice seroconverted after intramuscular inoculation with a single dose of 100 μg/100 μL of pCDSLE4-3 plasmid in deionized water.

Example 13. Preparation of recombinant plasmids containing coding sequences for St. Louis encephalitis virus prM and E antigens with JEV signal sequence.

Genomic RNA was extracted from 150 μL of Vero cell culture medium infected with MSI-7 strain of St. Louis encephalitis virus using the QIAamp™ Viral RNA Kit (Qiagen, Santa Clarita, CA). Extracted RNA was eluted and suspended in 80 μl of nuclease-free water, and used as a template for the amplification of St. Louis encephalitis virus prM and E gene coding sequences. Primer sequences were obtained from the work of Trent et al. (*Virology* 156: 293-304 (1987)). A cDNA fragment containing the genomic nucleotide region was amplified by the reverse transcriptase-polymerase chain reaction (RT-PCR). Restriction site AfeI was engineered at the 5' terminus of the cDNA by using amplimer SLE463 (SEQ ID NO:30). An in-frame translation termination codon , followed by a NotI restriction site was introduced at the 3' terminus of the cDNA by using amplimer cSLE2447 (SEQ ID NO:31). The RT-PCR product was purified by a QIAquick™ PCR Purification Kit (Qiagen).

The double-stranded amplicon, produced by use of the two amplimers above (SEQ ID NO:30 and SEQ ID NO:31), was digested with AfeI and NotI enzymes to generate a 2004 fragment of DNA (463 to 2466nt), and inserted into the AfeI and NotI sites of a pCBJESS-M vector to form pCBSLE (nucleotide sequence, SEQ ID NO:21; amino acid sequence, SEQ ID NO:22). The pCBJESS-M was derived from the pCBamp plasmid, that contained the cytomegalovirus early gene promoter and translational control element and an engineered, modified JE signal sequence element (SEQ ID NO:27). The JE signal sequence element comprises the modified JE signal sequence at -4 (Cys to Gly) and -2 (Gly to Ser) position in the original pCBJESS plasmid.

Automated DNA sequencing using an ABI prism 377 Sequencer (Applied Biosystems/Perkin Elmer, Foster City, CA) was used to confirm that the recombinant plasmid had a correct prM and E sequence as defined by Trent et al. (*Virology* 156: 293-304 (1987)).

Example 14. Preparation of recombinant plasmids containing coding sequences for yellow fever virus (YFV) prM and E proteins with JEV signal sequence. Genomic RNA was extracted from 150 μL of Vero cell culture medium infected with 17D-213 strain of yellow fever virus using the QIAamp™ Viral RNA Kit (Qiagen, Santa Clarita, CA). Extracted RNA was eluted and suspended in 80 μl of nuclease-free water, and used as a template for the amplification of yellow fever virus prM and E gene coding sequences. Primer sequences were obtained from the work of dos Santos et al. (*Virus Research* 35: 35-41 (1995)). A cDNA fragment containing the genomic nucleotide region was amplified by the reverse transcriptase-polymerase chain reaction (RT-PCR). Restriction site AfeI was engineered at the 5' terminus of the cDNA by using amplimer YF482 (SEQ ID NO:28). An in-frame translation termination codon, followed by a NotI restriction site was introduced at the 3' terminus of the cDNA by using amplimer cYF2433 (SEQ ID NO:29). The RT-PCR product was purified by a QIAquick™ PCR Purification Kit (Qiagen).

The double-stranded amplicon, produced by use of the two amplimers above

(SEQ ID NO:28 and SEQ ID NO:29), was digested with AfeI and NotI enzymes to generate a 1971 fragment of DNA (482 to 2452nt), and inserted into the AfeI and NotI sites of a pCBJESS-M vector to form pCBYF (nucleotide sequence, SEQ ID NO:23; amino acid sequence, SEQ ID NO:24). The pCBJESS-M was derived from the pCBamp plasmid, that contained the cytomegalovirus early gene promoter and translational control element and an engineered JE signal sequence element (SEQ ID NO:27). The JE signal sequence element comprises the modified JE signal sequence at –4 (Cys to Gly) and –2 (Gly to Ser) position of JESS in the pCBJESS plasmid.

Automated DNA sequencing using an ABI prism 377 Sequencer (Applied Biosystems/Perkin Elmer, Foster City, CA) was used to confirm that the recombinant plasmid had a correct prM and E sequence as defined by dos Santos et al. (*Virus Research* 35: 35-41 (1995)).

Example 15. Preparation of recombinant plasmids containing coding sequences for Powassan virus prM and E antigens with JEV signal sequence. Genomic RNA was extracted from 150 μL of Vero cell culture medium infected with LB strain of Powassan virus using the QIAamp™ Viral RNA Kit (Qiagen, Santa Clarita, CA). Extracted RNA was eluted and suspended in 80 μl of nuclease-free water, and used as a template for the amplification of Powassan virus prM and E gene coding sequences. Primer sequences were obtained from the work of Mandl et al. (*Virology* 194: 173-184 (1993)). A cDNA fragment containing the genomic nucleotide region was amplified by the reverse transcriptase-polymerase chain reaction (RT-PCR). Restriction site AfeI was engineered at the 5' terminus of the cDNA by using amplimer POW454 (SEQ ID NO:25). An in-frame translation termination codon, followed by a NotI restriction site was introduced at the 3' terminus of the cDNA by using amplimer cPOW2417 (SEQ ID NO:26). The RT-PCR product was purified by a QIAquick™ PCR Purification Kit (Qiagen).

The double-stranded amplicon, produced by use of the two amplimers above (SEQ ID NO:25 and SEQ ID NO:26), was digested with AfeI and NotI enzymes to generate a 1983 bp fragment of DNA (454 to 2436nt), and inserted into the AfeI and

NotI sites of a pCBJESS-M vector to form pCBPOW (nucleotide sequence, SEQ ID NO:19; amino acid sequence, SEQ ID NO:20). The pCBJESS-M was derived from the pCBamp plasmid, that contained the cytomegalovirus early gene promoter and translational control element and an engineered JE signal sequence element (SEQ ID NO:27). The JE signal sequence element comprises the modified JE signal sequence at –4 (Cys to Gly) and –2 (Gly to Ser) position of JESS in the pCBJESS plasmid.

Automated DNA sequencing using an ABI prism 377 Sequencer (Applied Biosystems/Perkin Elmer, Foster City, CA) was used to confirm that the recombinant plasmid had a correct prM and E sequence as defined by Mandl et al. (*Virology* 194:173-184, (1993)).

Example 16. Preparation of plasmids containing coding sequences for dengue serotype 2 structural proteins. Procedures such as those carried out for other flaviviruses (see Examples 1, 9 and 12-15) are to be followed to prepare vectors including nucleic acid TU's for dengue serotype 2 antigens. According to the examples, the amplimers used for construction of the vectors may be chosen to engineer the normal dengue virus signal sequence or they may be chosen so as to engineer a signal sequence from another flavivirus, such as a modified Japanese encephalitis virus signal sequence.

A plasmid containing the dengue serotype 2 gene region from prM to E is to be constructed. The dengue serotype 2 prM and E genes (Deubel et al., *Virology* 155:365-377 (1986); Gruenberg et al., *J. Gen. Virol.* 69: 1301-1398 (1988); Hahn et al., *Virology* 162:167-180 (1988)) are to be ligated into a plasmid such as pCDNA3, and then excised and cloned into vectors such as pCBamp, pCEP4, pREP4, or pRc/RSV (supplied by Invitrogen, Carlsbad, CA) to enable expression. If necessary, a dengue serotype 2 virus-specific sequence encoded in a cDNA sequence may be amplified using a procedure such as the polymerase chain reaction (PCR). Alternatively, if the viral RNA is the source of the gene region, a DNA sequence may be amplified by a RT-PCR procedure. A DNA fragment including an initiation codon at the 5' end, and a termination codon at the 3' end is to be cloned into an expression vector at an

appropriate restriction nuclease-specific site, in such a way that the cytomegalovirus (CMV) immediate early (IE) promoter, an initiation codon, and a terminator, are operably linked to the dengue serotype 2 virus sequence.

Example 17. Vaccination of mice using a dengue serotype 2 DNA vaccine. The dengue serotype 2 nucleic TU vaccine encoding the gene region from prM to E prepared in Example 16 is to be suspended in a suitable pharmaceutical carrier, such as water for injection or buffered physiological saline, and injected intramuscularly into groups of weanling mice. Control groups receive a comparable plasmid preparation lacking the dengue serotype 2 specific genes. The generation of dengue serotype 2-specific antibodies, and/or of dengue serotype 2-specific immune system cytotoxic cells, is to be assessed at fixed intervals thereafter, for example at weekly intervals. At about two to four months after administration of the nucleic acid TU vaccine, mice are to be challenged with dengue serotype 2 virus. Levels of viremia are to be assessed at appropriate intervals thereafter, such as every second day. Passive protection by maternal antibody is to be assessed as indicated in Example 8.

Example 18. Design and construction of Improved Signal Peptides. Signal peptides can determine the translocation and orientation of inserted protein, hence, the topology of prM and E proteins. The most common feature of signal peptides of eukaryotes consists of an 8 to 12 stretch of hydrophobic amino acids called the h-region (von Heijne, "Signal sequences. The limits of variation" *J. Mol. Biol.* 184: 99-105 (1985)). The region between the initiator Met and the h-region, which is known as the n-region, usually has one to five amino acids, and normally carries positively charged amino acids. Between the h-region and the cleavage site is the c-region, which consists of three to seven polar but mostly uncharged amino acid residues. During viral polyprotein synthesis, modulation of the signalase cleavage site from a cryptic to cleavable conformation at the junction of C and prM proteins depends on prior removal of the C protein by the viral protease complex, NS2B/NS3 (Lobigs, "Flavivirus premembrane protein cleavage and spike heterodimer secretion require the function of the viral proteinase NS3," *Proc. Natl. Acad. Sci. U S A.* 90: 6218-6222 (1993)). Thus,

it is critical to consider the effectiveness of the viral signal sequence when prM and E proteins are to be expressed alone by an expression plasmid.

Signal peptide differences in various plasmid constructs can account, at least in part, for the difference in protein translocation, cleavage site presentation and correct topology, thus, prM and E secretion and VLP formation. The modulation or optimization of these attributes can be improved by selection or use of signal sequences with properties that impart the desired characteristics. This can be accomplished by use of machine-learning computer programs using, for example, a hidden Markov model (HMM) trained on eukaryotes (see Henrik Nielsen et al., "Prediction of signal peptides and signal anchors by a hidden Markov model," In Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology (ISMB 6), AAAI Press, Menlo Park, California, pp. 122–130 (1998); Nielsen et al., "Machine learning approaches to the prediction of signal peptides and other protein sorting signals," *Protein Engineering* 12: 3-9 (1999); Nielsen et al., "A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites," *Int. J. Neural Sys.* 8: 581-599 (1997); "From sequence to sorting: Prediction of signal peptides," Henrik Nielsen, Ph.D. thesis. Defended at Department of Biochemistry, Stockholm University, Sweden (May 25 1999); each of which is incorporated herein by reference, particularly for the teaching related to the optimization of signal sequences using computer-enabled algorithms).

An example of the type of program used is that found at http://www.cbs.dtu.dk/services/SignalP-2.0/ as of April 3, 2002. The HMM described in the referred to and incorporated references was applied to calculate the signal peptide probability of the prM signal peptide sequences in different plasmid constructs (Table 7). SignalP-HMM searches correctly predicted the signal peptidase cleavage sites in all constructs. However, considerable differences in cleavage probability (ranging between 0.164 and 1.000) and in signal peptide probability (ranging between 0.165 to 1.00) were observed (Table 7). This is not surprising, as the cleavage site and signal peptide probability are known to also be influenced by the positively charged amino acids in the n-region, the length of the hydrophobic amino acid in the h-region and

amino acid composition in the c-region in the constructs (Chang et al., "Flavivirus DNA vaccines: current status and potential," *Annals of NY Acad. Sci.* 951: 272-285 (2001); Sakaguchi et al., "Functions of Signal and Signal-Anchor Sequences are Determined by the Balance Between the Hydrophobic Segment and the N-Terminal Charge," *Proc. Natl. Acad. Sci. USA* 89: 16-19 (1992)).

Three JE virus plasmid constructs, each derived from different strain of JE virus, showed different vaccine potentials (Lin et al., "DNA immunization with Japanese encephalitis virus nonstructural protein NS1 elicits protective immunity in mice," *J. Virol.* 72: 191-200 (1998); Konishi et al., "Induction of protective immunity against Japanese encephalitis in mice by immunization with a plasmid encoding Japanese encephalitis virus premembrane and envelope genes," *J. Virol.* 72: 4925-4930 (1998); Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice," *J. Virol.* 74: 4244-4252 (2000)). The signal peptide sequences in these constructs are different in the length of n-region which may or may not contain charged amino acids (Table 7). The n-region containing positively charged amino acids forms a short loop in the cytoplasmic side that causes the h-region (transmembrane helix) to be inserted in a tail orientation, exposing the signalase cleavage site. In our study, secreted VLPs containing prM/M and E proteins could be purified from culture medium of the pCDJE2-7 transformed cell line, JE4B, or pCBJE1-14 transiently transformed COS-1 cells. The gradient-purified VLPs and virions have identical immunological and biochemical properties. Processing efficiency from prM to mature M protein, the hallmark of flavivirus morphogenesis, is also similar between VLPs and virion particles. Thus, prM and E proteins expressed by pCDJE2-7 and pCBJE1-14 can be expressed as type I transmembrane proteins in the orientation similar to that of virion prM and E (Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice," *J. Virol.* 74: 4244-4252 (2000)). In contrast, the prM protein of pcDNA3JEME could be expressed as a type II membrane protein with its transmembrane h-region inserted in a head orientation because of the absence of positively charged amino acids in its n-region (Konishi et al., "Induction of protective immunity against Japanese encephalitis

in mice by immunization with a plasmid encoding Japanese encephalitis virus premembrane and envelope genes," *J. Virol.* 72: 4925-4930 (1998)). Efficient protein synthesis in combination with the expressed protein having the correct topology, particularly of the expressed prM and E, can enhance VLP formation and secretion, and thus promote the immunogenicity of the DNA vaccine (Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice," *J. Virol.* 74: 4244-4252 (2000)).

The use of computer-based computations, as described above, has been applied to optimize the design of expression plasmid. In particular, the predictive power of the SignalP-HMM program was applied to design the WN virus expression plasmid (Table 2) (Davis et al., "West Nile virus recombinant DNA vaccine protects mouse and horse from virus challenge and expresses in vitro a noninfectious recombinant antigen that can be used in enzyme-linked immunosorbent assays," *J. Virol.* 75: 4040-4047 (2001)). The pCBWN plasmid consists of a short version of JE virus signal peptide followed by WN virus prM-E gene sequence. Vaccine potential of this construct was amply demonstrated, as a single i.m. injection of pCBWN DNA not only induced a protective immunity but also prevented WN virus infection in mice and horses.

As discussed earlier, and as demonstrated in Examples 13-15, virus-encoded signal sequence from the same virus as the antigen-coding regions is not necessarily the optimal signal peptide available. Further, non-modified signal sequence is not necessarily optimal. For example, the signal peptide encoded in the pCBJE1-14 plasmid can be improved, as measured by the signal sequence probability by shortening n-region, by altering c-region sequence, or by the combination of both modifications (Figure 6). By way of illustration, a shortened version of JE virus signal peptide has been used for the expression of WN virus prM and E genes as described herein and in papers incorporated herein by reference for the teaching (Davis et al., "West Nile virus recombinant DNA vaccine protects mouse and horse from virus challenge and expresses in vitro a noninfectious recombinant antigen that can be used in enzyme-linked immunosorbent assays," *J. Virol.* 75: 4040-4047 (2001)). Dose titration studies

by single i.m. inoculation indicated that the pCBWN was at least 2-4 fold more immunogenic than pCBJE1-14 in mice.

Example 19. Multivalent vaccines. Multivalent and/or combination vaccines designed to immunize against multiple flaviviruses can also be produced. In preparation of a multivalent vaccine, monovalent vaccine components are prepared that include elements related to pathogens of interest, such as YF, different serotypes of DEN, JE, WN, SLE and TBE (RSSE and CEE) viruses or any other combination of flaviviruses. Design and production of DNA constructs as described in the other examples and in the specification are carried out as described. Combinations of appropriate vaccines can be made to provide multivalent or combination vaccines protective against multiple pathogens. Preliminary data from our group has demonstrated that i.m. injection of the combined pCBJE1-14 and pCBWN DNA vaccines induced JE virus- and WN virus-specific Nt antibodies in mice (Table 8). Each monovalent component, even if constructed using identical transcriptional and translational regulators, should preferably be tested in an analogous model system to ensure its vaccine potential. A combination vaccine cocktail can then be formulated. These vaccine cocktails can be tailored specifically for particular geographic regions. For example, a vaccine cocktail for tropical and subtropical Asia should include four serotypes of DEN, WN and JE virus vaccines. Similarly useful vaccine cocktails for Africa and Latin America should include four serotypes of DEN, WN and YF virus vaccines and four serotypes of DEN, Rocio and YF virus vaccines, respectively.

Example 20. Preparation and testing of recombinant Dengue virus type 2 vaccines.

a. Summary of example. A series of plasmids that encode the premembrane (prM) and envelope (E) proteins of the dengue virus type 2 (DEN-2) were constructed. These plasmids included an authentic DEN-2 prM-E construct (pCBD2-14-6)(SEQ ID NO:42) encoding the protein described by SEQ ID NO:43, a 90% DEN-2 E-10% Japanese encephalitis (JE) virus E chimeric construct (pCB9D2-1J-4-3)(SEQ ID NO:44) that encodes the protein described by SEQ ID NO:45 and an 80% DEN-2 E-20% JE E chimeric construct (pCB8D2-2J-2-9-1)(SEQ ID NO:46) that encodes the

protein described by SEQ ID NO:47. Monoclonal antibody (MAb) reactivity indicated that all three plasmids expressed E protein epitopes that reacted with a panel of domain 1, 2 and 3 antibodies. However, only the pCB8D2-2J-2-9-1 construct (SEQ ID NO:46) secreted high levels of prM, M (matured prM) and E into the media of the plasmid-transformed COS-1 cells. The major portion of the prM and E protein expressed by COS-1 cells transformed with pCBD2-14-6 plasmid (SEQ ID NO:42) and by COS-1 cells transformed with pCB9D2-4-3 plasmid (SEQ ID NO:44) remained membrane-bound. Replacement of 20% of the sequence encoding the E protein of DEN-2 E with sequence encoding the corresponding JE E protein sequence had no effect on MAbs reactivity.

In testing, groups of mice received two intramuscular immunization of selected plasmids at 0 and 3-wks, and the immune response was evaluated by determining specific neutralizing and ELISA antibody. The plasmid expressing secreted prM and E, which can form subviral-particles (SVPs), was superior to other constructs in stimulating an antibody response. Ninety percent neutralization titers ranging from 1:40 to >1:1000 were observed from the 7 of 9 serum specimens of pCB8D2-2J-2-9-1 immunized mice.

b. Importance of DEN-2 virus and vaccines. Dengue (DEN) fever is an acute infection that occurs in subtropical and tropical areas. It is one of the most important flaviviral diseases of humans. As noted earlier, there are four distinct DEN serotypes (DEN-1, DEN-2, DEN-3 and DEN-4) of dengue virus. Infection with any of these is usually either asymptomatic or only causes a self-limited febrile illness known as dengue fever (DF). However, in a small percentage of cases, dengue virus infection results in a much more serious disease, life-threatening dengue hemorrhage fever or dengue shock syndrome (DHF/DSS). Thus, while there are approximately 100 million cases of the relatively mild DF world-wide annually which are of limited concern, there are also an estimated 500,000 hospitalized DHF/DSS cases reported annually. To protect against this disease, a safe and effective DEN vaccine effective against all four serotypes is required for administration to children and non-immune adults in the DEN endemic and epidemic regions.

Safe vaccines must minimize the potential risk of serious infection by virulent viruses. Such virulent viruses can arise by gene reversion or recombination of some types of vaccines derived from attenuated vaccine viruses. Such occurrences did arise in the poliovirus eradication campaign (Guillot et al., "Natural Genetic Exchanges between Vaccine and Wild Poliovirus Strains in Humans," *J. Virol.* 74: 8434-8443 (2000); Liu et al., "Molecular Evolution of a Type 1 Wild-Vaccine Poliovirus Recombinant during Widespread Circulation in China," *J. Virol.* 74: 11153-11161 (2000)). Further, genomic sequencing of an American strain of yellow fever virus, TRINID79A, demonstrates that there is extensive similarity between this strain and the attenuated yellow fever vaccine virus, FNV (Chang et al., "Nucleotide sequence variation of the envelope protein gene identifies two distinct genotypes of yellow fever virus," *J. Virol.* 69: 5773-5780 (1995); Pisano et al., "Complete nucleotide sequence and phylogeny of an American strain of yellow fever virus, TRINID79A," *Arch. Virol.* 144: 1837-1843 (1999)). While not conclusive in and of itself, the similarity strongly suggests that TRINID 79A is derived from the FNV vaccine virus.

The use of DNA-based vaccines is a novel and promising immunization approach for the development of flavivirus vaccines (as described herein, in Chang et al., "Flavivirus DNA vaccines: current status and potential," *Ann. NY Acad. Sci.* 951: 272-285 (2001), and in references cited therein). In this example, a number of DEN-2 vaccines were produced and the immune response in mice after i.m. immunization of DEN-2 constructs were correlated with the efficiency of prM/M and E secretion. One construct that led to the secretion of significant amounts of prM/M and E antigens was shown to be capable of stimulating high titers of neutralizing antibodies in plasmid-vaccinated mice.

c. Materials and methods.

i. Cell culture and virus strains. COS-1 cells (ATCC, Manassas, VA; 1650-CRL) were grown at 37°C with 5 % CO_2 in Dulbecco's modified Eagle minimal essential medium (DMEM, GIBCO, Grand Island, NY) supplemented with 10% heat-inactivated fetal bovine serum (FBS, Hyclone Laboratories, Inc., Logan, UT), 1mM sodium pyruvate, 1mM non-essential amino acids, 30 ml/liter 7.5% $NaHCO_3$, 100 units/ml of penicillin, and 100 µg/ml of streptomycin. Vero and C6/36 cells were grown under the same conditions used for COS-1 cells. DEN-2 virus, strain-16681,

was used for cDNA cloning , IgG ELISA and the plaque reduction neutralization test (PRNT). Virus was propagated in the C6/36 cell culture. Virus used for immunological or biochemical studies was purified by precipitation with 7% polyethylene glycol (PEG-8000; Fisher Scientific, Fair Lawn, NJ) followed by ultracentrifugation on 30% glycerol-45% potassium-tartrate gradients (Obijeaki et al., "Segmented genome and nucleocapsid of La Crosse virus," *J. Virol.* 20:664-675 (1976)).

ii. Plasmid construction. Genomic RNA was extracted from 150 µl of C6/36 cell culture medium infected with DEN-2 16681 strain using the QIAamp™ Viral RNA Kit (Qiagen, Santa Clarita, CA). Extracted RNA was resuspended in 80 µl of diethyl pyrocarbonate-treated water (DEPC, Sigma, ST. Louis, MO) and then used as a template in reverse transcriptase-PCR (RT-PCR) amplification of DEN-2 virus prM and E genes. Primer sequences (Table 9) were designed based on published sequences (Gadkari et al., "Critical evaluation of Kyasanur Forest disease virus neutralizing antibodies found in bats (a preliminary report)," *Indian J. Med. Res.* 64: 64-67 (1976); Kinney et al., "Construction of infectious cDNA clones for dengue 2 virus: strain 16681 and its attenuated vaccine derivative, strain PDK-53," *Virology* 230: 300-308 (1997)). The recognition and cleavage site for restriction enzyme KasI was incorporated at the 5' terminus of the cDNA amplicon. An in-frame termination codon, followed by a NotI restriction site, was introduced at the 3'terminus of the cDNA amplicon. The DEN-2 virus cDNA amplicon was digested with KasI and NotI enzymes, and was then inserted into the KasI and NotI sites of a pCBJESS vector to form the 100% DEN-2 E plasmid, pCBD2-14-6 (SEQ ID NO:42).

To construct the 90% and 80% DEN-2 E plasmids, the 100% DEN-2 plasmid, pCBD2-14-6 (SEQ ID NO:42), and the JE plasmid, pCBJE1-14 (SEQ ID NO:17), were used as the PCR templates to amplify DEN-2 and JE DNA sequence, respectively. Sets of primers used in amplification reactions to obtain the DEN-2 and JE gene fragments are listed in Table 9. T7 and SP6 priming sites are found in the pCBamp plasmid, derived from original pCDNA-3 plasmid (Invitrogen, Carlsbad, CA), and can be utilized as desired or as required. PCR amplified DNA fragments for the 90% DEN-2-

10% JE E protein gene were digested with BxtXI restriction endonuclease, ligated using T4 DNA ligase, digested with KasI and NotI enzyme, and inserted into the KasI and NotI sites of the pCBJESS vector to obtain the plasmid, pCB9D2-1J-4-3 (SEQ ID NO:44). PCR amplified DNA fragments for the 80% DEN-2-20% JE E gene were digested with BsmBI, ligated with T4 DNA ligase, digested with KasI and NotI enzyme, and inserted into the KasI and NotI sites of the pCBJESS vector to obtain the plasmid, pCB8D2-2J-2-9-1 (SEQ ID NO:46). Schematic representations of the three plasmid constructs are shown in Figure 7. The 90% DEN-2-10% JE E and the 80% DEN-2-20% JE E protein junction regions, respectively, are shown in Table 9.

Automated DNA sequencing was performed on an ABI Prism 377 Sequencer (Applied Biosystems/Perkin Elmer, Foster City, CA) according to the manufacture's recommended procedures. Recombinant plasmids with the correct prM and E sequences were identified using sequence analysis.

iii. Transient expression of DEN-2 recombinant antigen in COS-1 cells by electroporation. COS-1 cells were electroporated separately with each DEN-2 plasmid or green fluorescent protein expression plasmid control (pEGFP, Clonetech, San Francisco, CA) using the protocol described elsewhere in the examples and in Chang et al., ("A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice," *J. Virol.* 74: 4244-4252 (2000)). Electroporated cells were seeded onto 75 cm^2 culture flasks, and kept at 37°C and 5% CO_2. Six hours following electroporation the growth media were replaced with a maintenance media containing 2% fetal bovine serum. Tissue culture medium and cells were harvested separately 48 hours following electroporation for antigen characterization.

iv. Epitope mapping using DEN-2 E-specific monoclonal antibodies. Forty-eight hours following electroporation, adherent cells were trypsinized, resuspended in PBS containing 5% goat serum, spotted on a 12 well spot-slide and air dried. Cells adhered to the spot-slide were fixed with acetone for 10 minutes at –20°C and allowed to air dry. E-protein specific monoclonal antibodies (MAb) were used to detect protein

expression by the indirect immunofluorescence antibody assay (IFA), as described previously (Table 10; Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice," *J. Virol.* 74: 4244-4252 (2000)).

v. Characterization of the recombinant DEN-2 virus antigen. Tissue culture medium was harvested 48 hours following electroporation. Antigen-capture (Ag-capture) ELISA was used to detect secreted DEN-2 virus antigen in the culture medium of transiently transformed COS-1 cells. The MAb 4G2 and horseradish peroxidase-conjugated MAb 6B6C-1 were used to capture the DEN virus antigens and detect captured antigen, respectively (Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice," *J. Virol.* 74: 4244-4252 (2000); Hunt et al., "A recombinant particulate antigen of Japanese encephalitis virus produced in stably-transformed cells is an effective noninfectious antigen and subunit immunogen," *J. Virol. Methods.* 97: 133-149 (2001)).

Forty-eight hours following electroporation, transformed cells for each plasmid were trypsinized and resuspended in PBS as aliquots containing 5×10^6 cells. These cell samples were processed for membrane protein extraction using the Mem-PER mammalian membrane protein extraction reagent kit (Pierce, Rockford, IL) following the manufacturer's suggested protocol. Both hydrophobic and hydrophilic proteins are isolated. This procedure was developed for enrichment of integral membrane protein found in the hydrophobic phase. Both hydrophobic and hydrophilic fractions were analyzed by Ag-capture ELISA for DEN-2 recombinant antigen.

Recombinant antigen in the medium was concentrated by precipitation with 10% polyethylene glycol (PEG)-8000. The precipitant was resuspended in TNE buffer (50 mM Tris, 100 mM NaCl, 10 mM EDTA, pH 7.5) to $1/100^{th}$ of the original volume, clarified by centrifugation, and stored at 4° C. Recombinant antigen concentrated by PEG precipitation and resuspended in TNE buffer was extracted with 4.0 % ethanol to remove residual PEG (Hunt et al., "A recombinant particulate antigen of Japanese

encephalitis virus produced in stably-transformed cells is an effective noninfectious antigen and subunit immunogen," *J. Virol. Methods.* 97: 133-149 (2001)). Ethanol-extracted antigens, hydrophobic membrane proteins from the transformed cells and gradient-purified DEN-2 virions were analyzed on a NuPAGE, 4-12% Bis-Tris gradient gel in an Excel Plus Electrophoresis Apparatus™ (Invitrogen Corp., Carlsbad,CA), followed by electroblotting onto nitrocellulose membranes using a Excel Plus Blot Unit (Invitrogen Corp.)./DEN-2 virus-specific protein was detected by western blot using DEN-2 virus specific MAbs 1A6A-8 (E specific) and 1A2A-1 (capsid specific), as well as rabbit antiserum specific for DEN-2 prM and mouse serum specific for a peptide composed of amino acid 1-34 of the DEN-2 M protein, and normal mouse ascitic fluid was used as negative control (Murray et al., "Processing of the dengue virus type 2 proteins prM and C-prM," *J. Gen. Virol.* 74 (Pt 2): 175-182 (1993); Roehrig et al., "Monoclonal antibody mapping of the envelope glycoprotein of the dengue 2 virus, Jamaica," *Virology* 246: 317-328 (1998)).

vi. Mouse vaccination. Groups of ten, three-wk-old female ICR out-bred mice were used in the study. Mice were i.m. injected with pCBD2-14-6, pCB9D2-1J-4-3, pCB8D2-2J-2-9-1 or pEGFP on week-0 and week-3 at a dose of 100 μg in a volume of 100 μl per mouse. The plasmid DNA was purified from XL-1 blue cells with EndoFree Plasmid Giga Kits™ (Qiagen) and resuspended in PBS, pH 7.5, at a concentration of 1.0 μg/μl. Mice that received 100 μg of pEGFP were used as plasmid-vaccinated controls. Mice were bled every 3 wks following injection, and the DEN-2 virus specific antibody response was evaluated by use of indirect ELISA and PRNT.

vii. Serological tests. Pre- and post-vaccination serum specimens were tested for antibody binding ability to purified DEN-2 virion by ELISA, for neutralizing (Nt) antibody by PRNT, and for antibodies that recognize purified DEN-2 virus proteins by Western blotting. PRNT was performed with Vero cells, as previously described (Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice," *J. Virol.* 74: 4244-4252 (2000)), using DEN-2 (strain-16681) and JE (strain-Nakayama) virus. Endpoints

were determined at a 90% plaque-reduction level (Hunt et al., "A recombinant particulate antigen of Japanese encephalitis virus produced in stably-transformed cells is an effective noninfectious antigen and subunit immunogen," *J. Virol. Methods.* 97: 133-149 (2001)).

d. Results.

i. Transient expression of DEN-2 virus recombinant antigen. Expression of the prM and E genes of DEN-2 virus or a chimeric E gene from a combination of DEN-2 and JE virus sequences (80% DEN-20% JE or 90% DEN-10% JE) was accomplished by separate transformations of each of the three recombinant DEN-2 DNA plasmids into COS-1 cells. The basic plasmid design was based on results from previous studies with JE virus and WN virus recombinant plasmids in which plasmid-transformed cells expressed, and secreted authentic viral proteins into the cell culture fluid (Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice," *J. Virol.* 74: 4244-4252 (2000); Davis et al., "West Nile virus recombinant DNA vaccine protects mouse and horse from virus challenge and expresses in vitro a noninfectious recombinant antigen that can be used in enzyme-linked immunosorbent assays," *J. Virol.* 75: 4040-4047 (2001)). Transient expression of DEN-2 recombinant proteins was initially assessed by Ag-capture ELISA of cell culture supernatants and by IFA of acetone-fixed, transformed COS-1 cells (Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice," *J. Virol.* 74: 4244-4252 (2000)). The point of optimum antigen expression was determined to be 48 hours following electroporation.

ii. Epitope mapping of the E protein expressed by transiently-transformed COS-1 cells. The DEN-2 protein expressed by each of the recombinant plasmids was evaluated by IFA using a panel of murine MAbs with known reactivity to DEN-2 virus (Table 10; Henchal et al., "Epitopic analysis of antigenic determinants on the surface of dengue-2 virions using monoclonal antibodies," *Am. J. Trop. Med. Hyg.* 34: 162-169 (1985); Roehrig et al., "Monoclonal antibody mapping of the envelope glycoprotein of the dengue 2 virus, Jamaica," *Virology* 246: 317-328 (1998)). The MAb panel included

antibodies reactive with each of the three antigenic domains of the E protein of flaviviruses as well as prM and C proteins. (Mandl et al., "Antigenic structure of the flavivirus envelope protein E at the molecular level, using tick-borne encephalitis virus as a model," *J. Virol.* 63: 564-571 (1989); Rey et al., "The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution," *Nature* 375: 291-298 (1995)). The MAbs specific for flavivirus antigenic Domains 2 and 3 showed nearly identical qualitative reactivity with DEN-2 virus and each of the three plasmid-expressed proteins. One of the Domain 1-specific MAbs, 1B4C-2, also showed a similar reactivity pattern with all expressed proteins. However, two of the Domain 1-specific MAbs, 2B3A-1 and 9A4D-1, were much less reactive with E protein expressed by plasmids pCBD2-14-6 and pCB9D2-1J-4-3 as shown by endpoint titration (values in parentheses, Table 10). Comparison of the endpoint titers revealed the apparent poor expression of epitopes C3 and C4 in constructs containing 100% DEN-2 E and 90% DEN-2 E-10% JE E. MAb 2H2, specific for prM, had the same reactivity with antigen expressed by all three plasmids. Anti-C MAb 1A2A-1 reacted well with DEN-2 virus and had low level, non-specific reactivity with the plasmid-expressed viral proteins, which included prM and E, but not C.

iii. Comparison of secreted protein and membrane-bound protein produced by each of the three DEN-2 recombinant plasmids. Similar amounts of cell culture fluid were harvested from COS-1 cells 48 hours post-transformation for each recombinant DEN-2 plasmid. Secreted recombinant antigen found in the culture fluid was concentrated 100-fold by PEG precipitation, followed by ethanol extraction to remove PEG which interfered with subsequent analysis by polyacrylamide gel electrophoresis. The relative amount of secreted antigen expressed by each plasmid was determined by Ag-capture ELISA analysis of both PEG-precipitated and ethanol-extracted cell culture fluid preparations (Table 11). Secreted antigen was detected only from cells transfected with pCB8D2-2J-2-9-1 (SEQ ID NO:34), which contained 80% DEN-2 E and 20% JE E genes. The recombinant plasmids containing either 100% DEN-2 E or 90% DEN-2 E-10% JE E genes produced no ELISA-detectable antigen in the culture fluid, despite efforts to concentrate expressed protein.

Western blot analysis was also used to evaluate the production of secreted antigen by each of the DEN-2 recombinant plasmids. For comparison purposes, equivalent volumes of PEG-precipitated, ethanol-extracted cell culture supernatant were run on NuPAGE gradient gels, electroblotted to nitrocellulose, and analyzed using MAbs or polyclonal antisera capable of reacting with all DEN-2 structural proteins (Figure 8A). Western blot analysis showed greater sensitivity in detecting recombinant antigen than Ag-capture ELISA since DEN-2-specific proteins were detected in the culture fluid from two of the plasmids, pCB8D2-2J-2-9-1 and pCB9D2-1J-4-3 (SEQ ID NOS:46 and 44, respectively). Plasmid pCB8D2-2J-2-9-1 (SEQ ID NO:46) expressed the greatest amount of secreted antigen which was shown to be composed of E, prM, and M proteins. Relatively less secreted antigen was produced by pCB9D2-1J-4-3 (SEQ ID NO:44) and barely detectable levels were found for the pCBD2-14-6 (SEQ ID NO:42) preparation, which appeared to contain relatively less expressed E protein, especially if the non-specific reactivity of the E-specific MAb, 1A6A-8, on the control pEGFP was taken into consideration (Figure 8A, lanes a, b for 14-6 and GFP).

Since E, prM, and M are membrane-associated proteins throughout their intracellular synthesis, any assessment of the expression of these proteins by the three recombinant DEN-2 plasmids should include an evaluation of cell membrane preparations from plasmid-transformed cells. The Mem-PER Mammalian Membrane Protein Extraction Reagent kit (Pierce) was used to isolate the integral membrane proteins from equivalent numbers of cells transformed by each of the recombinant plasmids. Hydrophobic proteins were separated from hydrophilic proteins by phase partitioning. Preliminary analysis by Ag-capture ELISA indicated that the hydrophilic protein fraction was non-reactive; however, the hydrophobic protein fractions from COS-1 cells transformed with each of the recombinant DEN-2 plasmids had similar titers in ELISA tests (Table 11). These results indicated that recombinant antigen encoded by all three plasmids was expressed following transformation, but that the recombinant antigens expressed were not all secreted at the same level.

Confirmation of the Ag-capture results for the hydrophobic protein fractions was accomplished by western blot (Figure 8B). Equivalent volumes of hydrophobic

protein fractions from each of the plasmid-transformed cells were diluted according to the manufacturer's recommendations for SDS-polyacrylamide gel electrophoresis in order to reduce band and lane distortion. Immunoblotting with E-, prM-, C-, and M-specific MAbs or polyclonal antisera demonstrated that all three recombinant DEN-2 plasmids induced the production of similar amounts of recombinant antigen composed of E and prM. No M protein was detected, either because it was not processed from prM or because the levels were too low to be detected. Despite efforts to reduce band distortion, high levels of detergent in the hydrophobic protein samples apparently caused E and prM to run in a slightly aberrant manner (slower migration) compared to samples without such high concentrations of detergents (compare E and prM migration in Figure 8A and 8B).

iv. Comparison of the immune response in mice vaccinated with three different DEN-2 recombinant DNA plasmids. Three-week old ICR mice were immunized by i.m. injection with 100 µg of pCB8D2-2J-2-9-1 (SEQ ID NO:46), pCB9D2-1J-4-3 (SEQ ID NO:44), pCBD2-14-6 (SEQ ID NO:42), or pEGFP on weeks 0 and 3. Mice were bled 3, 6, and 9 weeks after the primary immunization. Individual and pooled sera were tested by indirect ELISA, using screening dilutions of 1:100 and 1:400 at 3 and 6 weeks post-vaccination and endpoint titrations at 9 weeks post-vaccination. Nine-week sera were also tested by PRNT with both DEN-2 and JE viruses. The ELISA results showed that after one immunization (3-week sera), all mice given pCB8D2-2J-2-9-1 had seroconverted, whereas only 50% of pCB9D2-1J-4-3 and 20% of pCBD2-14-6-vaccinated mice reacted with DEN-2 virus (Table 12). By 9 weeks post-vaccination, all mice vaccinated with either pCB8D2-2J-2-9-1 or pCBD2-1J-4-3 demonstrated anti-DEN-2 ELISA reactivity; however, the geometric mean titers differed significantly (titers of 1:20,000 versus 1:708, respectively). Only 40% of pCBD2-14-6-immunized mice had anti-DEN-2 ELISA titers greater than 1:100. A western blot of pooled 9-week sera from pCB8D2-2J-2-9-1-immunized mice on purified DEN-2 virus showed that the immunodominant response was to the E glycoprotein. Slight reactivity reactivity to prM and M were also detected.

More significantly, in terms of evaluating the vaccine potential of the three DEN-2 plasmids, the induction of virus-neutralizing antibody in 7of 9 mice immunized

with pCB8D2-2J-2-9-1 (SEQ ID NO:46) was observed based on a 90% plaque-reduction endpoint (Table 10). However, if a 50% neutralization endpoint is used, then all 9 of 9 sera have PRNT titers of ³ 1:40. Ninety percent neutralization titers ranged from 1:40 to >1:1000 for the 7 sera with neutralizing activity. None of the mice immunized with pCB9D2-1J-4-3 produced neutralizing antibody, and only 1 of 10 sera from pCBD2-14-6-vaccinated mice neutralized virus, but at a titer of only 1:8.

Since two of the recombinant plasmids, namely pCB9D2-IJ-4-3 (SEQ ID NO:44) and pCB8D2-2J-2-9-1 (SEQ ID NO:46) contained JE virus E-gene sequences, all sera were also evaluated for the presence of JE virus neutralizing activity. However, no such activity was detected at the 90% neutralization endpoint for mice in any of the immunization groups. Not surprisingly, mice immunized with the control plasmid pEGFP exhibited no reactivity to either DEN-2 or JE viruses.

e. Discussion. The same steps used earlier for the JE and WN vaccines was initially used to construct a recombinant DEN-2 plasmid, pCBD2-14-6 (SEQ ID NO:42), consisting of the authentic DEN-2 prM and E gene region. Antigenic mapping of DEN-2 proteins expressed by COS-1 cells transformed by this plasmid, using a panel of MAb by IFA, indicated the prM and E protein had a compatible fluorescent intensity and a similar MAb reactivity as virus infected cells (Table 10). However, these COS-1 cells transformed by the plasmid encoding an authentic DEN-2 prM and E region failed to secrete detectable DEN-2 antigen into the culture fluid (as measured by antigen-capture ELISA). In addition, vaccination using the plasmid encoding an authentic DEN-2 prM and E region failed to stimulate anti-DEN-2 virus neutralizing antibody in i.m. immunized-mice (Table 13). Interestingly, transformation of cells by pCBD2-14-6 resulted in a punctuated-globular fluorescent staining that suggested that the C-terminus of the E protein of DEN-2 may contribute to that protein's membrane retention signal. This IFA staining pattern was not observed in either the JE or WN construct-transformed cells (Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice," *J. Virol.* 74: 4244-4252 (2000); Davis et al., "West Nile virus recombinant DNA vaccine protects mouse and horse from virus challenge and

expresses in vitro a noninfectious recombinant antigen that can be used in enzyme-linked immunosorbent assays," *J. Virol.* 75: 4040-4047 (2001)). Therefore, in light of the observation made in accordance with the teachings of the present application, two additional plasmids, pCB9D2-1J-4-3 (SEQ ID NO:44) and pCB8D2-2J-2-9-1 (SEQ ID NO:46), in which appropriate manipulation of the DNA sequence was made for 10 % or 20 % of C-terminal E of DEN-2 to be replaced with the corresponding region of JE virus E protein, respectively. Relative effectiveness of the different constructs in stimulating detectable anti-DEN-2 ELISA antibody in vaccinated mice is shown in Table 13.

These results are consistent with the model that interactions between prM and E can influence the processes of particle assembly and secretion. Support for this model can be found in a study of tick-borne encephalitis virus that putatively suggests that interactions between prM and the ectodomain of E are involved in prM-mediated intracellular transport of prM-E, thus, secretion of virus-like particle (Allison et al., "Mapping of functional elements in the stem-anchor region of tick-borne encephalitis virus envelope protein E," *J. Virol.* 73: 5605-5612 (1999)).

In the present example, replacing a C-terminal portion of DEN-2 E protein with JE E protein, corresponding to TBE H1^{prM} to TM2, resulted in DEN-2 prM protein and chimeric E protein secretion. However, in contrast, replacement of TM1 and TM2, in TBE, caused only a minor improvement in antigen secretion. The major portion of the prM and E protein expressed by either COS-1 transformed pCBD2-14-6 and pCB9D2-4-3 plasmid remained membrane-bound (Table 13). These results indicated that an unidentified membrane retention sequence is located in the C-terminal stem region of DEN-2 E protein. Substitution of this C-terminal stem region with sequence from the JE virus removes or renders ineffective this retention sequence.

It has been asserted by others that the prM protein is essential for maintaining proper conformation and secretion of E protein during prM-E maturation (Aberle et al., "A DNA immunization model study with constructs expressing the tick-borne encephalitis virus envelope protein E in different physical forms," *J. Immunol.* 163:

6756-6761 (1999), Allison et al., "Synthesis and secretion of recombinant tick-borne encephalitis virus protein E in soluble and particulate form," *J. Virol.* 69:5816-5820 (1995)). Further, it has also been demonstrated that the ectodomain of the E protein interacts with prM. This interaction has been estimated to involve amino acid sequence within amino acid residues 200-327 of E in the Murray Valley encephalitis virus (Guirakhoo et al., "The Murray Valley encephalitis virus prM protein confers acid resistance to virus particles and alters the expression of epitopes within the R2 domain of E glycoprotein," *Virology* 191: 921-931 (1992)).

Proper prM and E interactions and retained integrity of the E protein's structure are likely maintained in the protein expressed by all three DEN-2 constructs, at least insofar as they are required for immunoreactivity. Furthermore, the replacement of the C-terminus 20 % E in the pCB8D2-2J-2-9-1 resulted in a protein that maintained 395 amino acids of authentic DEN-2 E. Any such modification is expected to have minimal influence on E and prM-E interactions and their influence on the antigenic nature of the chimeric E protein. As replacing the C-terminus region of DEN-2 E with JE stem-anchor sequence had no effect on MAbs reactivity (Table 10), retention of the DEN-2 sequence so replaced can only be optional in attaining a DEN-2 specific immunological response.

Previously, it has been shown that a plasmid construct encoding a secreted subviral particle of tick-borne encephalitis virus prM and E protein was superior, in terms of the extent and functionality of the antibody response and in terms of response to virus challenge, to other constructs that encoded C-terminally truncated soluble E-dimer that was secreted, full-length E that was not secreted, or a truncated E that was not efficiently secreted (Aberle et al., "A DNA immunization model study with constructs expressing the tick-borne encephalitis virus envelope protein E in different physical forms," *J. Immunol.* 163: 6756-6761 (1999)). However, here we have demonstrated that the vaccine potency of DEN-2 DNA is correlated with the secretion of prM/M and E (Table 13). The morphology and physical character of secreted prM and E were not demonstrated in this study. However, the prM and E, secreted by the pCB8D2-2J-2-9-1 construct, does likely form a viral-like particle. It is believed that

the presentation of multiple protective antigens on the particle surface improves vaccine potency of this construct.

Previous attempts for the development of DEN-2 virus DNA vaccine has resulted in varying degrees of success (Kochel et al., "Inoculation of plasmids expressing the dengue-2 envelope gene elicit neutralizing antibodies in mice," *Vaccine* 15: 547-552 (1997); Konishi et al., "A DNA vaccine expressing dengue type 2 virus premembrane and envelope genes induces neutralizing antibody and memory B cells in mice," *Vaccine* 18: 1133-1139 (2000)). To improve the level of effectiveness, differing strategies have been adopted. For example, coimmunization of the immuno-stimulatory CpG motif containing pUC19 plasmid, plasmid expressing murine GM-CSF in the vaccine regimen, or replacing C-terminus 43 amino acid of E with lysosome-associated membrane retention sequence improved the antibody response to the DEN-2 vaccine has been used (Porter et al., "Protective efficacy of a dengue 2 DNA vaccine in mice and the effect of CpG immuno-stimulatory motifs on antibody responses," *Arch. Virol.* 143: 997-1003 (1998); Raviprakash et al., "Synergistic Neutralizing Antibody Response to a Dengue Virus Type 2 DNA Vaccine by Incorporation of Lysosome-Associated Membrane Protein Sequences and Use of Plasmid Expressing GM-CSF," *Virology* 290: 74-82 (2001)). Unmethylated CpG motifs directly activates macrophages, natural killer cells and lymphocytes to secrete cytokines and chemokines, and supports the development of immune responses mediated by Th1 cytokines (Manders et al., "Immunology of DNA vaccines: CpG motifs and antigen presentation," *Inflamm. Res.* 49: 199-205 (2000)). However, inclusion of the CpG might bias the host's cytokine profile and thereby contribute to the both the development of Th-1-mediated organ-specific autoimmune disorders and interfere with immune homeostasis (Smith et al., "The regulation of DNA vaccines," *Curr. Opin. Biotech.* 12: 299-303 (2001)). There is also evidence in mice that excess levels of cytokine, although increasing the response of certain T-helper cells, can decrease or shutdown the response of other actors in the immune response, leading to generalized immunosuppression or chronic inflammation (Robertson et al., "Assuring the quality, safety, and efficacy of DNA vaccines," *Mol. Biotechnol.* 17: 143-149 (2001)). Correspondingly, the safety and efficacy of flavivirus DNA immunization

could be benefited by manipulation of an expression plasmid to enhance transcription and translation and targeting prM and E proteins for secretion that promote correct polyprotein processing and assembly (Chang et al., "Flavivirus DNA vaccines: current status and potential," *Ann. NY Acad. Sci.* 951: 272-285 (2001)). Future improvements could be focused on enhancing DNA uptake by antigen presenting cells or by muscle cells (Rodriguez et al., "Enhancing DNA immunization," *Virology* 268: 233-238 (2000)).

Table 1. Transient expression of JE prM and E proteins by various recombinant plasmids in two transferred cell lines.

Vector					Recombinant Plasmid	IFA intensity/percentage of antigen-positive cells*	
	Promotor	Intron	Poly (A)	ORI		COS-1	COS-7
pCDNA3	CMV	No	BGH	SV40	pCDJE2-7	3+/40	3+/35
pCBamp	CMV	No	BGH	No	pCBJE1-14	3+/45	nd
pC1Bamp	CMV	Yes	BGH	No	pC1BJES14	3+/39	nd
pCEP4	CMV	No	SV40	OriP	pCEJE	2+/4	2+/3
pREP4	RSV	No	SV40	OriP	pREJE	1+/3	1+/2
pRc/RSV	RSV	No	BGH	SV40	pRCJE	1+/3	1+/3
pCDNA3	CMV	No	BGH	SV40	pCDNA3/CAT	-	-

*Various cell lines were transformed with pCDNA3/CAT (negative control), pCDJE2-7, pCBJE1-14, pC1BJES14, pCEJEm pREJE, or pRCJE. Cells ere trypsinized 48 hours later and tested by an indirect immunofluorescent antibody assay (IFA) with JE virus-specific HIAF. Data are presented as the intensity (scale of 1+ to 4+) and the percentage of IFA positive cells. The pCDNA3/CAT transformed cells were used as the negative control.

Table 2. Characterization of proteins expressed by a pCDJE2-7 stably transformed clone (JE-4B) of COS-1 cells with JE virus-reactive antibodies.

Mab or antiserum	Biological Activity of Mab		Immunofluorescent intensity of cells	
	Specificity	Biological Function	JEV infected	4B
Mab:				
MC3	JEV Specific		2+	2+
2F2	JEV Specific	HI, N	4+	4+
112	JEV Specific		4+	4+
503	JEV Specific	N	4+	3+
109	Subgroup	HI	2+	1+
N.04	Subgroup	HI, N	3+	4+
201	Subgroup		1+	1+
203	Subgroup		4+	3+
204	Subgroup		2+	2+
301	Subgroup	HI	2+	2+
504	Flavivirus		4+	4+
6B6C-1	Flavivirus		2+	2+
3B4C-4	VEE		-	-
H1AF:				
Anti-JEV			4+	3+
Anti-WEE			-	-
PBS			-	-

Table 3. Persistence of the immune response in mice immunized with pCDJE2-7 or JE-VEX vaccine.

	ELISA Titer (log_{10})						$PRNT_{80\%}$ Titer		
	3 wks	6 wks	9 wks	23 wks	40 wks	60 wks*	3 wks	6 wks	9 wks
1x pCDJE2-7	2.6-3.2	3.8-5.0	3.8-4.4	>3.2	>3.2	2.4, 2.4, 3.8, 4.4	<20	20	40-160
2x pCDJE2-7	2.6-3.8	4.4	3.8-4.4	>3.2	>3.2	2.6, 3.8, 3.8	<20	20-40	40-160
2x JE-VAX	2.6-3.8	4.4-5.0	3.8-5.6	>3.2	>3.2	<2, <2, <2, 4.4	<20	20-40	20-160
2x pCDNA3/CAT	<2	<2	<2	ND	ND	<2	<20	<20	<20

Mice were inoculated with 1 or 2, 100 μg/dose plasmid DNA, or 1/5 human dose of JE-VAX vaccine. Sera were collected for testing prior to the second immunization.

* Individual serum titers.

Table 4. The age-dependent percent seropositive rate in mice following vaccination with various JEV vaccines.

	3-day old		3-week old	
	3 weeks PV	7 weeks PV	3 weeks PV	7 weeks PV
JE-VAX	0	0	100	100
pCDNA3/CAT	0	0	0	0
pCDJE2-7	40	60	90	90
pC1BJES14	10	60	80	100
pCBJE1-14	80	100	100	100

Table 5. Protection from JEV challenge in 8 week old mice following vaccination at 3 days old with various JEV vaccines.

Vaccine	Pre-challenge JEV seroconversion	Days post-challenge survival rate (%)				
		6	7	8	9	21
JE-VAX	0	100	100	60	40	40
pCDNA3/CAT	0	100	80	30	30	30
pCDJE2-7	60	100	100	100	100	100
pC1BJES14	60	100	100	100	100	100
pCBJE1-14	100	100	100	100	100	100

Table 6. Evaluation of the ability of maternal antibody from JEV-nucleic acid-vaccinated female mice to protect their pups from fatal JEV encephalitis.

Vaccinated mother		JEV challenged pups		
Vaccine	$PRNT_{50}$	Challenge age (days)	No. survival[1]	ELISA[2]
1 x pCDJE2-7	40	4	0/11	
2 x pCDJE2-7	80	4	12/12	12/12
2 x JE-VAX	20	3	0/16	
2 x pCDNA-3/CAT	<10	5	0/14	
1 x pCDJE2-7	20	15	5/11	5/5
2 x pCDJE2-7	40	14	8/12	7/8
2 x JE-VAX	80	13	5/5	5/5
2 x pCDNA-3/CAT	<10	14	0/14	

Mice were inoculated intramuscularly with 1 or 2, 100 μg dose of plasmid DNA, or subcutaneously with two, 1/5 human dose of JE-VAX vaccine. Sera were collected 9 weeks post-vaccination for PRNT testing prior to mating with non-immune male.

[1]: No Survivors/total for each litter.

[2]: Number of JEV ELISA-antibody-positive animals (titer ≥ 1:400)/No. of survivors; sera were collected for testing 12 weeks after challenge.

Table 7. Characteristic of the signal peptides and their vaccine of potentials among flavivirus DNA vaccine constructs

Plasmid	Signal peptide sequence preceding prM protein	Signal peptide probability[a] SP	AP	Csite	Immunization protocol/protection
pSLE1	YLETIRRRPSKRRGGTRSLLGLAALIGLASS/LQLLSTYQG (SEQ ID NO:32)	0.702	0.292	0.352	im x 2/Partial
pJME	MMLASLAVVIACAGA/MKLSNFQGK (SEQ ID NO:33)	0.998	0.000	0.778	im x 2/Partial
pCJEME	MNRGSIMWLASLAVVIACAGA/MKLSNFQGK (SEQ ID NO:34)	0.985	0.012	0.785	im x 2/100%
pCBJE1-14	MGRRQNRRGGNRGSIMWLASLAVVIACAGA/MKLSNFQGK (SEQ ID NO:35)	0.791	0.199	0.623	im x 1/100%
pcDNA3prM-E	MSKKRGGSRTSVLMVIFMLIGFAAA/LKLSNFQGK (SEQ ID NO:36)	0.721	0.277	0.622	im x 4 / Partial gg x 2-4/100%
pCBWN	MGKRSAGSIMWLASLAVVIACAGA/VTLSNFQGK (SEQ ID NO:37)	0.976	0.024	0.526	im x 1/100%
p1012D2ME	MNVLRGFRKRIGRMLNILNRRRRTAGMIIMLIPTVMA/FHLTTRNGE (SEQ ID NO:38)	0.165	0.778	0.164	id x 2/None
SV-PE	MVGLQKRGKRRSATDWMSWLLVITLLGMTLA/ATVRKERGD (SEQ ID NO:39)	0.943	0.056	0.899	im or gg x 2/100%
pWRG7077-RSSE	MGWLLVVVLLGVTLA/ATVRKERGD (SEQ ID NO:40)	1.000	0.000	0.912	gg x 2/100%
pWRG7077-CEE	MSWLLVITLLGMTIA/ATVRKERGD (SEQ ID NO:41)	0.999	0.000	0.821	gg x 2/100%

[a] The SignalP HMM program was applied to calculate signal peptide (SP), anchor peptide (AP) and signalase cleavage site (C site) probabilities. Single amino acid codes were used, and charged amino acids were highlighted by underline-bold letters. The signalase cleavage site separating SP and prM is indicated by "/". DNA vaccines were inoculated by intramuscular (im), intradermal (id), or gene gun (gg) method.

WO 02/081754
PCT/US02/10764

Table 8. Neutralizing antibody (Nt) responses in mice immunized with different doses of the combined WN and JE virus DNA vaccines.

	pCBWN + pCBJE1-14				pCB Control
Dose per plasmid (μg)	100 +100	40 + 40	20 +20	10 +10	100
Percentage of mice with Nt :					
WN virus/JE virus:	100 / 100	100 / 70	70 / 0	60 /0	0 / 0
Range of $PRNT_{90}$ titer:					
WN virus:	1:320 - 1:80	1:80 - 1:20	1:80 - <1:10	1:20 - <1:10	<1:10
JE virus:	1:40 - 1:10	1:10 - <1:10	<1:10	<1:10	<1:10

Groups of ten, three-week-old female ICR out-bred mice were i.m. injected with a single dose of Combined plasmid DNAs as indicated. The serum specimens collected 12-week after immunization were assayed by the plaque-reduction neutralization test (PRNT). The end-point titers against JE and WN virus were calculated based on the 90% percent plaque reduction using JE virus (strain SA-14) and West Nile virus (strain NY-6480), respectively.

Table 9. Oligonucleotides used to construct DEN-2 virus prM-E expression plasmids, and the junction region of chimeric DEN-2 and JE E indicated.

100% DEN-2 prM-E:	
D2KasI-438[a]	5' TGTGCA***GGCGCC***TTCCATTTAACCACACGTAACG (SEQ ID NO:48)
CD2NotI-2402	5' TCGA***GCGGCCGC***TCAACTAATTAGGCCTGCACCATGACTC (SEQ ID NO:49)
90% DEN-2 E & 10% JE E:	
T7	5' CTTATCGAAATTAATACGACTCACTATAGG (SEQ ID NO:50)
CD2BstXI-2244	5' ATAGATTGCT***CCA***AA***C***ACT***TGG***TGG (SEQ ID NO:51)
JE-2281	5' ACTCCATAGGAAAAGCCGTTCACC (SEQ ID NO:52)
CSP6	5' GCGAGCTCTAGCATTTAGGTGACACTATAG (SEQ ID NO:53)
	DEN-2↩ ↪JE
90-10 Junction:	**Leu His Gln Val Phe Gly Gly Ala Phe Arg Thr (SEQ ID NO:55)**
	CTC CAC CAA GTG TTT GGT GGT GCC TTC AGA ACA (SEQ ID NO:54)
80% DEN-2 E & 20% JE E:	
T7	5' CTTATCGAAATTAATACGACTCACTATAGG (SEQ ID NO:56)
CD2BsmBI-2097	5' GAATT***CGTCTC***ACTTCCTTTCTTAAACCAGTTGAGCTTC (SEQ ID NO:57)
JEBsmBI-2175	5' GGAATT***CGTCTC***GGAAGCACGCTGGGCAAGG (SEQ ID NO:58)
CSP6	5' GCGAGCTCTAGCATTTAGGTGACACTATAG 3' (SEQ ID NO:59)
	DEN-2↩ ↪JE
80-20 Junction:	**Asn Trp Lys Lys Gly Ser Thr Leu Gly Lys Ala (SEQ ID NO:61)**
	AAC TGG TTT AAG AAA GGA AGC ACG CTG GGC GCC (SEQ ID NO:60)

[a] Restriction enzyme sites encoded in oligonucleotides were highlighted by the bold, italic and underline.

Table 10. Characterization of DEN-2 E glycoprotein epitopes expressed by the recombinant DEN-2 plasmids as determined by the indirect fluorescent antibody assay (IFA).

Antibody			Controls[a]		Plasmid construct[a]		
MAb (Epitope)[b]	Antigenic Domain[c]	PRNT[d]	DEN-2 infected cells	Normal cells	pCBD2-14-6	pCB9D2-1I-4-3	pCB8D2-2I-2-9-1
4G2 (A1)	2	+/–	4+	-	4+	4+	4+
4E5 (A2)	2	Yes	3+	-	3-4+	3-4+	2-3+
1B7 (A5)	2	Yes	3-4+	-	4+	4+	2-3+
1B4C-2(C1)	1	No	3-4+ (8000)	-	2-3+ (4000)	2-3+ (4000)	2-3+ (8000)
2B3A-1 (C3)	1	No	3-4+ (≥3200)	-	3+ (100)	2+ (100)	2-3+ (≥3200)
9A4D-1 (C4)	1	No	3-4+	-	2-3+ (400)	1-3+ (400)	3+ (≥12800)
3H5 (B2)	3	Yes	4+	-	4+	4+	4+
10A4D-2 (B3)	3	Yes	2-3+	-	3-4+	3-4+	2-3+
1A1D-2 (B4)	3	Yes	4+	-	3-4+	4+	3-4+
9D12-6		Yes	2-4+		2-3+	2-3+	3-4+
2H2	prM	No	4+	-	4+	3-4+	3-4+
1A2A-1	Capsid	No	2-3+	-	1+	2+	1-2+

[a]IFA substrates were acetone-fixed COS-1 cells, either infected with DEN-2 16681, uninfected controls, or transformed with a DEN-2 recombinant plasmid.

[b]Monoclonal antibodies were used at a predetermined optimum dilution based on reactivity with DEN-2 16681 virus. For some MAbs, endpoint titers, shown in parentheses, are reported and for others, only qualitative values are reported based on a scale from 1+ to 4+, with 3-4+ considered positive, 2+ equivocal, and 1+ negative.

[c]Antigenic domains based on the E-glycoprotein of TBE virus (Mandl et al., "Antigenic structure of the flavivirus envelope protein E at the molecular level, using tick-borne encephalitis virus as a model," *J. Virol.* 63: 564-571 (1989); Rey et al., "The envelope glycoprotein from tick-borne encephalitis virus at 2 A resolution," *Nature* 375: 291-298 (1995)).

[d]Plaque-reduction neutralization activity at a 1:100 dilution of ascitic fluid, using a 90% plaque-reduction endpoint, except for 4G2 and 9D12-6, for which a 50% neutralization endpoint is reported (Henchal et al., "Epitopic analysis of antigenic determinants on the surface of dengue-2 virions using monoclonal antibodies," *Am. J. Trop. Med. Hyg.* 34: 162-169 (1985); Roehrig et al., "Monoclonal antibody mapping of the envelope glycoprotein of the dengue 2 virus, Jamaica," *Virology* 246: 317-328 (1998)).

Table 11. Detection of secreted and membrane-bound DEN-2 recombinant protein by antigen-capture ELISA.

Plasmid	Sample type	Endpoint ELISA titer
pCBD2-14-6	PEG-precipitated culture fluid[a]	<1:10
pCBD2-14-6	PEG-precipitated, ethanol-extracted culture fluid[b]	<1:20
pCBD2-14-6	Hydrophobic membrane protein preparation[c]	1:160
pCB9D2-1J-4-3	PEG-precipitated culture fluid[a]	<1:10
pCB9D2-1J-4-3	PEG-precipitated, ethanol-extracted culture fluid[b]	<1:20
pCB9D2-1J-4-3	Hydrophobic membrane protein preparation[c]	1:80
pCB8D2-2J-2-9-1	PEG-precipitated culture fluid[a]	1:640
pCB8D2-2J-2-9-1	PEG-precipitated, ethanol-extracted culture fluid[b]	1:80
pCB8D2-2J-2-9-1	Hydrophobic membrane protein preparation[c]	1:80
pEGFP	PEG-precipitated culture fluid[a]	<1:10
pEGFP	PEG-precipitated, ethanol-extracted culture fluid[b]	<1:10
pEGFP	Hydrophobic membrane protein preparation[c]	<1:10

[a]Culture supernatant from plasmid-transformed cells was precipitated with 10% polyethylene glycol (PEG) and resuspended in 1/100th of orginal volume.

[b]PEG-precipitated culture supernatant was extracted with 4% ethanol to remove PEG and the pellet was resuspended in 1/5 of the volume extracted.

[c]Hydrophobic membrane fractions were prepared as described in Materials and Methods.

Table 12. Immunogenicity of three DEN-2 recombinant plasmids in ICR mice.

Plasmid DNA[b]	Mouse #	ELISA on DEN-2 virus					PRNT on DEN-2 virus[c]	PRNT on JE virus[d]
		Serum 3 wks, p.v.[e]		Serum 6 wks, p.v.[e]		Endpoint titer 9 wks, p.v.	Endpoint titer 9 wks, p.v.	Endpoint titer 9 wks, p.v.
		1:100	1:400	1:100	1:400			
pCBD3-U-3-9-1	Pool, 1,2,4-10	ND[f]	ND	+	+	64,000	ND	ND
	1	+	+	+	+	64,000	>1000	<5
	2	+	+	+	+	32,000	>1000	<5
	4	+	+	+	+	16,000	200	<5
	5	+	+	+	+	4,000	<10	<5
	6	+	+	+	+	16,000	200	<5
	7	+	-	+	+	64,000	100	<5
	8	+	-	+	+	8,000	40	<5
	9	+	+	+	+	6,400	<5	<5
	10	+	+	+	+	64,000	>1000	<5
pCBD3-U-4-3	Pool, 1-10	ND	ND	+	+	1,000	ND	<5[g]
	1	-	-	+	-	400	<10	ND
	2	+	-	+	+	200	<10	ND
	3	+	+	+	+	4,000	<5	<5
	4	+	-	+	-	200	<10	ND
	5	+	-	+	+	400	<10	ND
	6	+	+	+	-	4,000	<5	5
	7	-	+/-	-	-	100	<10	ND
	8	-	-	-	-	200	<10	ND
	9	+	-	+	-	4,000	<5	<5
	10	-	-	+	+	4,000	<5	<5
pCBD3-14-8	Pool, 1-10	ND	ND	+	-	200	<5[g]	<5[g]
	1	-	-	-	-	400	<10	ND
	2,3,6-9	-	-	-	-	<100	ND	ND
	4	+	+	+	+	1,000	<5	<5
	5	-	-	+	-	2,000	5	<5
	10	+	-	-	-	<100	ND	ND
pCDFF	Pool, 1-10	-	ND	-	ND	<100	<5	<5

[a]PRNT, plaque-reduction neutralization test, 90% neutralization endpoint.

[b]Mice were immunized intramuscularly with 100 μg plasmid DNA on weeks 0 and 3.

[c]ELISA screens used sera diluted 1:100 and 1:400.

[d]ND, not done.

[e]Pool, 1,2,4,5,7,8.

[f]Pool, 2,3,6-10.

[g]Pool, 1-3,6-10.

Table 13. Summary of the characteristics of three DEN-2 recombinant plasmids.

Plasmid	IFA[a]		Ag-capture ELISA titer		ELISA titer on DEN-2[b]		DEN-2 PRNT[c]
	+/–	Globular/Diffuse	Secreted antigen	Hydrophobic membrane protein prep	No. sera ≥ 1:100	Pooled sera titer	No. sera ≥ 1:10
pCB8D2-2J-2-9-1	+	Diffuse	1:640	1:80	9/9	1:64000	7/9[d]
pCB9D2-1J-4-3	+	Globular	<1:10	1:80	10/10	1:1000	0/10
pCBD2-14-6	+	Globular	<1:10	1:160	3/10	1:200	0/10

[a]Indirect fluorescent antibody assay (IFA) staining characteristics, + or –, and diffuse or globular pattern.

[b]Anti-DEN-2 ELISA titer of sera from mice immunized with the recombinant plasmids. Sera were collected 9 weeks post-vaccination (weeks 0 and 3). Number of mice with titer of ≥ 1:100/ total number of mice is shown, including the endpoint ELISA titer of the pooled serum sample.

[c]Number of mice with plaque-reduction neutralization titers (PRNT, 90% reduction) ≥ 1:10/ total number of mice. Sera were collected 9 weeks post-vaccination.

[d]Of the 7 mice with neutralizing antibody, 3 mice had PRNT titers of >1:1000, 3 had titers of ≥1:100<1:1000, and one had a titer of 1:40.

What is claimed is:

1. An isolated nucleic acid comprising a transcriptional unit encoding a signal sequence of a structural protein of a first flavivirus and an immunogenic flavivirus antigen of a second flavivirus, wherein the transcriptional unit directs the synthesis of the antigen.

2. The nucleic acid of claim 1, wherein the signal sequence is a Japanese encephalitis virus signal sequence.

3. The nucleic acid of claim 1, wherein the immunogenic flavivirus antigen is of a flavivirus selected from the group consisting of yellow fever virus, dengue serotype 1 virus, dengue serotype 2 virus, dengue serotype 3 virus, dengue serotype 4 virus, Japanese encephalitis virus, Powassan virus and West Nile virus.

4. The nucleic acid of claim 1, wherein the transcriptional unit encodes a signal sequence of Japanese encephalitis virus and an M protein and an E protein of West Nile virus.

5. The nucleic acid of claim 1, wherein the transcriptional unit encodes a signal sequence of Japanese encephalitis virus and an M protein and an E protein of yellow fever virus.

6. The nucleic acid of claim 1, wherein the transcriptional unit encodes a signal sequence of Japanese encephalitis virus and an M protein and an E protein of St. Louis encephalitis virus.

7. The nucleic acid of claim 1, wherein the transcriptional unit encodes a signal sequence of Japanese encephalitis virus and an M protein and an E protein of Powassan virus.

8. The nucleic acid of claim 1, wherein the antigen is selected from the group

consisting of an M protein of a flavivirus, an E protein of a flavivirus, both an M protein and an E protein of a flavivirus, a portion of an M protein of a flavivirus, a portion of an E protein of a flavivirus and both a portion of an M protein of a flavivirus and a portion of an E protein of a flavivirus or any combination thereof.

9. The nucleic acid of claim 8, wherein the antigen is both the M protein and the E protein of a flavivirus.

10. The nucleic acid of claim 1, wherein the nucleic acid is DNA.

11. The nucleic acid of claim 10, comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:15, SEQ ID NO:19, SEQ ID NO:21 and SEQ ID NO:23.

12. The nucleic acid of claim 1, wherein the transcriptional unit comprises a control sequence disposed appropriately such that it operably controls the synthesis of the antigen.

13. The nucleic acid of claim 12, wherein the control sequence is the cytomegalovirus immediate early promoter.

14. The nucleic acid of claim 1, comprising a Kozak consensus sequence located at a translational start site for a polypeptide comprising the antigen encoded by the TU.

15. The nucleic acid of claim 1 wherein the transcriptional unit comprises a poly-A terminator.

16. A cell comprising the nucleic acid of claim 1.

17. A composition comprising the nucleic acid of claim 1 and a pharmaceutically acceptable carrier.

18. A method of immunizing a subject against infection by a flavivirus, comprising administering to the subject an effective amount of the composition of claim 17.

19. The method of claim 18, wherein the flavivirus antigen is of a flavivirus selected from the group consisting of yellow fever virus, dengue serotype 1 virus, dengue serotype 2 virus, dengue serotype 3 virus, dengue serotype 4 virus, Japanese encephalitis virus, Powassan virus and West Nile virus.

20. The method of claim 18, wherein the antigen is selected from the group consisting of an M protein of a flavivirus, an E protein of a flavivirus, both an M protein and an E protein of a flavivirus, a portion of an M protein of a flavivirus, a portion of an E protein of a flavivirus and both a portion of an M protein of a flavivirus and a portion of an E protein of a flavivirus or any combination thereof.

21. The method of claim 20, wherein the antigen is both the M protein and the E protein of a flavivirus, and wherein a cell within the body of the subject, after incorporating the nucleic acid within it, secretes subviral particles comprising the M protein and the E protein.

22. The method of claim 18, wherein the transcriptional unit encodes a signal sequence of Japanese encephalitis virus, and an M protein and an E protein of West Nile virus.

23. The method of claim 18, wherein the transcriptional unit encodes a signal sequence of Japanese encephalitis virus, and an M protein and an E protein of yellow fever virus.

24. The method of claim 18, wherein the transcriptional unit encodes a signal sequence of Japanese encephalitis virus, and an M protein and an E protein of St. Louis encephalitis virus.

25. The method of claim 18, wherein the transcriptional unit encodes a signal sequence of Japanese encephalitis virus, and an M protein and an E protein of Powassan virus.

26. The method of claim 18, comprising administering the composition to the subject in a single dose.

27. The method of claim 18, wherein the composition is administered via a parenteral route.

28. The nucleic acid of claim 1, wherein the antigen is a St. Louis encephalitis virus antigen.

29. The method of claim 18, wherein the antigen is a St. Louis encephalitis virus antigen.

30. The nucleic acid of claim 1, wherein the antigen is a Japanese encephalitis virus antigen.

31. The method of claim 18, wherein the antigen is a Japanese encephalitis virus antigen.

32. The nucleic acid of claim 1, wherein the antigen is a yellow fever virus antigen.

33. The method of claim 18, wherein the antigen is a yellow fever virus antigen.

34. The nucleic acid of claim 1, wherein the antigen is a dengue virus antigen.

35. The method of claim 18, wherein the antigen is a dengue virus antigen.

36. The nucleic acid of claim 1, wherein the antigen is a West Nile virus antigen.

37. The method of claim 18, wherein the antigen is a West Nile virus antigen.

38. An antigen produced from the nucleic acid of claim 1.

39. A method of detecting a flavivirus antibody in a sample, comprising:

(a) contacting the sample with the antigen of claim 38 under conditions whereby an antigen/antibody complex can form; and

(b) detecting antigen/antibody complex formation, thereby detecting a flavivirus antibody in the sample.

40. An antibody produced in response to immunization by the antigen of claim 38.

41. A method of detecting a flavivirus antigen in a sample, comprising:

(a) contacting the sample with the antibody of claim 40 under conditions whereby an antigen/antibody complex can form; and

(b) detecting antigen/antibody complex formation, thereby detecting a flavivirus antigen in a sample.

42. A method of diagnosing a flavivirus infection in a subject, comprising:

(a) contacting a sample from the subject with the antigen of claim 38 under conditions whereby an antigen/antibody complex can form; and

(b) detecting antigen/antibody complex formation, thereby diagnosing a flavivirus infection in a subject.

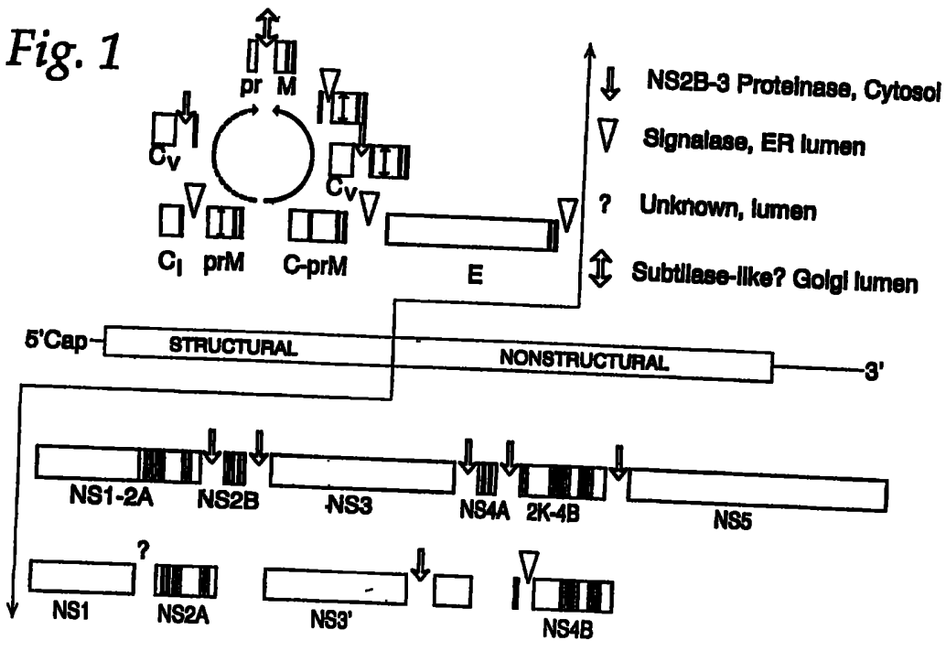
43. A method of diagnosing a flavivirus infection in a subject, comprising:

(a) contacting a sample from the subject with the antibody of claim 40 under conditions whereby an antigen/antibody complex can form; and

(b) detecting antigen/antibody complex formation, thereby diagnosing a flavivirus infection in a subject.

PCT/US02/10754

Fig. 1



WO 02/081754

PCT/US02/10764

WO 02/081754

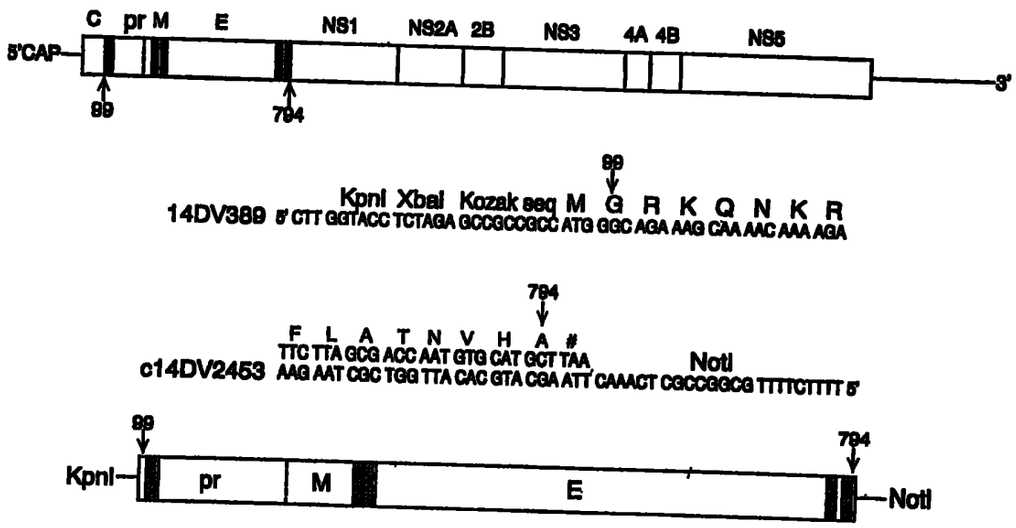


Fig. 2

PCT/US92/10764

WO 93/19764

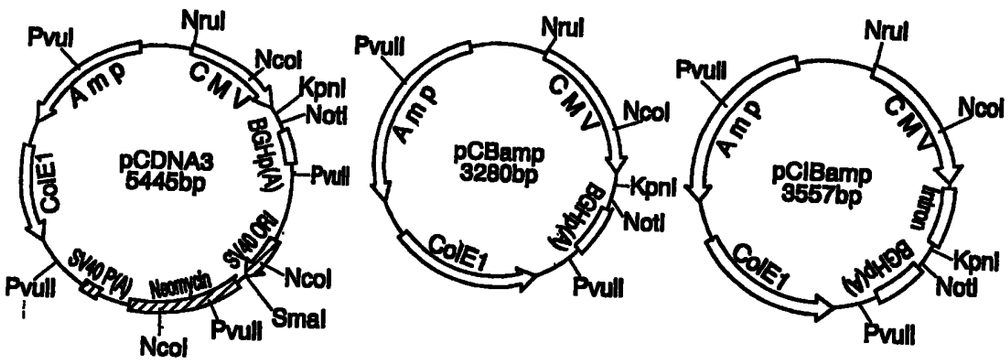


Fig. 3

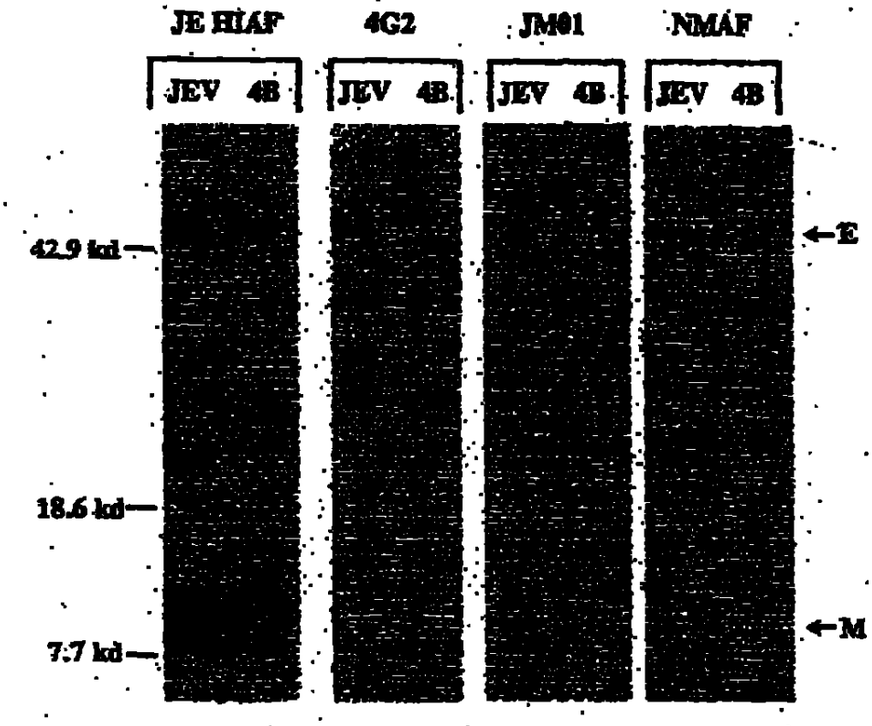

JE HIAF
4G2
JM01
NMAF
JEV 4B
JEV 4B
JEV 4B
JEV 4B
42.9 kd
18.6 kd
7.7 kd
E
M

Fig. 4

- 14 -

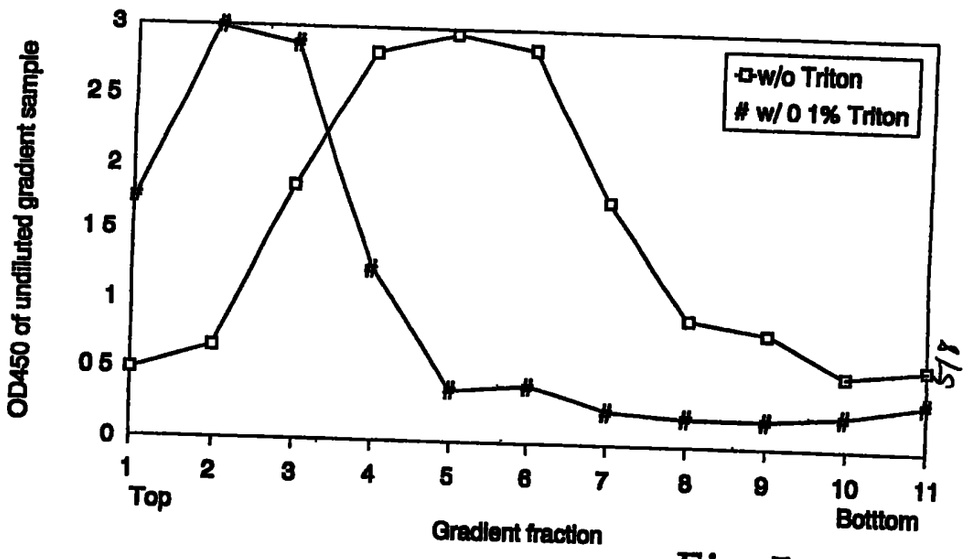
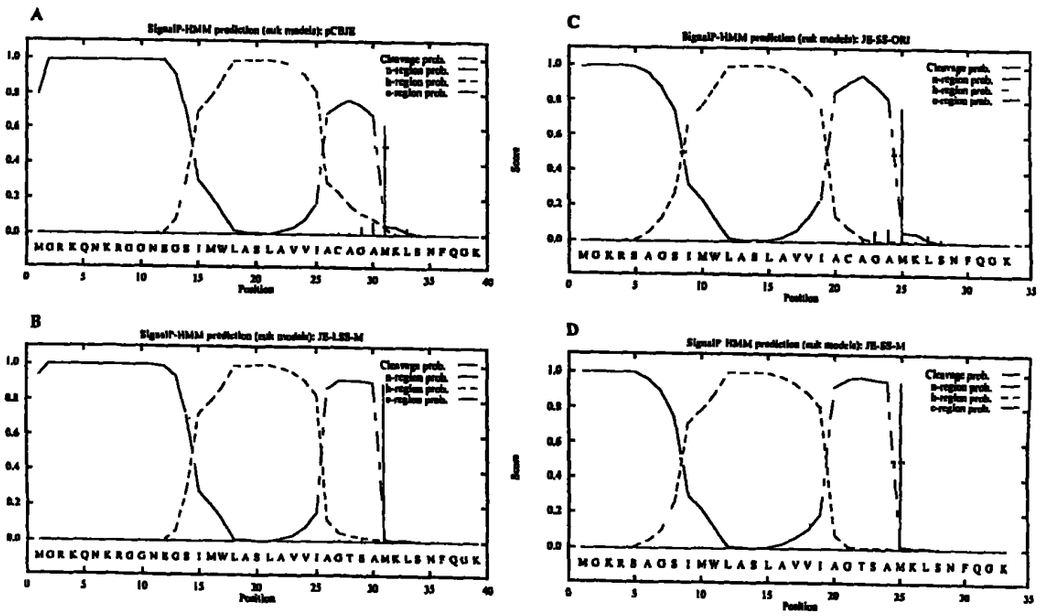

OD450 of undiluted gradient sample
3
2 5
2
1 5
1
0 5
0
1
2
3
4
5
6
7
8
9
10
11
Top
Botttom
Gradient fraction
w/o Triton
w/ 0 1% Triton

Fig. 5

FIG 6


A
SignalP-HMM prediction (mk models): pCBJE
Cleavage prob.
n-region prob.
h-region prob.
c-region prob.
M G R K Q N K R G G N E G S I M W L A S L A V V I A C A G A M K L S N F Q G K
Score
Position
C
SignalP-HMM prediction (mk models): JB-SS-ORI
M G K R S A G S I M W L A S L A V V I A C A G A M K L S N F Q G K
B
SignalP-HMM prediction (mk models): JB-LSS-M
M G R K Q N K R G G N E G S I M W L A S L A V V I A G T S A M K L S N F Q G K
D
SignalP-HMM prediction (mk models): JB-SS-M
M G K R S A G S I M W L A S L A V V I A G T S A M K L S N F Q G K

PCT/US02/10764

WO 02/081754

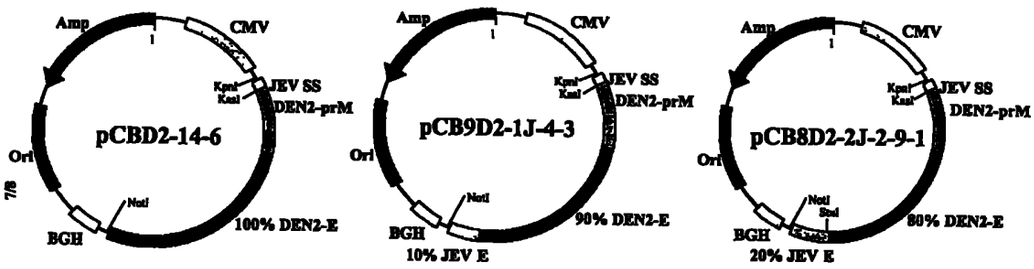


FIG 7

010113
PCT/US02/10764
8/8
WO 02/081754

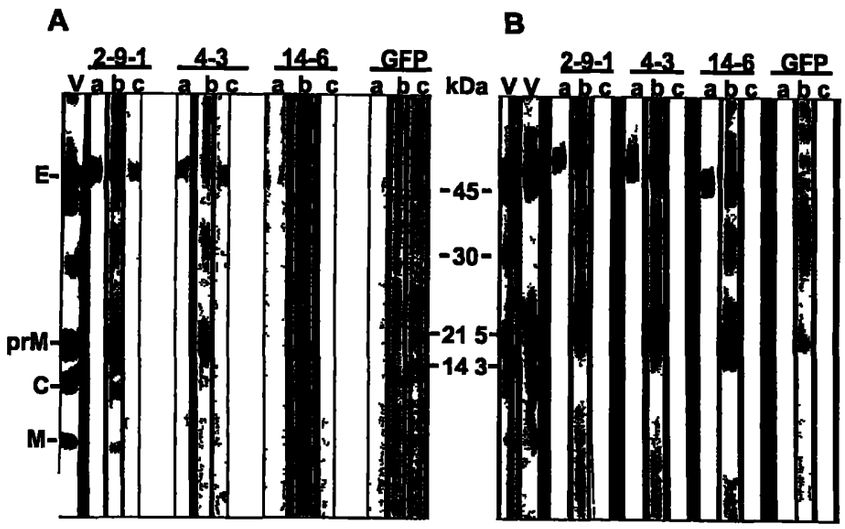


FIG 8

00114

SEQUENCE LISTING

<110> The Government of the United States of America, as represented by the Secretary, Department of Health and Human Services, c/o Centers for Disease Control and Prevention

Chang, Gwong-Jen J

<120> Nucleic Acid Vaccines for Prevention of Flavivirus Infection

<130> 14114.0332P2

<150> 09/826,115
<151> 2001-04-04

<150> 09/701,536
<151> 2000-11-29

<150> PCT/US99/12398
<151> 1999-06-03

<150> 60/087,908
<151> 1998-06-04

<160> 61

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note = synthetic construct

<221> CDS
<222> (25)...(48)

<221> misc_feature
<222> 1-48
<223> Amplimer 14DV389

<400> 1
cttggtacct ctagagccgc cgcc atg ggc aga aag caa aac aaa aga 48
Met Gly Arg Lys Gln Asn Lys Arg
1 5

<210> 2
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note = synthetic construct

<400> 2
Met Gly Arg Lys Gln Asn Lys Arg
1 5

<210> 3
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-50
<223> Amplimer c14DV2453

<400> 3
ttttcttttg cggccgctca aacttaagca tgcacattgg tcgctaagaa 50

<210> 4
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (25)...(48)

<221> misc_feature
<222> (1)...(48)
<223> Amplimer YFDV389

<400> 4
cttggtacct ctagagccgc cgcc atg cgt tcc cat gat gtt ctg act 48
Met Arg Ser His Asp Val Leu Thr
1 5

<210> 5
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 5
Met Arg Ser His Asp Val Leu Thr
1 5

<210> 6
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-41
<223> Amplimer cYFDV2452

<400> 6
ttttcttttg cggccgctca cgcccaaact cctagagaaa c 41

<210> 7
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (25)...(54)

<221> misc_feature
<222> 1-54
<223> Amplimer SLEDV410

<400> 7
cttggtacct ctagagccgc cgcc atg tct aaa aaa aga gga ggg acc aga 51
Met Ser Lys Lys Arg Gly Gly Thr Arg
1 5

<210> 8
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 8
Met Ser Lys Lys Arg Gly Gly Thr Arg
1 5

<210> 9
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-38
<223> Amplimer cSLEDV2449

<400> 9
ttttcttttg cggccgctta ggcttgcacg ctggttgc 38

<210> 10
<211> 7500
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (916)...(3009)

<221> misc_feature
<222> 1-7500
<223> pCDJE 2-7

<400> 10
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc 900
gagctcgccg ccgcc atg ggc aga aag caa aac aaa aga gga gga aat gaa 951
Met Gly Arg Lys Gln Asn Lys Arg Gly Gly Asn Glu
1 5 10

ggc tca atc atg tgg ctc gcg agc ttg gca gtt gtc ata gct tgt gcg 999
Gly Ser Ile Met Trp Leu Ala Ser Leu Ala Val Val Ile Ala Cys Ala
15 20 25

gga gcc atg aag ttg tcg aat ttc cag ggg aag ctt ttg atg acc atc 1047
Gly Ala Met Lys Leu Ser Asn Phe Gln Gly Lys Leu Leu Met Thr Ile
30 35 40

aac aac acg gac att gca gac gtt atc gtg att ccc acc tca aaa gga 1095
Asn Asn Thr Asp Ile Ala Asp Val Ile Val Ile Pro Thr Ser Lys Gly
45 50 55 60

gag aac aga tgc tgg gtc cgg gca atc gac gtc ggc tac atg tgt gag 1143
Glu Asn Arg Cys Trp Val Arg Ala Ile Asp Val Gly Tyr Met Cys Glu
65 70 75

gac act atc acg tac gaa tgt cct aag ctt acc atg ggc aat gat cca 1191
Asp Thr Ile Thr Tyr Glu Cys Pro Lys Leu Thr Met Gly Asn Asp Pro
80 85 90

gag gat gtg gat tgc tgg tgt gac aac caa gaa gtc tac gtc caa tat 1239
Glu Asp Val Asp Cys Trp Cys Asp Asn Gln Glu Val Tyr Val Gln Tyr
95 100 105

```
gga cgg tgc acg cgg acc agg cat tcc aag cga agc agg aga tcc gtg     1287
Gly Arg Cys Thr Arg Thr Arg His Ser Lys Arg Ser Arg Arg Ser Val
    110                 115                 120

tcg gtc caa aca cat ggg gag agt tca cta gtg aat aaa aaa gag gct     1335
Ser Val Gln Thr His Gly Glu Ser Ser Leu Val Asn Lys Lys Glu Ala
125                 130                 135                 140

tgg ctg gat tca acg aaa gcc aca cga tat ctc atg aaa act gag aac     1383
Trp Leu Asp Ser Thr Lys Ala Thr Arg Tyr Leu Met Lys Thr Glu Asn
                145                 150                 155

tgg atc ata agg aat cct ggc tat gct ttc ctg gcg gcg gta ctt ggc     1431
Trp Ile Ile Arg Asn Pro Gly Tyr Ala Phe Leu Ala Ala Val Leu Gly
            160                 165                 170

tgg atg ctt ggc agt aac aac ggt caa cgc gtg gta ttt acc atc ctc     1479
Trp Met Leu Gly Ser Asn Asn Gly Gln Arg Val Val Phe Thr Ile Leu
        175                 180                 185

ctg ctg ttg gtc gct ccg gct tac agt ttt aat tgt ctg gga atg ggc     1527
Leu Leu Leu Val Ala Pro Ala Tyr Ser Phe Asn Cys Leu Gly Met Gly
    190                 195                 200

aat cgt gac ttc ata gaa gga gcc agt gga gcc act tgg gtg gac ttg     1575
Asn Arg Asp Phe Ile Glu Gly Ala Ser Gly Ala Thr Trp Val Asp Leu
205                 210                 215                 220

gtg ctg gaa gga gat agc tgc ttg aca atc atg gca aac gac aaa cca     1623
Val Leu Glu Gly Asp Ser Cys Leu Thr Ile Met Ala Asn Asp Lys Pro
                225                 230                 235

aca ttg gac gtc cgc atg att aac atc gaa gct agc caa ctt gct gag     1671
Thr Leu Asp Val Arg Met Ile Asn Ile Glu Ala Ser Gln Leu Ala Glu
            240                 245                 250

gtc aga agt tac tgc tat cat gct tca gtc act gac atc tcg acg gtg     1719
Val Arg Ser Tyr Cys Tyr His Ala Ser Val Thr Asp Ile Ser Thr Val
        255                 260                 265

gct cgg tgc ccc acg act gga gaa gcc cac aac gag aag cga gct gat     1767
Ala Arg Cys Pro Thr Thr Gly Glu Ala His Asn Glu Lys Arg Ala Asp
    270                 275                 280

agt agc tat gtg tgc aaa caa ggc ttc act gac cgt ggg tgg ggc aac     1815
Ser Ser Tyr Val Cys Lys Gln Gly Phe Thr Asp Arg Gly Trp Gly Asn
285                 290                 295                 300

gga tgt gga ctt ttc ggg aag gga agc att gac aca tgt gca aaa ttc     1863
Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Asp Thr Cys Ala Lys Phe
                305                 310                 315

tcc tgc acc agt aaa gcg att ggg aga aca atc cag cca gaa aac atc     1911
Ser Cys Thr Ser Lys Ala Ile Gly Arg Thr Ile Gln Pro Glu Asn Ile
            320                 325                 330

aaa tac gaa gtt ggc att ttt gtg cat gga acc acc act tcg gaa aac     1959
Lys Tyr Glu Val Gly Ile Phe Val His Gly Thr Thr Thr Ser Glu Asn
        335                 340                 345
```

```
cat ggg aat tat tca gcg caa gtt ggg gcg tcc cag gcg gca aag ttt     2007
His Gly Asn Tyr Ser Ala Gln Val Gly Ala Ser Gln Ala Ala Lys Phe
    350                 355                 360

aca gta aca ccc aat gct cct tcg ata acc ctc aaa ctt ggt gac tac     2055
Thr Val Thr Pro Asn Ala Pro Ser Ile Thr Leu Lys Leu Gly Asp Tyr
365                 370                 375                 380

gga gaa gtc aca ctg gac tgt gag cca agg agt gga ctg aac act gaa     2103
Gly Glu Val Thr Leu Asp Cys Glu Pro Arg Ser Gly Leu Asn Thr Glu
                385                 390                 395

gcg ttt tac gtc atg acc gtg ggg tca aag tca ttt ctg gtc cat agg     2151
Ala Phe Tyr Val Met Thr Val Gly Ser Lys Ser Phe Leu Val His Arg
            400                 405                 410

gag tgg ttt cat gac ctc gct ctc ccc tgg acg tcc cct tcg agc aca     2199
Glu Trp Phe His Asp Leu Ala Leu Pro Trp Thr Ser Pro Ser Ser Thr
        415                 420                 425

gcg tgg aga aac aga gaa ctc ctc atg gaa ttt gaa gag gcg cac gcc     2247
Ala Trp Arg Asn Arg Glu Leu Leu Met Glu Phe Glu Glu Ala His Ala
    430                 435                 440

aca aaa cag tcc gtt gtt gct ctt ggg tca cag gaa gga ggc ctc cat     2295
Thr Lys Gln Ser Val Val Ala Leu Gly Ser Gln Glu Gly Gly Leu His
445                 450                 455                 460

cag gcg ttg gca gga gcc atc gtg gtg gag tac tca agc tca gtg aag     2343
Gln Ala Leu Ala Gly Ala Ile Val Val Glu Tyr Ser Ser Ser Val Lys
                465                 470                 475

tta aca tca ggc cac ctg aaa tgt agg ctg aaa atg gac aaa ctg gct     2391
Leu Thr Ser Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Ala
            480                 485                 490

ctg aaa ggc aca acc tat ggc atg tgt aca gaa aaa ttc tcg ttc gcg     2439
Leu Lys Gly Thr Thr Tyr Gly Met Cys Thr Glu Lys Phe Ser Phe Ala
        495                 500                 505

aaa aat ccg gcg gac act ggt cac gga aca gtt gtc att gaa ctc tcc     2487
Lys Asn Pro Ala Asp Thr Gly His Gly Thr Val Val Ile Glu Leu Ser
    510                 515                 520

tac tct ggg agt gat ggc ccc tgc aaa att ccg att gct tcc gtt gcg     2535
Tyr Ser Gly Ser Asp Gly Pro Cys Lys Ile Pro Ile Ala Ser Val Ala
525                 530                 535                 540

agc ctc aat gac atg acc ccc gtt ggg cgg ctg gtg aca gtg aac ccc     2583
Ser Leu Asn Asp Met Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro
                545                 550                 555

ttc gtc gcg act tcc agt gcc agc tca aag gtg ctg gtc gag atg gaa     2631
Phe Val Ala Thr Ser Ser Ala Ser Ser Lys Val Leu Val Glu Met Glu
            560                 565                 570

ccc ccc ttc gga gac tcc tac atc gta gtt gga agg gga gac aag cag     2679
Pro Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Asp Lys Gln
        575                 580                 585
```

```
atc aac cac cat tgg cac aaa gct gga agc acg ctg ggc aag gcc ttt     2727
Ile Asn His His Trp His Lys Ala Gly Ser Thr Leu Gly Lys Ala Phe
    590                 595                 600

tca aca act ttg aag gga gct caa aga ctg gca gcg ttg ggc gac aca     2775
Ser Thr Thr Leu Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr
605                 610                 615                 620

gcc tgg gac ttt ggc tct att gga ggg gtc ttc aac tcc ata gga aaa     2823
Ala Trp Asp Phe Gly Ser Ile Gly Gly Val Phe Asn Ser Ile Gly Lys
                625                 630                 635

gcc gtt cac caa gtg ttt ggt ggt gcc ttc aga aca ctc ttt ggg gga     2871
Ala Val His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly
            640                 645                 650

atg tct tgg atc aca caa ggg cta atg ggt gcc cta ctg ctc tgg atg     2919
Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met
        655                 660                 665

ggc gtc aac gca cga gac cga tca att gct ttg gcc ttc tta gcc aca     2967
Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr
    670                 675                 680

ggg ggt gtg ctc gtg ttc tta gcg acc aat gtg cat gct taa             3009
Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala  *
685                 690                 695

ttagtttgag cggccgctcg agcatgcatc tagagggccc tattctatag tgtcacctaa   3069
atgctagagc tcgctgatca gcctcgactg tgccttctag ttgccagcca tctgttgttt   3129
gcccctcccc cgtgccttcc ttgaccctgg aaggtgccac tcccactgtc ctttcctaat   3189
aaaatgagga aattgcatcg cattgtctga gtaggtgtca ttctattctg gggggtgggg   3249
tggggcagga cagcaagggg gaggattggg aagacaatag caggcatgct ggggatgcgg   3309
tgggctctat ggcttctgag gcggaaagaa ccagctgggg ctctaggggg tatccccacg   3369
cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta   3429
cacttgccag cgccctagcg cccgctcctt tcgctttctt cccttccttt ctcgccacgt   3489
tcgccggctt tccccgtcaa gctctaaatc gggggcatccc tttagggttc cgatttagtg   3549
ctttacggca cctcgacccc aaaaaacttg attagggtga tggttcacgt agtgggccat   3609
cgccctgata gacggttttt cgccctttga cgttggagtc cacgttcttt aatagtggac   3669
tcttgttcca aactggaaca acactcaacc ctatctcggt ctattctttt gatttataag   3729
ggattttggg gatttcggcc tattggttaa aaaatgagct gatttaacaa aaatttaacg   3789
cgaattaatt ctgtggaatg tgtgtcagtt agggtgtgga aagtccccag gctccccagg   3849
caggcagaag tatgcaaagc atgcatctca attagtcagc aaccaggtgt ggaaagtccc   3909
caggctcccc agcaggcaga agtatgcaaa gcatgcatct caattagtca gcaaccatag   3969
tcccgccccc aactccgccc atcccgcccc taactccgcc cagttccgcc cattctccgc   4029
cccatggctg actaattttt tttatttatg cagaggccga ggccgcctct gcctctgagc   4089
tattccagaa gtagtgagga ggcttttttg gaggcctagg cttttgcaaa aagctcccgg   4149
gagcttgtat atccatttcc ggatctgatc aagagacagg atgaggatcg tttcgcatga   4209
ttgaacaaga tggattgcac gcaggttctc cggccgcttg ggtggagagg ctattcggct   4269
atgactgggc acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc   4329
aggggcgccc ggttcttttt gtcaagaccg acctgtccgg tgccctgaat gaactgcagg   4389
acgaggcagc gcggctatcg tggctggcca cgacgggcgt tccttgcgca gctgtgctcg   4449
acgttgtcac tgaagcggga agggactggc tgctattggg cgaagtgccg gggcaggatc   4509
tcctgtcatc tcaccttgct cctgccgaga aagtatccat catggctgat gcaatgcggc   4569
ggctgcatac gcttgatccg gctacctgcc cattcgacca ccaagcgaaa catcgcatcg   4629
agcgagcacg tactcggatg gaagccggtc ttgtcgatca ggatgatctg gacgaagagc   4689
atcaggggct cgcgccagcc gaactgttcg ccaggctcaa ggcgcgcatg cccgacggcg   4749
aggatctcgt cgtgacccat ggcgatgcct gcttgccgaa tatcatggtg gaaaatggcc   4809
gcttttctgg attcatcgac tgtggccggc tgggtgtggc ggaccgctat caggacatag   4869
cgttggctac ccgtgatatt gctgaagagc ttggcggcga atgggctgac cgcttcctcg   4929
tgctttacgg tatcgccgct cccgattcgc agcgcatcgc cttctatcgc cttcttgacg   4989
```

```
agttcttctg agcgggactc tggggttcga aatgaccgac caagcgacgc ccaacctgcc   5049
atcacgagat ttcgattcca ccgccgcctt ctatgaaagg ttgggcttcg gaatcgtttt   5109
ccgggacgcc ggctggatga tcctccagcg cggggatctc atgctggagt tcttcgccca   5169
ccccaacttg tttattgcag cttataatgg ttacaaataa agcaatagca tcacaaattt   5229
cacaaataaa gcatttttt cactgcattc tagttgtggt ttgtccaaac tcatcaatgt   5289
atcttatcat gtctgtatac cgtcgacctc tagctagagc ttggcgtaat catggtcata   5349
gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag   5409
cataaagtgt aaagcctggg gtgcctaatg agtgagctaa ctcacattaa ttgcgttgcg   5469
ctcactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca   5529
acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc tcactgactc   5589
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg   5649
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa   5709
ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gcccccctga   5769
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag   5829
ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct   5889
taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc aatgctcacg   5949
ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc   6009
ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt   6069
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta   6129
tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggac   6189
agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc   6249
ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat   6309
tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc   6369
tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt   6429
cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta   6489
aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct   6549
atttcgttca tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg   6609
cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga   6669
tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt   6729
atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt   6789
taatagtttg cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt   6849
tggtatggct tcattcagct ccggttccca acgatcaagg cgagttacat gatccccat   6909
gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc gttgtcagaa gtaagttggc   6969
cgcagtgtta tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc   7029
cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat   7089
gcggcgaccg agttgctctt gcccggcgtc aatacgggat ataccgcgc cacatagcag   7149
aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt   7209
accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc   7269
ttttactttc accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa   7329
gggaataagg gcgacacgga aatgttgaat actcatactc ttcctttttc aatattattg   7389
aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa   7449
taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg t             7500
```

<210> 11
<211> 697
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<223> pCDJE 2-7

<400> 11

```
Met Gly Arg Lys Gln Asn Lys Arg Gly Gly Asn Glu Gly Ser Ile Met
 1               5                  10                  15
Trp Leu Ala Ser Leu Ala Val Val Ile Ala Cys Ala Gly Ala Met Lys
            20                  25                  30
Leu Ser Asn Phe Gln Gly Lys Leu Leu Met Thr Ile Asn Asn Thr Asp
        35                  40                  45
```

Ile Ala Asp Val Ile Val Ile Pro Thr Ser Lys Gly Glu Asn Arg Cys
50 55 60
Trp Val Arg Ala Ile Asp Val Gly Tyr Met Cys Glu Asp Thr Ile Thr
65 70 75 80
Tyr Glu Cys Pro Lys Leu Thr Met Gly Asn Asp Pro Glu Asp Val Asp
85 90 95
Cys Trp Cys Asp Asn Gln Glu Val Tyr Val Gln Tyr Gly Arg Cys Thr
100 105 110
Arg Thr Arg His Ser Lys Arg Ser Arg Arg Ser Val Ser Val Gln Thr
115 120 125
His Gly Glu Ser Ser Leu Val Asn Lys Lys Glu Ala Trp Leu Asp Ser
130 135 140
Thr Lys Ala Thr Arg Tyr Leu Met Lys Thr Glu Asn Trp Ile Ile Arg
145 150 155 160
Asn Pro Gly Tyr Ala Phe Leu Ala Ala Val Leu Gly Trp Met Leu Gly
165 170 175
Ser Asn Asn Gly Gln Arg Val Val Phe Thr Ile Leu Leu Leu Leu Val
180 185 190
Ala Pro Ala Tyr Ser Phe Asn Cys Leu Gly Met Gly Asn Arg Asp Phe
195 200 205
Ile Glu Gly Ala Ser Gly Ala Thr Trp Val Asp Leu Val Leu Glu Gly
210 215 220
Asp Ser Cys Leu Thr Ile Met Ala Asn Asp Lys Pro Thr Leu Asp Val
225 230 235 240
Arg Met Ile Asn Ile Glu Ala Ser Gln Leu Ala Glu Val Arg Ser Tyr
245 250 255
Cys Tyr His Ala Ser Val Thr Asp Ile Ser Thr Val Ala Arg Cys Pro
260 265 270
Thr Thr Gly Glu Ala His Asn Glu Lys Arg Ala Asp Ser Ser Tyr Val
275 280 285
Cys Lys Gln Gly Phe Thr Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu
290 295 300
Phe Gly Lys Gly Ser Ile Asp Thr Cys Ala Lys Phe Ser Cys Thr Ser
305 310 315 320
Lys Ala Ile Gly Arg Thr Ile Gln Pro Glu Asn Ile Lys Tyr Glu Val
325 330 335
Gly Ile Phe Val His Gly Thr Thr Thr Ser Glu Asn His Gly Asn Tyr
340 345 350
Ser Ala Gln Val Gly Ala Ser Gln Ala Ala Lys Phe Thr Val Thr Pro
355 360 365
Asn Ala Pro Ser Ile Thr Leu Lys Leu Gly Asp Tyr Gly Glu Val Thr
370 375 380
Leu Asp Cys Glu Pro Arg Ser Gly Leu Asn Thr Glu Ala Phe Tyr Val
385 390 395 400
Met Thr Val Gly Ser Lys Ser Phe Leu Val His Arg Glu Trp Phe His
405 410 415
Asp Leu Ala Leu Pro Trp Thr Ser Pro Ser Ser Thr Ala Trp Arg Asn
420 425 430
Arg Glu Leu Leu Met Glu Phe Glu Glu Ala His Ala Thr Lys Gln Ser
435 440 445
Val Val Ala Leu Gly Ser Gln Glu Gly Gly Leu His Gln Ala Leu Ala
450 455 460
Gly Ala Ile Val Val Glu Tyr Ser Ser Ser Val Lys Leu Thr Ser Gly
465 470 475 480
His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Ala Leu Lys Gly Thr
485 490 495
Thr Tyr Gly Met Cys Thr Glu Lys Phe Ser Phe Ala Lys Asn Pro Ala
500 505 510
Asp Thr Gly His Gly Thr Val Val Ile Glu Leu Ser Tyr Ser Gly Ser
515 520 525

```
Asp Gly Pro Cys Lys Ile Pro Ile Ala Ser Val Ala Ser Leu Asn Asp
    530                 535                 540
Met Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val Ala Thr
545                 550                 555                 560
Ser Ser Ala Ser Ser Lys Val Leu Val Glu Met Glu Pro Pro Phe Gly
                565                 570                 575
Asp Ser Tyr Ile Val Val Gly Arg Gly Asp Lys Gln Ile Asn His His
            580                 585                 590
Trp His Lys Ala Gly Ser Thr Leu Gly Lys Ala Phe Ser Thr Thr Leu
        595                 600                 605
Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr Ala Trp Asp Phe
    610                 615                 620
Gly Ser Ile Gly Gly Val Phe Asn Ser Ile Gly Lys Ala Val His Gln
625                 630                 635                 640
Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly Met Ser Trp Ile
                645                 650                 655
Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met Gly Val Asn Ala
            660                 665                 670
Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr Gly Gly Val Leu
        675                 680                 685
Val Phe Leu Ala Thr Asn Val His Ala
    690                 695
```

<210> 12
<211> 46
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-46
<223> NN 466

<400> 12
cttggtaccc gtctcggcgc cgtgaccctc tcgaacttcc agggca 46

<210> 13
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-43
<223> CNN2444

<400> 13
agaggcactt gcacgtgcgg acttccgccg gcgaaaaaga aaa 43

<210> 14
<211> 24
<212> PRT
<213> Artificial Sequence

<220>

<223> Description of artificial sequence; note =
synthetic construct

<223> JE Signal

<400> 14
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
1 5 10 15
Val Val Ile Ala Cys Ala Gly Ala
20

<210> 15
<211> 5308
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (911)...(2987)

<221> misc_feature
<222> (1)...(5308)
<223> pCBMN

<400> 15
gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggtttttgg cagtacatca atgggcgtgg atagcggttt 660
gactcacggg gatttccaag tctccaccce attgacgtca atgggagttt gttttggcac 720
caaaatcaac gggactttcc aaaatgtcgt aacaactccg ccccattgac gcaaatgggc 780
ggtaggcgtg tacggtggga ggtctatata agcagagctc tctggctaac tagagaaccc 840
actgcttact ggcttatcga aattaatacg actcactata gggagaccca agcttggtac 900
cgccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg 949
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala
1 5 10

agc ttg gca gtt gtc ata gct tgt gca ggc gcc gtg acc ctc tcg aac 997
Ser Leu Ala Val Val Ile Ala Cys Ala Gly Ala Val Thr Leu Ser Asn
15 20 25

ttc cag ggc aag gtg atg atg acg gta aat gct act gac gtc aca gat 1045
Phe Gln Gly Lys Val Met Met Thr Val Asn Ala Thr Asp Val Thr Asp
30 35 40 45

gtc atc acg att cca aca gct gct gga aag aac cta tgc att gtc aga 1093
Val Ile Thr Ile Pro Thr Ala Ala Gly Lys Asn Leu Cys Ile Val Arg
50 55 60

```
gca atg gat gtg gga tac atg tgc gat gat act atc act tat gaa tgc      1141
Ala Met Asp Val Gly Tyr Met Cys Asp Asp Thr Ile Thr Tyr Glu Cys
               65                  70                  75

cca gtg ctg tcg gct ggt aat gat cca gaa gac atc gac tgt tgg tgc      1189
Pro Val Leu Ser Ala Gly Asn Asp Pro Glu Asp Ile Asp Cys Trp Cys
          80                  85                  90

aca aag tca gca gtc tac gtc agg tat gga aga tgc acc aag aca cgc      1237
Thr Lys Ser Ala Val Tyr Val Arg Tyr Gly Arg Cys Thr Lys Thr Arg
     95                 100                 105

cac tca aga cgc agt cgg agg tca ctg aca gtg cag aca cac gga gaa      1285
His Ser Arg Arg Ser Arg Arg Ser Leu Thr Val Gln Thr His Gly Glu
110                 115                 120                 125

agc act cta gcg aac aag aag ggg gct tgg atg gac agc acc aag gcc      1333
Ser Thr Leu Ala Asn Lys Lys Gly Ala Trp Met Asp Ser Thr Lys Ala
                130                 135                 140

aga agg tat ttg gta aaa aca gaa tca tgg atc ttg agg aac cct gga      1381
Thr Arg Tyr Leu Val Lys Thr Glu Ser Trp Ile Leu Arg Asn Pro Gly
            145                 150                 155

tat gcc ctg gtg gca gcc gtc att ggt tgg atg ctt ggg agc aac acc      1429
Tyr Ala Leu Val Ala Ala Val Ile Gly Trp Met Leu Gly Ser Asn Thr
        160                 165                 170

atg cag aga gtt gtg ttt gtc gtg cta ttg ctt ttg gtg gcc cca gct      1477
Met Gln Arg Val Val Phe Val Val Leu Leu Leu Leu Val Ala Pro Ala
    175                 180                 185

tac agc ttc aac tgc ctt gga atg agc aac aga gac ttc ttg gaa gga      1525
Tyr Ser Phe Asn Cys Leu Gly Met Ser Asn Arg Asp Phe Leu Glu Gly
190                 195                 200                 205

gtg tct gga gca aca tgg gtg gat ttg gtt ctc gaa ggc gac agc tgc      1573
Val Ser Gly Ala Thr Trp Val Asp Leu Val Leu Glu Gly Asp Ser Cys
                210                 215                 220

gtg act atc atg tct aag gac aag cct acc atc gat gtg aag atg atg      1621
Val Thr Ile Met Ser Lys Asp Lys Pro Thr Ile Asp Val Lys Met Met
            225                 230                 235

aat atg gag gcg gcc aac ctg gca gag gtc cgc agt tat tgc tat ttg      1669
Asn Met Glu Ala Ala Asn Leu Ala Glu Val Arg Ser Tyr Cys Tyr Leu
        240                 245                 250

gct acc gtc agc gat ctc tcc acc aaa gct gcg tgc ccg acc atg gga      1717
Ala Thr Val Ser Asp Leu Ser Thr Lys Ala Ala Cys Pro Thr Met Gly
    255                 260                 265

gaa gct cac aat gac aaa cgt gct gac cca gct ttt gtg tgc aga caa      1765
Glu Ala His Asn Asp Lys Arg Ala Asp Pro Ala Phe Val Cys Arg Gln
270                 275                 280                 285

gga gtg gtg gac agg ggc tgg ggc aac ggt tgc gga cta ttt ggc aaa      1813
Gly Val Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys
                290                 295                 300
```

```
gga agc att gac aca tgc gcc aaa ttt gcc tgc tct acc aag gca ata      1861
Gly Ser Ile Asp Thr Cys Ala Lys Phe Ala Cys Ser Thr Lys Ala Ile
            305                 310                 315

gga aga acc atc ttg aaa gag aat atc aag tac gaa gtg gcc att ttt      1909
Gly Arg Thr Ile Leu Lys Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe
        320                 325                 330

gtc cat gga cca act act gtg gag tcg cac gga aac tac tcc aca cag      1957
Val His Gly Pro Thr Thr Val Glu Ser His Gly Asn Tyr Ser Thr Gln
    335                 340                 345

gtt gga gcc act cag gca ggg aga ttc agc atc act cct gcg gcg cct      2005
Val Gly Ala Thr Gln Ala Gly Arg Phe Ser Ile Thr Pro Ala Ala Pro
350                 355                 360                 365

tca tac aca cta aag ctt gga gaa tat gga gag gtg aca gtg gac tgt      2053
Ser Tyr Thr Leu Lys Leu Gly Glu Tyr Gly Glu Val Thr Val Asp Cys
                370                 375                 380

gaa cca cgg tca ggg att gac acc aat gca tac tac gtg atg act gtt      2101
Glu Pro Arg Ser Gly Ile Asp Thr Asn Ala Tyr Tyr Val Met Thr Val
            385                 390                 395

gga aca aag acg ttc ttg gtc cat cgt gag tgg ttc atg gac ctc aac      2149
Gly Thr Lys Thr Phe Leu Val His Arg Glu Trp Phe Met Asp Leu Asn
        400                 405                 410

ctc cct tgg agc agt gct gga agt act gtg tgg agg aac aga gag acg      2197
Leu Pro Trp Ser Ser Ala Gly Ser Thr Val Trp Arg Asn Arg Glu Thr
    415                 420                 425

tta atg gag ttt gag gaa cca cac gcc acg aag cag tct gtg ata gca      2245
Leu Met Glu Phe Glu Glu Pro His Ala Thr Lys Gln Ser Val Ile Ala
430                 435                 440                 445

ttg ggc tca caa gag gga gct ctg cat caa gct ttg gct gga gcc att      2293
Leu Gly Ser Gln Glu Gly Ala Leu His Gln Ala Leu Ala Gly Ala Ile
                450                 455                 460

cct gtg gaa ttt tca agc aac act gtc aag ttg acg tcg ggt cat ttg      2341
Pro Val Glu Phe Ser Ser Asn Thr Val Lys Leu Thr Ser Gly His Leu
            465                 470                 475

aag tgt aga gtg aag atg gaa aaa ttg cag ttg aag gga aca acc tat      2389
Lys Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr
        480                 485                 490

ggc gtc tgt tca aag gct ttc aag ttt ctt ggg act ccc gcg gac aca      2437
Gly Val Cys Ser Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr
    495                 500                 505

ggt cac ggc act gtg gtg ttg gaa ttg cag tac act ggc acg gat gga      2485
Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly
510                 515                 520                 525

cct tgc aaa gtt cct atc tcg tca gtg gct tca ttg aac gac cta acg      2533
Pro Cys Lys Val Pro Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr
                530                 535                 540
```

```
cca gtg ggc aga ttg gtc act gtc aac cct ttt gtt tca gtg gcc acg      2581
Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val Ser Val Ala Thr
            545                 550                 555

gcc aac gct aag gtc ctg att gaa ttg gaa cca ccc ttt gga gac tca      2629
Ala Asn Ala Lys Val Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser
        560                 565                 570

tac ata gtg gtg ggc aga gga gaa caa cag atc aat cac cat tgg cac      2677
Tyr Ile Val Val Gly Arg Gly Glu Gln Gln Ile Asn His His Trp His
    575                 580                 585

aag tct gga agc agc att ggc aaa gcc ttt aca acc acc ctc aaa gga      2725
Lys Ser Gly Ser Ser Ile Gly Lys Ala Phe Thr Thr Thr Leu Lys Gly
590                 595                 600                 605

gcg cag aga cta gcc gct cta gga gac aca gct tgg gac ttt gga tca      2773
Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
                610                 615                 620

gtt gga ggg gtg ttc acc tca gtt ggg aag gct gtc cat caa gtg ttc      2821
Val Gly Gly Val Phe Thr Ser Val Gly Lys Ala Val His Gln Val Phe
            625                 630                 635

gga gga gca ttc cgc tca ctg ttc gga ggc atg tcc tgg ata acg caa      2869
Gly Gly Ala Phe Arg Ser Leu Phe Gly Gly Met Ser Trp Ile Thr Gln
        640                 645                 650

gga ttg ctg ggg gct ctc ctg ttg tgg atg ggc atc aat gct cgt gat      2917
Gly Leu Leu Gly Ala Leu Leu Leu Trp Met Gly Ile Asn Ala Arg Asp
    655                 660                 665

agg tcc ata gct ctc acg ttt ctc gca gtt gga gga gtt ctg ctc ttc      2965
Arg Ser Ile Ala Leu Thr Phe Leu Ala Val Gly Gly Val Leu Leu Phe
670                 675                 680                 685

ctc tcc gtg aac gtg cac gcc t gaaggcggcc gctcgagcat gcatctagag      3017
Leu Ser Val Asn Val His Ala
                690

ggccctattc tatagtgtca cctaaatgct agagctcgct gatcagcctc gactgtgcct   3077
tctagttgcc agccatctgt tgtttgcccc tccccccgtgc cttccttgac cctggaaggt   3137
gccactccca ctgtcctttc ctaataaaat gaggaaattg catcgcattg tctgagtagg   3197
tgtcattcta ttctgggggg tggggtgggg caggacagca agggggagga ttgggaagac   3257
aatagcaggc atgctgggga tgcggtgggc tctatggctt ctgaggcggga aagaaccagc   3317
tgcattaatg aatcggccaa cgcgcggggga gaggcggttt gcgtattggg cgctcttccg   3377
cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc   3437
actcaaaggc ggtaatacgg ttatccacag aatcaggggga taacgcagga aagaacatgt   3497
gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc   3557
ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa   3617
acccgacagg actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc   3677
ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg   3737
cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc   3797
tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc   3857
gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca   3917
ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact   3977
acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca gttaccttcg   4037
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt   4097
ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct   4157
tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga   4217
gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa   4277
```

tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac 4337
ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga 4397
taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc 4457
cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca 4517
gaagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta 4577
gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg 4637
tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc 4697
gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt cctccgatcg 4757
ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt 4817
ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac tcaaccaagt 4877
cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca atacgggata 4937
ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc 4997
gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac 5057
ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa 5117
ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct 5177
tcctttttca atattattga agcatttatc agggttattg tctcatgagc ggatacatat 5237
ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc 5297
cacctgacgt c 5308

<210> 16
<211> 692
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<223> pCBHN

<400> 16

Met	Gly	Lys	Arg	Ser	Ala	Gly	Ser	Ile	Met	Trp	Leu	Ala	Ser	Leu	Ala
1				5				10					15		
Val	Val	Ile	Ala	Cys	Ala	Gly	Ala	Val	Thr	Leu	Ser	Asn	Phe	Gln	Gly
			20					25					30		
Lys	Val	Met	Met	Thr	Val	Asn	Ala	Thr	Asp	Val	Thr	Asp	Val	Ile	Thr
		35					40					45			
Ile	Pro	Thr	Ala	Ala	Gly	Lys	Asn	Leu	Cys	Ile	Val	Arg	Ala	Met	Asp
	50					55					60				
Val	Gly	Tyr	Met	Cys	Asp	Asp	Thr	Ile	Thr	Tyr	Glu	Cys	Pro	Val	Leu
65					70					75					80
Ser	Ala	Gly	Asn	Asp	Pro	Glu	Asp	Ile	Asp	Cys	Trp	Cys	Thr	Lys	Ser
				85					90					95	
Ala	Val	Tyr	Val	Arg	Tyr	Gly	Arg	Cys	Thr	Lys	Thr	Arg	His	Ser	Arg
			100					105					110		
Arg	Ser	Arg	Arg	Ser	Leu	Thr	Val	Gln	Thr	His	Gly	Glu	Ser	Thr	Leu
		115					120					125			
Ala	Asn	Lys	Lys	Gly	Ala	Trp	Met	Asp	Ser	Thr	Lys	Ala	Thr	Arg	Tyr
	130					135					140				
Leu	Val	Lys	Thr	Glu	Ser	Trp	Ile	Leu	Arg	Asn	Pro	Gly	Tyr	Ala	Leu
145					150					155					160
Val	Ala	Ala	Val	Ile	Gly	Trp	Met	Leu	Gly	Ser	Asn	Thr	Met	Gln	Arg
				165					170					175	
Val	Val	Phe	Val	Val	Leu	Leu	Leu	Leu	Val	Ala	Pro	Ala	Tyr	Ser	Phe
			180					185					190		
Asn	Cys	Leu	Gly	Met	Ser	Asn	Arg	Asp	Phe	Leu	Glu	Gly	Val	Ser	Gly
		195					200					205			
Ala	Thr	Trp	Val	Asp	Leu	Val	Leu	Glu	Gly	Asp	Ser	Cys	Val	Thr	Ile
	210					215					220				
Met	Ser	Lys	Asp	Lys	Pro	Thr	Ile	Asp	Val	Lys	Met	Met	Asn	Met	Glu
225					230					235					240

Ala Ala Asn Leu Ala Glu Val Arg Ser Tyr Cys Tyr Leu Ala Thr Val
245 250 255
Ser Asp Leu Ser Thr Lys Ala Ala Cys Pro Thr Met Gly Glu Ala His
260 265 270
Asn Asp Lys Arg Ala Asp Pro Ala Phe Val Cys Arg Gln Gly Val Val
275 280 285
Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile
290 295 300
Asp Thr Cys Ala Lys Phe Ala Cys Ser Thr Lys Ala Ile Gly Arg Thr
305 310 315 320
Ile Leu Lys Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe Val His Gly
325 330 335
Pro Thr Thr Val Glu Ser His Gly Asn Tyr Ser Thr Gln Val Gly Ala
340 345 350
Thr Gln Ala Gly Arg Phe Ser Ile Thr Pro Ala Ala Pro Ser Tyr Thr
355 360 365
Leu Lys Leu Gly Glu Tyr Gly Glu Val Thr Val Asp Cys Glu Pro Arg
370 375 380
Ser Gly Ile Asp Thr Asn Ala Tyr Tyr Val Met Thr Val Gly Thr Lys
385 390 395 400
Thr Phe Leu Val His Arg Glu Trp Phe Met Asp Leu Asn Leu Pro Trp
405 410 415
Ser Ser Ala Gly Ser Thr Val Trp Arg Asn Arg Glu Thr Leu Met Glu
420 425 430
Phe Glu Glu Pro His Ala Thr Lys Gln Ser Val Ile Ala Leu Gly Ser
435 440 445
Gln Glu Gly Ala Leu His Gln Ala Leu Ala Gly Ala Ile Pro Val Glu
450 455 460
Phe Ser Ser Asn Thr Val Lys Leu Thr Ser Gly His Leu Lys Cys Arg
465 470 475 480
Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys
485 490 495
Ser Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr Gly His Gly
500 505 510
Thr Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly Pro Cys Lys
515 520 525
Val Pro Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr Pro Val Gly
530 535 540
Arg Leu Val Thr Val Asn Pro Phe Val Ser Val Ala Thr Ala Asn Ala
545 550 555 560
Lys Val Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val
565 570 575
Val Gly Arg Gly Glu Gln Gln Ile Asn His His Trp His Lys Ser Gly
580 585 590
Ser Ser Ile Gly Lys Ala Phe Thr Thr Thr Leu Lys Gly Ala Gln Arg
595 600 605
Leu Ala Ala Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Val Gly Gly
610 615 620
Val Phe Thr Ser Val Gly Lys Ala Val His Gln Val Phe Gly Gly Ala
625 630 635 640
Phe Arg Ser Leu Phe Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Leu
645 650 655
Gly Ala Leu Leu Leu Trp Met Gly Ile Asn Ala Arg Asp Arg Ser Ile
660 665 670
Ala Leu Thr Phe Leu Ala Val Gly Gly Val Leu Leu Phe Leu Ser Val
675 680 685
Asn Val His Ala
690

<210> 17
<211> 5334

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (916)...(3007)

<221> misc_feature
<222> (1)...(5334)
<223> pCBJE 1-14

<400> 17

```
gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg      60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg     120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc     180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt     240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata     300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc     360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc     420
attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt     480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt     540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca     600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg     660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc     720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg     780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca     840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc     900
tctagagccg ccgcc atg ggc aga aag caa aac aaa aga gga gga aat gaa       951
                 Met Gly Arg Lys Gln Asn Lys Arg Gly Gly Asn Glu
                  1               5                  10

ggc tca atc atg tgg ctc gcg agc ttg gca gtt gtc ata gct tgt gcg       999
Gly Ser Ile Met Trp Leu Ala Ser Leu Ala Val Val Ile Ala Cys Ala
         15                  20                  25

gga gcc atg aag ttg tcg aat ttc cag ggg aag ctt ttg atg acc atc      1047
Gly Ala Met Lys Leu Ser Asn Phe Gln Gly Lys Leu Leu Met Thr Ile
     30                  35                  40

aac aac acg gac att gca gac gtt atc gtg att ccc acc tca aaa gga      1095
Asn Asn Thr Asp Ile Ala Asp Val Ile Val Ile Pro Thr Ser Lys Gly
 45                  50                  55                  60

gag aac aga tgc tgg gtc cgg gca atc gac gtc ggc tac atg tgt gag      1143
Glu Asn Arg Cys Trp Val Arg Ala Ile Asp Val Gly Tyr Met Cys Glu
                 65                  70                  75

gac act atc acg tac gaa tgt cct aag ctt acc atg ggc aat gat cca      1191
Asp Thr Ile Thr Tyr Glu Cys Pro Lys Leu Thr Met Gly Asn Asp Pro
             80                  85                  90

gag gat gtg gat tgc tgg tgt gac aac caa gaa gtc tac gtc caa tat      1239
Glu Asp Val Asp Cys Trp Cys Asp Asn Gln Glu Val Tyr Val Gln Tyr
         95                  100                 105
```

```
gga cgg tgc acg cgg acc agg cat tcc aag cga agc agg aga tcc gtg     1287
Gly Arg Cys Thr Arg Thr Arg His Ser Lys Arg Ser Arg Arg Ser Val
    110                 115                 120

tcg gtc caa aca cat ggg gag agt tca cta gtg aat aaa aaa gag gct     1335
Ser Val Gln Thr His Gly Glu Ser Ser Leu Val Asn Lys Lys Glu Ala
125                 130                 135                 140

tgg ctg gat tca acg aaa gcc aca cga tat ctc atg aaa act gag aac     1383
Trp Leu Asp Ser Thr Lys Ala Thr Arg Tyr Leu Met Lys Thr Glu Asn
                145                 150                 155

tgg atc ata agg aat cct ggc tat gct ttc ctg gcg gcg gta ctt ggc     1431
Trp Ile Ile Arg Asn Pro Gly Tyr Ala Phe Leu Ala Ala Val Leu Gly
            160                 165                 170

tgg atg ctt ggc agt aac aac ggt caa cgc gtg gta ttt acc atc ctc     1479
Trp Met Leu Gly Ser Asn Asn Gly Gln Arg Val Val Phe Thr Ile Leu
        175                 180                 185

ctg ctg ttg gtc gct ccg gct tac agt ttt aat tgt ctg gga atg ggc     1527
Leu Leu Leu Val Ala Pro Ala Tyr Ser Phe Asn Cys Leu Gly Met Gly
    190                 195                 200

aat cgt gac ttc ata gaa gga gcc agt gga gcc act tgg gtg gac ttg     1575
Asn Arg Asp Phe Ile Glu Gly Ala Ser Gly Ala Thr Trp Val Asp Leu
205                 210                 215                 220

gtg ctg gaa gga gat agc tgc ttg aca atc atg gca aac gac aaa cca     1623
Val Leu Glu Gly Asp Ser Cys Leu Thr Ile Met Ala Asn Asp Lys Pro
                225                 230                 235

aca ttg gac gtc cgc atg att aac atc gaa gct agc caa ctt gct gag     1671
Thr Leu Asp Val Arg Met Ile Asn Ile Glu Ala Ser Gln Leu Ala Glu
            240                 245                 250

gtc aga agt tac tgc tat cat gct tca gtc act gac atc tcg acg gtg     1719
Val Arg Ser Tyr Cys Tyr His Ala Ser Val Thr Asp Ile Ser Thr Val
        255                 260                 265

gct cgg tgc ccc acg act gga gaa gcc cac aac gag aag cga gct gat     1767
Ala Arg Cys Pro Thr Thr Gly Glu Ala His Asn Glu Lys Arg Ala Asp
    270                 275                 280

agt agc tat gtg tgc aaa caa ggc ttc act gac cgt ggg tgg ggc aac     1815
Ser Ser Tyr Val Cys Lys Gln Gly Phe Thr Asp Arg Gly Trp Gly Asn
285                 290                 295                 300

gga tgt gga ctt ttc ggg aag gga agc att gac aca tgt gca aaa ttc     1863
Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Asp Thr Cys Ala Lys Phe
                305                 310                 315

tcc tgc acc agt aaa gcg att ggg aga aca atc cag cca gaa aac atc     1911
Ser Cys Thr Ser Lys Ala Ile Gly Arg Thr Ile Gln Pro Glu Asn Ile
            320                 325                 330

aaa tac gaa gtt ggc att ttt gtg cat gga acc acc act tcg gaa aac     1959
Lys Tyr Glu Val Gly Ile Phe Val His Gly Thr Thr Thr Ser Glu Asn
        335                 340                 345
```

```
cat ggg aat tat tca gcg caa gtt ggg gcg tcc cag gcg gca aag ttt     2007
His Gly Asn Tyr Ser Ala Gln Val Gly Ala Ser Gln Ala Ala Lys Phe
    350                 355                 360

aca gta aca ccc aat gct cct tcg ata acc ctc aaa ctt ggt gac tac     2055
Thr Val Thr Pro Asn Ala Pro Ser Ile Thr Leu Lys Leu Gly Asp Tyr
365                 370                 375                 380

gga gaa gtc aca ctg gac tgt gag cca agg agt gga ctg aac act gaa     2103
Gly Glu Val Thr Leu Asp Cys Glu Pro Arg Ser Gly Leu Asn Thr Glu
                385                 390                 395

gcg ttt tac gtc atg aac gtg ggg tca aag tca ttt ctg gtc cat agg     2151
Ala Phe Tyr Val Met Thr Val Gly Ser Lys Ser Phe Leu Val His Arg
            400                 405                 410

gag tgg ttt cat gac ctc gct ctc ccc tgg acg tcc cct tcg agc aca     2199
Glu Trp Phe His Asp Leu Ala Leu Pro Trp Thr Ser Pro Ser Ser Thr
        415                 420                 425

gcg tgg aga aac aga gaa ctc ctc atg gaa ttt gaa gag gcg cac gcc     2247
Ala Trp Arg Asn Arg Glu Leu Leu Met Glu Phe Glu Glu Ala His Ala
    430                 435                 440

aca aaa cag tcc gtt gtt gct ctt ggg tca cag gaa gga ggc ctc cat     2295
Thr Lys Gln Ser Val Val Ala Leu Gly Ser Gln Glu Gly Gly Leu His
445                 450                 455                 460

cag gcg ttg gca gga gcc atc gtg gtg gag tac tca agc tca gtg aag     2343
Gln Ala Leu Ala Gly Ala Ile Val Val Glu Tyr Ser Ser Ser Val Lys
                465                 470                 475

tta aca tca ggc cac ctg aaa tgt agg ctg aaa atg gac aaa ctg gct     2391
Leu Thr Ser Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Ala
            480                 485                 490

ctg aaa ggc aca acc tat ggc atg tgt aca gaa aaa ttc tcg ttc gcg     2439
Leu Lys Gly Thr Thr Tyr Gly Met Cys Thr Glu Lys Phe Ser Phe Ala
        495                 500                 505

aaa aat ccg gcg gac act ggt cac gga aca gtt gtc att gaa ctc tcc     2487
Lys Asn Pro Ala Asp Thr Gly His Gly Thr Val Val Ile Glu Leu Ser
    510                 515                 520

tac tct ggg agt gat ggc ccc tgc aaa att ccg att gct tcc gtt gcg     2535
Tyr Ser Gly Ser Asp Gly Pro Cys Lys Ile Pro Ile Ala Ser Val Ala
525                 530                 535                 540

agc ctc aat gac atg acc ccc gtt ggg cgg ctg gtg aca gtg aac ccc     2583
Ser Leu Asn Asp Met Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro
                545                 550                 555

ttc gtc gcg act tcc agt gcc agc tca aag gtg ctg gtc gag atg gaa     2631
Phe Val Ala Thr Ser Ser Ala Ser Ser Lys Val Leu Val Glu Met Glu
            560                 565                 570

ccc ccc ttc gga gac tcc tac atc gta gtt gga agg gga gac aag cag     2679
Pro Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Asp Lys Gln
        575                 580                 585
```

```
atc aac cac cat tgg cac aaa gct gga agc acg ctg ggc aag gcc ttt     2727
Ile Asn His His Trp His Lys Ala Gly Ser Thr Leu Gly Lys Ala Phe
    590                 595                 600

tca aca act ttg aag gga gct caa aga ctg gca gcg ttg ggc gac aca     2775
Ser Thr Thr Leu Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr
605                 610                 615                 620

gcc tgg gac ttt ggc tct att gga ggg gtc ttc aac tcc ata gga aaa     2823
Ala Trp Asp Phe Gly Ser Ile Gly Gly Val Phe Asn Ser Ile Gly Lys
                625                 630                 635

gcc gtt cac caa gtg ttt ggt ggt gcc ttc aga aca ctc ttt ggg gga     2871
Ala Val His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly
            640                 645                 650

atg tct tgg atc aca caa ggg cta atg ggt gcc cta ctg ctc tgg atg     2919
Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met
        655                 660                 665

ggc gtc aac gca cga gac cga tca att gct ttg gcc ttc tta gcc aca     2967
Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr
    670                 675                 680

ggg ggt gtg ctc gtg ttc tta gcg acc aat gtg cat gct t aattagtttg    3017
Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala
685                 690                 695

agcggccgct cgagcatgca tctagagggc cctattctat agtgtcacct aaatgctaga   3077
gctcgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt ttgcccctcc   3137
cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tccttttcta ataaaatgag   3197
gaaattgcat cgcattgtct gagtaggtgt cattctattc tgggggggtgg ggtggggcag  3257
gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct   3317
atggcttctg aggcggaaag aaccagctgc attaatgaat cggccaacgc gcggggagag   3377
gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg cgctcggtcg   3437
ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta tccacagaat   3497
caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta   3557
aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa   3617
atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc   3677
ccccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt  3737
ccgcctttct cccttcgggg agcgtggcgc tttctcatag ctcacgctgt aggtatctca   3797
gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccgc gttcagcccg   3857
accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat   3917
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta   3977
cagagttctt gaagtggtgg cctaactacg gctacactag aagaacagta tttggtatct   4037
gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac   4097
aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa   4157
aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa   4217
actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt   4277
taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact tggtctgaca   4337
gttaccaatg cttaatcagt gaggcacCta tctcagcgat ctgtctattt cgttcatcca   4397
tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta ccatctggcc   4457
ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta tcagcaataa   4517
accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc   4577
agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca   4637
acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat   4697
tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag   4757
cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac   4817
tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt   4877
ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg cgaccgagtt   4937
gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc   4997
```

tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat 5057
ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt actttcacca 5117
gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaagggа ataagggcga 5177
cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg 5237
gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg 5297
ttccgcgcac atttccccga aaagtgccac ctgacgt 5334

<210> 18
<211> 697
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<223> pCBJE 1-14

<400> 18

Met	Gly	Arg	Lys	Gln	Asn	Lys	Arg	Gly	Gly	Asn	Glu	Gly	Ser	Ile	Met
1				5					10					15	
Trp	Leu	Ala	Ser	Leu	Ala	Val	Val	Ile	Ala	Cys	Ala	Gly	Ala	Met	Lys
			20					25					30		
Leu	Ser	Asn	Phe	Gln	Gly	Lys	Leu	Leu	Met	Thr	Ile	Asn	Asn	Thr	Asp
		35					40					45			
Ile	Ala	Asp	Val	Ile	Val	Ile	Pro	Thr	Ser	Lys	Gly	Glu	Asn	Arg	Cys
	50					55					60				
Trp	Val	Arg	Ala	Ile	Asp	Val	Gly	Tyr	Met	Cys	Glu	Asp	Thr	Ile	Thr
65					70					75					80
Tyr	Glu	Cys	Pro	Lys	Leu	Thr	Met	Gly	Asn	Asp	Pro	Glu	Asp	Val	Asp
				85					90					95	
Cys	Trp	Cys	Asp	Asn	Gln	Glu	Val	Tyr	Val	Gln	Tyr	Gly	Arg	Cys	Thr
			100					105					110		
Arg	Thr	Arg	His	Ser	Lys	Arg	Ser	Arg	Arg	Ser	Val	Ser	Val	Gln	Thr
		115					120					125			
His	Gly	Glu	Ser	Ser	Leu	Val	Asn	Lys	Lys	Glu	Ala	Trp	Leu	Asp	Ser
	130					135					140				
Thr	Lys	Ala	Thr	Arg	Tyr	Leu	Met	Lys	Thr	Glu	Asn	Trp	Ile	Ile	Arg
145					150					155					160
Asn	Pro	Gly	Tyr	Ala	Phe	Leu	Ala	Ala	Val	Leu	Gly	Trp	Met	Leu	Gly
				165					170					175	
Ser	Asn	Asn	Gly	Gln	Arg	Val	Val	Phe	Thr	Ile	Leu	Leu	Leu	Leu	Val
			180					185					190		
Ala	Pro	Ala	Tyr	Ser	Phe	Asn	Cys	Leu	Gly	Met	Gly	Asn	Arg	Asp	Phe
		195					200					205			
Ile	Glu	Gly	Ala	Ser	Gly	Ala	Thr	Trp	Val	Asp	Leu	Val	Leu	Glu	Gly
	210					215					220				
Asp	Ser	Cys	Leu	Thr	Ile	Met	Ala	Asn	Asp	Lys	Pro	Thr	Leu	Asp	Val
225					230					235					240
Arg	Met	Ile	Asn	Ile	Glu	Ala	Ser	Gln	Leu	Ala	Glu	Val	Arg	Ser	Tyr
				245					250					255	
Cys	Tyr	His	Ala	Ser	Val	Thr	Asp	Ile	Ser	Thr	Val	Ala	Arg	Cys	Pro
			260					265					270		
Thr	Thr	Gly	Glu	Ala	His	Asn	Glu	Lys	Arg	Ala	Asp	Ser	Ser	Tyr	Val
		275					280					285			
Cys	Lys	Gln	Gly	Phe	Thr	Asp	Arg	Gly	Trp	Gly	Asn	Gly	Cys	Gly	Leu
	290					295					300				
Phe	Gly	Lys	Gly	Ser	Ile	Asp	Thr	Cys	Ala	Lys	Phe	Ser	Cys	Thr	Ser
305					310					315					320
Lys	Ala	Ile	Gly	Arg	Thr	Ile	Gln	Pro	Glu	Asn	Ile	Lys	Tyr	Glu	Val
				325					330					335	

```
Gly Ile Phe Val His Gly Thr Thr Thr Ser Glu Asn His Gly Asn Tyr
            340                 345                 350
Ser Ala Gln Val Gly Ala Ser Gln Ala Ala Lys Phe Thr Val Thr Pro
        355                 360                 365
Asn Ala Pro Ser Ile Thr Leu Lys Leu Gly Asp Tyr Gly Glu Val Thr
    370                 375                 380
Leu Asp Cys Glu Pro Arg Ser Gly Leu Asn Thr Glu Ala Phe Tyr Val
385                 390                 395                 400
Met Thr Val Gly Ser Lys Ser Phe Leu Val His Arg Glu Trp Phe His
                405                 410                 415
Asp Leu Ala Leu Pro Trp Thr Ser Pro Ser Ser Thr Ala Trp Arg Asn
            420                 425                 430
Arg Glu Leu Leu Met Glu Phe Glu Glu Ala His Ala Thr Lys Gln Ser
        435                 440                 445
Val Val Ala Leu Gly Ser Gln Glu Gly Gly Leu His Gln Ala Leu Ala
    450                 455                 460
Gly Ala Ile Val Val Glu Tyr Ser Ser Ser Val Lys Leu Thr Ser Gly
465                 470                 475                 480
His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Ala Leu Lys Gly Thr
                485                 490                 495
Thr Tyr Gly Met Cys Thr Glu Lys Phe Ser Phe Ala Lys Asn Pro Ala
            500                 505                 510
Asp Thr Gly His Gly Thr Val Val Ile Glu Leu Ser Tyr Ser Gly Ser
        515                 520                 525
Asp Gly Pro Cys Lys Ile Pro Ile Ala Ser Val Ala Ser Leu Asn Asp
    530                 535                 540
Met Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val Ala Thr
545                 550                 555                 560
Ser Ser Ala Ser Ser Lys Val Leu Val Glu Met Glu Pro Pro Phe Gly
                565                 570                 575
Asp Ser Tyr Ile Val Val Gly Arg Gly Asp Lys Gln Ile Asn His His
            580                 585                 590
Trp His Lys Ala Gly Ser Thr Leu Gly Lys Ala Phe Ser Thr Thr Leu
        595                 600                 605
Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr Ala Trp Asp Phe
    610                 615                 620
Gly Ser Ile Gly Gly Val Phe Asn Ser Ile Gly Lys Ala Val His Gln
625                 630                 635                 640
Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly Met Ser Trp Ile
                645                 650                 655
Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met Gly Val Asn Ala
            660                 665                 670
Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr Gly Gly Val Leu
        675                 680                 685
Val Phe Leu Ala Thr Asn Val His Ala
    690                 695
```

<210> 19
<211> 5283
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (910)...(2965)

<400> 19
gacggatcgg gagatctccc gatccccctat ggtcgactct cagtacaatc tgctctgatg 60

```
cagcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg    120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc    180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt    240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata    300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc    360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc    420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt    480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt    540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca    600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg    660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc    720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg    780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacccа    840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc    900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc    951
          Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser
            1               5                  10

ttg gca gtt gtc ata gct ggt aca agc gct acc acc atc cac cgg gac      999
Leu Ala Val Val Ile Ala Gly Thr Ser Ala Thr Thr Ile His Arg Asp
 15                  20                  25                  30

agg gaa gga tac atg gtt atg cgg gcc agt gga agg gac gct gca agc     1047
Arg Glu Gly Tyr Met Val Met Arg Ala Ser Gly Arg Asp Ala Ala Ser
                35                  40                  45

cag gtc agg gta caa aac gga acg tgc gtc atc ctg gca aca gac atg     1095
Gln Val Arg Val Gln Asn Gly Thr Cys Val Ile Leu Ala Thr Asp Met
            50                  55                  60

gga gag tgg tgt gaa gat tca atc acc tac tct tgc gtc acg att gac     1143
Gly Glu Trp Cys Glu Asp Ser Ile Thr Tyr Ser Cys Val Thr Ile Asp
        65                  70                  75

cag gag gaa gaa ccc gtt gac gtg gac tgc ttc tgc cga ggt gtt gat     1191
Gln Glu Glu Glu Pro Val Asp Val Asp Cys Phe Cys Arg Gly Val Asp
    80                  85                  90

agg gtt aag tta gag tat gga cgc tgt gga agg caa gct gga tct agg     1239
Arg Val Lys Leu Glu Tyr Gly Arg Cys Gly Arg Gln Ala Gly Ser Arg
 95                  100                 105                 110

ggg aaa agg tct gtg gtc att cca aca cat gca caa aaa gac atg gtc     1287
Gly Lys Arg Ser Val Val Ile Pro Thr His Ala Gln Lys Asp Met Val
                115                 120                 125

ggg cga ggt cat gca tgg ctt aaa ggt gac aat att cga gat cat gtc     1335
Gly Arg Gly His Ala Trp Leu Lys Gly Asp Asn Ile Arg Asp His Val
            130                 135                 140

acc cga gtc gag ggc tgg atg tgg aag aac aag ctt cta act gcc gcc     1383
Thr Arg Val Glu Gly Trp Met Trp Lys Asn Lys Leu Leu Thr Ala Ala
        145                 150                 155

att gtg gcc ttg gct tgg ctc atg gtt gat agt tgg atg gcc aga gtg     1431
Ile Val Ala Leu Ala Trp Leu Met Val Asp Ser Trp Met Ala Arg Val
    160                 165                 170

act gtc atc ctc ttg gcg ttg agt cta ggg cca gtg tac gcc acg agg     1479
Thr Val Ile Leu Leu Ala Leu Ser Leu Gly Pro Val Tyr Ala Thr Arg
175                 180                 185                 190
```

```
tgc acg cat ctt gag aac aga gat ttt gtg aca gga act caa ggg acc      1527
Cys Thr His Leu Glu Asn Arg Asp Phe Val Thr Gly Thr Gln Gly Thr
                195                 200                 205

acc aga gtg tcc cta gtt ttg gaa ctt gga ggc tgc gtg acc atc aca      1575
Thr Arg Val Ser Leu Val Leu Glu Leu Gly Gly Cys Val Thr Ile Thr
            210                 215                 220

gct gag ggc aag cca tcc att gat gta tgg ctc gaa gac att ttt cag      1623
Ala Glu Gly Lys Pro Ser Ile Asp Val Trp Leu Glu Asp Ile Phe Gln
        225                 230                 235

gaa agc ccg gct gaa acc aga gaa tac tgc ctg cac gcc aaa ttg acc      1671
Glu Ser Pro Ala Glu Thr Arg Glu Tyr Cys Leu His Ala Lys Leu Thr
    240                 245                 250

aac aca aaa gtg gag gct cgc tgt cca acc act gga ccg gcg aca ctt      1719
Asn Thr Lys Val Glu Ala Arg Cys Pro Thr Thr Gly Pro Ala Thr Leu
255                 260                 265                 270

ccg gag gag cat cag gct aat atg gtg tgc aag aga gac caa agc gac      1767
Pro Glu Glu His Gln Ala Asn Met Val Cys Lys Arg Asp Gln Ser Asp
                275                 280                 285

cgt gga tgg gga aac cac tgc ggg ttt ttt ggg aag ggc agt ata gtg      1815
Arg Gly Trp Gly Asn His Cys Gly Phe Phe Gly Lys Gly Ser Ile Val
            290                 295                 300

gct tgt gca aag ttt gaa tgc gag gaa gca aaa aaa gct gtg ggc cac      1863
Ala Cys Ala Lys Phe Glu Cys Glu Glu Ala Lys Lys Ala Val Gly His
        305                 310                 315

gtc tat gac tcc aca aag atc acg tat gtt gtc aag gtt gag ccc cac      1911
Val Tyr Asp Ser Thr Lys Ile Thr Tyr Val Val Lys Val Glu Pro His
    320                 325                 330

aca ggg gat tac ttg gct gca aat gag acc aat tca aac agg aaa tca      1959
Thr Gly Asp Tyr Leu Ala Ala Asn Glu Thr Asn Ser Asn Arg Lys Ser
335                 340                 345                 350

gca cag ttt acg gtg gca tcc gag aaa gtg atc ctg cgg ctc ggc gac      2007
Ala Gln Phe Thr Val Ala Ser Glu Lys Val Ile Leu Arg Leu Gly Asp
                355                 360                 365

tat gga gat gtg tcg ctg acg tgt aaa gtg gca agt ggg att gat gtc      2055
Tyr Gly Asp Val Ser Leu Thr Cys Lys Val Ala Ser Gly Ile Asp Val
            370                 375                 380

gcc caa act gtg gtg atg tca ctc gac agc agc aag gac cac ctg cct      2103
Ala Gln Thr Val Val Met Ser Leu Asp Ser Ser Lys Asp His Leu Pro
        385                 390                 395

tct gca tgg caa gtg cac cgt gac tgg ttt gag gac ttg gcg ctg ccc      2151
Ser Ala Trp Gln Val His Arg Asp Trp Phe Glu Asp Leu Ala Leu Pro
    400                 405                 410

tgg aaa cac aag gac aac caa gat tgg aac agt gtg gag aaa ctt gtg      2199
Trp Lys His Lys Asp Asn Gln Asp Trp Asn Ser Val Glu Lys Leu Val
415                 420                 425                 430
```

```
gaa ttt gga cca cca cat gct gtg aaa atg gat gtt ttc aat ctg ggg     2247
Glu Phe Gly Pro Pro His Ala Val Lys Met Asp Val Phe Asn Leu Gly
                435                 440                 445

gac cag acg gct gtg ctg ctc aaa tca ctg gca gga gtt ccg ctg gcc     2295
Asp Gln Thr Ala Val Leu Leu Lys Ser Leu Ala Gly Val Pro Leu Ala
            450                 455                 460

agt gtg gag ggc cag aaa tac cac ctg aaa agc ggc cat gtt act tgt     2343
Ser Val Glu Gly Gln Lys Tyr His Leu Lys Ser Gly His Val Thr Cys
        465                 470                 475

gat gtg gga ctg gaa aag ctg aaa ctg aaa ggc aca acc tac tcc atg     2391
Asp Val Gly Leu Glu Lys Leu Lys Leu Lys Gly Thr Thr Tyr Ser Met
    480                 485                 490

tgt gac aaa gca aag ttc aaa tgg aag aga gtt cct gtg gac agc ggc     2439
Cys Asp Lys Ala Lys Phe Lys Trp Lys Arg Val Pro Val Asp Ser Gly
495                 500                 505                 510

cat gac aca gta gtc atg gag gta tca tac aca gga agc gac aag cca     2487
His Asp Thr Val Val Met Glu Val Ser Tyr Thr Gly Ser Asp Lys Pro
                515                 520                 525

tgt cgg atc ccg gtg cgg gct gtg gca cat ggt gtc cca gcg gtt aat     2535
Cys Arg Ile Pro Val Arg Ala Val Ala His Gly Val Pro Ala Val Asn
            530                 535                 540

gta gcc atg ctc ata acc ccc aat cca acc att gaa aca aat ggt ggc     2583
Val Ala Met Leu Ile Thr Pro Asn Pro Thr Ile Glu Thr Asn Gly Gly
        545                 550                 555

gga ttc ata gaa atg cag ctg cca cca ggg gat aac atc atc tat gtg     2631
Gly Phe Ile Glu Met Gln Leu Pro Pro Gly Asp Asn Ile Ile Tyr Val
    560                 565                 570

gga gac ctt agc cag cag tgg ttt cag aaa ggc agt acc att ggt aga     2679
Gly Asp Leu Ser Gln Gln Trp Phe Gln Lys Gly Ser Thr Ile Gly Arg
575                 580                 585                 590

atg ttt gaa aaa acc cgc agg gga ttg gaa agg ctc tct gtg gtt gga     2727
Met Phe Glu Lys Thr Arg Arg Gly Leu Glu Arg Leu Ser Val Val Gly
                595                 600                 605

gaa cat gca tgg gac ttt ggc tca gta ggc ggg gta ctg tct tct gtg     2775
Glu His Ala Trp Asp Phe Gly Ser Val Gly Gly Val Leu Ser Ser Val
            610                 615                 620

ggg aag gca atc cac acg gtg ctg ggg gga gct ttc aac acc ctt ttt     2823
Gly Lys Ala Ile His Thr Val Leu Gly Gly Ala Phe Asn Thr Leu Phe
        625                 630                 635

ggg ggg gtt gga ttc atc cct aag atg ctg ctg ggg gtt gct ctg gtc     2871
Gly Gly Val Gly Phe Ile Pro Lys Met Leu Leu Gly Val Ala Leu Val
    640                 645                 650

tgg ttg gga cta aat gcc agg aat cca acg atg tcc atg acg ttt ctt     2919
Trp Leu Gly Leu Asn Ala Arg Asn Pro Thr Met Ser Met Thr Phe Leu
655                 660                 665                 670
```

```
gct gtg ggg gct ttg aca ctg atg atg aca atg gga gtt ggg gca t       2965
Ala Val Gly Ala Leu Thr Leu Met Met Thr Met Gly Val Gly Ala
                675                 680                 685

gagcggccgc tcgagcatgc atctagaggg ccctattcta tagtgtcacc taaatgctag   3025
agctcgctga tcagcctcga ctgtgccttc tagttgccag ccatctgttg tttgcccctc   3085
ccccgtgcct tccttgaccc tggaaggtgc cactcccact gtcctttcct aataaaatga   3145
ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg gggtggggca   3205
ggacagcaag ggggaggatt gggaagacaa tagcaggcat gctggggatg cggtgggctc   3265
tatggcttct gaggcggaaa gaacagctgc attaatgaat cggccaacgc gcggggagag   3325
gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg cgctcggtcg   3385
ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta tccacagaat   3445
caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta   3505
aaaaggccgc gttgctggcg ttttttccata ggctccgccc ccctgacgag catcacaaaa   3565
atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc   3625
cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt   3685
ccgcctttct cccttcggga agcgtggcgc tttctcaatg ctcacgctgt aggtatctca   3745
gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccg   3805
accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat   3865
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta   3925
cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta tttggtatct   3985
gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac   4045
aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa   4105
aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa   4165
actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt   4225
taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact tggtctgaca   4285
gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca   4345
tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta ccatctggcc   4405
ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta tcagcaataa   4465
accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc   4525
agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca   4585
acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat   4645
tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag   4705
cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac   4765
tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt   4825
ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg cgaccgagtt   4885
gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc   4945
tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat   5005
ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt actttcacca   5065
gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataagggcga   5125
cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg   5185
gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg   5245
ttccgcgcac atttccccga aaagtgccac ctgacgtc                           5283
```

<210> 20
<211> 681
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 20

```
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
 1               5                  10                  15
Val Val Ile Ala Gly Thr Ser Ala Val Thr Leu Val Arg Lys Asn Arg
            20                  25                  30
Trp Leu Leu Leu Asn Val Thr Ser Glu Asp Leu Gly Lys Thr Phe Ser
        35                  40                  45
```

```
Val Gly Thr Gly Asn Cys Thr Thr Asn Ile Leu Glu Ala Lys Tyr Trp
    50                  55                  60
Cys Pro Asp Ser Met Glu Tyr Asn Cys Pro Asn Leu Ser Pro Arg Glu
65                  70                  75                  80
Glu Pro Asp Asp Ile Asp Cys Trp Cys Tyr Gly Val Glu Asn Val Arg
                85                  90                  95
Val Ala Tyr Gly Lys Cys Asp Ser Ala Gly Arg Ser Arg Arg Ser Arg
            100                 105                 110
Arg Ala Ile Asp Leu Pro Thr His Glu Asn His Gly Leu Lys Thr Arg
        115                 120                 125
Gln Glu Lys Trp Met Thr Gly Arg Met Gly Glu Arg Gln Leu Gln Lys
    130                 135                 140
Ile Glu Arg Trp Phe Val Arg Asn Pro Phe Phe Ala Val Thr Ala Leu
145                 150                 155                 160
Thr Ile Ala Tyr Leu Val Gly Ser Asn Met Thr Gln Arg Val Val Ile
                165                 170                 175
Ala Leu Leu Val Leu Ala Val Gly Pro Ala Tyr Ser Ala His Cys Ile
            180                 185                 190
Gly Ile Thr Asp Arg Asp Phe Ile Glu Gly Val His Gly Gly Thr Trp
        195                 200                 205
Val Ser Ala Thr Leu Glu Gln Asp Lys Cys Val Thr Val Met Ala Pro
    210                 215                 220
Asp Lys Pro Ser Leu Asp Ile Ser Leu Glu Thr Val Ala Ile Asp Arg
225                 230                 235                 240
Pro Ala Glu Val Arg Lys Val Cys Tyr Asn Ala Val Leu Thr His Val
                245                 250                 255
Lys Ile Asn Asp Lys Cys Pro Ser Thr Gly Glu Ala His Leu Ala Glu
            260                 265                 270
Glu Asn Glu Gly Asp Asn Ala Cys Lys Arg Thr Tyr Ser Asp Arg Gly
        275                 280                 285
Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Val Ala Cys
    290                 295                 300
Ala Lys Phe Thr Cys Ala Lys Ser Met Ser Leu Phe Glu Val Asp Gln
305                 310                 315                 320
Thr Lys Ile Gln Tyr Val Ile Arg Ala Gln Leu His Val Gly Ala Lys
                325                 330                 335
Gln Glu Asn Trp Thr Thr Asp Ile Lys Thr Leu Lys Phe Asp Ala Leu
            340                 345                 350
Ser Gly Ser Gln Glu Val Glu Phe Ile Gly Tyr Gly Lys Ala Thr Leu
        355                 360                 365
Glu Cys Gln Val Gln Thr Ala Val Asp Phe Gly Asn Ser Tyr Ile Ala
    370                 375                 380
Glu Met Glu Thr Glu Ser Trp Ile Val Asp Arg Gln Trp Ala Gln Asp
385                 390                 395                 400
Leu Thr Leu Pro Trp Gln Ser Gly Ser Gly Gly Val Trp Arg Glu Met
                405                 410                 415
His His Leu Val Glu Phe Glu Pro Pro His Ala Ala Thr Ile Arg Val
            420                 425                 430
Leu Ala Leu Gly Asn Gln Glu Gly Ser Leu Lys Thr Ala Leu Thr Gly
        435                 440                 445
Ala Met Arg Val Thr Lys Asp Thr Asn Asp Asn Asn Leu Tyr Lys Leu
    450                 455                 460
His Gly Gly His Val Ser Cys Arg Val Lys Leu Ser Ala Leu Thr Leu
465                 470                 475                 480
Lys Gly Thr Ser Tyr Lys Ile Cys Thr Asp Lys Met Phe Phe Val Lys
                485                 490                 495
Asn Pro Thr Asp Thr Gly His Gly Thr Val Val Met Gln Val Lys Val
            500                 505                 510
Ser Lys Gly Ala Pro Cys Arg Ile Pro Val Ile Val Ala Asp Asp Leu
        515                 520                 525
```

```
Thr Ala Ala Ile Asn Lys Gly Ile Leu Val Thr Val Asn Pro Ile Ala
    530                 535                 540
Ser Thr Asn Asp Asp Glu Val Leu Ile Glu Val Asn Pro Pro Phe Gly
545                 550                 555                 560
Asp Ser Tyr Ile Ile Val Gly Arg Gly Asp Ser Arg Leu Thr Tyr Gln
                565                 570                 575
Trp His Lys Glu Gly Ser Ser Ile Gly Lys Leu Phe Thr Gln Thr Met
            580                 585                 590
Lys Gly Val Glu Arg Leu Ala Val Met Gly Asp Thr Ala Trp Asp Phe
        595                 600                 605
Ser Ser Ala Gly Gly Phe Phe Thr Ser Val Gly Lys Gly Ile His Thr
    610                 615                 620
Val Phe Gly Ser Ala Phe Gln Gly Leu Phe Gly Gly Leu Asn Trp Ile
625                 630                 635                 640
Thr Lys Val Ile Met Gly Ala Val Leu Ile Trp Val Gly Ile Asn Thr
                645                 650                 655
Arg Asn Met Thr Met Ser Met Ser Met Ile Leu Val Gly Val Ile Met
            660                 665                 670
Met Phe Leu Ser Leu Gly Val Gly Ala
        675                 680
```

<210> 21
<211> 5304
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (910)...(2986)

<400> 21

```
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg     60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg    120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc    180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt    240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata    300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc    360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc    420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt    480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt    540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca    600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg    660
actcaagggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc    720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg    780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca    840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc    900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc    951
          Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser
           1               5                  10

ttg gca gtt gtc ata gct ggt aca agc gct ttg cag tta tca acc tat     999
Leu Ala Val Val Ile Ala Gly Thr Ser Ala Leu Gln Leu Ser Thr Tyr
 15                  20                  25                  30

cag ggg aaa gtg tta atg tca atc aac aag act gac gct caa agc gcc    1047
Gln Gly Lys Val Leu Met Ser Ile Asn Lys Thr Asp Ala Gln Ser Ala
                35                  40                  45
```

```
ata aac att cct agt gcc aac gga gca aac act tgc att gtg agg gct     1095
Ile Asn Ile Pro Ser Ala Asn Gly Ala Asn Thr Cys Ile Val Arg Ala
             50                  55                  60

cta gat gtg ggg gtc atg tgc aaa gat gac atc aca tac ctg tgc cca     1143
Leu Asp Val Gly Val Met Cys Lys Asp Asp Ile Thr Tyr Leu Cys Pro
         65                  70                  75

gtg ctt tca gcg gga aat gat ccc gag gac att gac tgt tgg tgt gac     1191
Val Leu Ser Ala Gly Asn Asp Pro Glu Asp Ile Asp Cys Trp Cys Asp
     80                  85                  90

gtc gaa gag gtg tgg gtg cac tac ggc aga tgc acg cgc atg gga cat     1239
Val Glu Glu Val Trp Val His Tyr Gly Arg Cys Thr Arg Met Gly His
 95                 100                 105                 110

tcg agg cgt agc cga cgg tca atc tct gtg cag cat cat gga gat tcc     1287
Ser Arg Arg Ser Arg Arg Ser Ile Ser Val Gln His His Gly Asp Ser
                115                 120                 125

aca ctg gca aca aag aac acg cca tgg ttg gac acc gtg aaa acc acc     1335
Thr Leu Ala Thr Lys Asn Thr Pro Trp Leu Asp Thr Val Lys Thr Thr
            130                 135                 140

aaa tac ttg aca aaa gta gaa aac tgg gtt ttg cgc aat cct gga tat     1383
Lys Tyr Leu Thr Lys Val Glu Asn Trp Val Leu Arg Asn Pro Gly Tyr
        145                 150                 155

gcc cta gtt gcg ctg gcg att gga tgg atg ctc ggt agc aac aac aca     1431
Ala Leu Val Ala Leu Ala Ile Gly Trp Met Leu Gly Ser Asn Asn Thr
    160                 165                 170

cag aga gtg gtt ttt gtg atc atg ctg atg ctg att gct ccg gca tac     1479
Gln Arg Val Val Phe Val Ile Met Leu Met Leu Ile Ala Pro Ala Tyr
175                 180                 185                 190

agc ttc aac tgt ctg gga aca tca aac agg gac ttt gtc gag gga gcc     1527
Ser Phe Asn Cys Leu Gly Thr Ser Asn Arg Asp Phe Val Glu Gly Ala
                195                 200                 205

agt ggg gca aca tgg att gac ttg gta ctt gaa ggg gga agc tgt gtc     1575
Ser Gly Ala Thr Trp Ile Asp Leu Val Leu Glu Gly Gly Ser Cys Val
            210                 215                 220

aca gtg atg gca cca gag aaa cca aca ctg gac ttc aaa gtg atg aag     1623
Thr Val Met Ala Pro Glu Lys Pro Thr Leu Asp Phe Lys Val Met Lys
        225                 230                 235

atg gag gct acc gag tta gcc act gtg cgt gag tat tgt tac gaa gca     1671
Met Glu Ala Thr Glu Leu Ala Thr Val Arg Glu Tyr Cys Tyr Glu Ala
    240                 245                 250

acc ttg gac acg ctg tca aca gtg gca agg tgc ccc aca aca gga gaa     1719
Thr Leu Asp Thr Leu Ser Thr Val Ala Arg Cys Pro Thr Thr Gly Glu
255                 260                 265                 270

gct cac aac acc aaa agg agt gac cca aca ttt gtc tgc aaa aga gat     1767
Ala His Asn Thr Lys Arg Ser Asp Pro Thr Phe Val Cys Lys Arg Asp
                275                 280                 285
```

```
gtt gtg gac cgc gga tgg ggt aac gga tgt ggt ctg ttt gga aaa ggg      1815
Val Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly
            290                 295                 300

agc att gac aca tgc gct aag ttc aca tgc aaa aac aag gca aca ggg      1863
Ser Ile Asp Thr Cys Ala Lys Phe Thr Cys Lys Asn Lys Ala Thr Gly
        305                 310                 315

aag acg atc ttg aga gaa aac atc aag tat gag gtt gca atc ttt gtg      1911
Lys Thr Ile Leu Arg Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe Val
    320                 325                 330

cat ggt tca acg gac tct acg tca cat ggc aat tac tct gag cag att      1959
His Gly Ser Thr Asp Ser Thr Ser His Gly Asn Tyr Ser Glu Gln Ile
335                 340                 345                 350

gga aaa aac caa gcg gct aga ttc acc ata agc ccg caa gca ccg tcc      2007
Gly Lys Asn Gln Ala Ala Arg Phe Thr Ile Ser Pro Gln Ala Pro Ser
                355                 360                 365

ttt acg gcc aac atg ggc gag tat gga aca gtt acc att gat tgt gaa      2055
Phe Thr Ala Asn Met Gly Glu Tyr Gly Thr Val Thr Ile Asp Cys Glu
            370                 375                 380

gca aga tca gga atc aac acg gag gat tat tat gtt ttc act gtc aag      2103
Ala Arg Ser Gly Ile Asn Thr Glu Asp Tyr Tyr Val Phe Thr Val Lys
        385                 390                 395

gag aag tca tgg cta gtg aac agg gac tgg ttt cac gac ttg aac ctt      2151
Glu Lys Ser Trp Leu Val Asn Arg Asp Trp Phe His Asp Leu Asn Leu
    400                 405                 410

cca tgg acg agc cct gcc aca act gat tgg cgc aac aga gaa aca ctg      2199
Pro Trp Thr Ser Pro Ala Thr Thr Asp Trp Arg Asn Arg Glu Thr Leu
415                 420                 425                 430

gtg gaa ttt gag gaa ccg cat gcc acc aag caa act gta gta gcc cta      2247
Val Glu Phe Glu Glu Pro His Ala Thr Lys Gln Thr Val Val Ala Leu
                435                 440                 445

gga tcg caa gaa ggt gcc ctg cac aca gca ttg gct gga gcc att cca      2295
Gly Ser Gln Glu Gly Ala Leu His Thr Ala Leu Ala Gly Ala Ile Pro
            450                 455                 460

gcc act gtt agc agc tca acc cta acc ttg caa tca ggg cat ttg aaa      2343
Ala Thr Val Ser Ser Ser Thr Leu Thr Leu Gln Ser Gly His Leu Lys
        465                 470                 475

tgc aga gct aag ctt gac aag gtc aaa atc aag gga acg aca tat ggc      2391
Cys Arg Ala Lys Leu Asp Lys Val Lys Ile Lys Gly Thr Thr Tyr Gly
    480                 485                 490

atg tgt gac tct gcc ttc acc ttc agc aag aac cca act gac aca ggg      2439
Met Cys Asp Ser Ala Phe Thr Phe Ser Lys Asn Pro Thr Asp Thr Gly
495                 500                 505                 510

cac ggg aca gtg att gtg gaa ctg cag tat act gga agc aac gga ccc      2487
His Gly Thr Val Ile Val Glu Leu Gln Tyr Thr Gly Ser Asn Gly Pro
                515                 520                 525
```

```
tgc cga gtt ccc atc tcc gtg act gca aac ctc atg gat ttg aca ccg      2535
Cys Arg Val Pro Ile Ser Val Thr Ala Asn Leu Met Asp Leu Thr Pro
            530                 535                 540

gtt gga aga ttg gtc acg gtc aat ccc ttt ata agc aca ggg gga gcg      2583
Val Gly Arg Leu Val Thr Val Asn Pro Phe Ile Ser Thr Gly Gly Ala
        545                 550                 555

aac aac aag gtc atg atc gaa gtt gaa cca ccc ttt ggc gat tct tac      2631
Asn Asn Lys Val Met Ile Glu Val Glu Pro Pro Phe Gly Asp Ser Tyr
    560                 565                 570

atc gtc gtc gga aga ggc acc acc cag att aac tac cac tgg cac aaa      2679
Ile Val Val Gly Arg Gly Thr Thr Gln Ile Asn Tyr His Trp His Lys
575                 580                 585                 590

gag gga agc agc att ggg aag gct ttg gcg acc aca tgg aaa gga gcc      2727
Glu Gly Ser Ser Ile Gly Lys Ala Leu Ala Thr Thr Trp Lys Gly Ala
                595                 600                 605

caa cgg cta gcc gtc tta ggg gac aca gcg tgg gac ttt gga tct att      2775
Gln Arg Leu Ala Val Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Ile
            610                 615                 620

gga gga gtt ttc aat tca att ggc aaa gct gtc cac caa gtt ttc gga      2823
Gly Gly Val Phe Asn Ser Ile Gly Lys Ala Val His Gln Val Phe Gly
        625                 630                 635

gga gcg ttc agg act ctg ttc ggg gga atg tcc tgg atc aca cag ggg      2871
Gly Ala Phe Arg Thr Leu Phe Gly Gly Met Ser Trp Ile Thr Gln Gly
    640                 645                 650

cta ctt gga gct ctt ctc ctg tgg atg ggg ttg cag gcc cgc gac agg      2919
Leu Leu Gly Ala Leu Leu Leu Trp Met Gly Leu Gln Ala Arg Asp Arg
655                 660                 665                 670

agc atc tcg ctg act cta ctg gct gtc gga ggg att ctc atc ttt ctg      2967
Ser Ile Ser Leu Thr Leu Leu Ala Val Gly Gly Ile Leu Ile Phe Leu
                675                 680                 685

gca acc agc gtg caa gcc t gagcggccgc tcgagcatgc atctagaggg           3016
Ala Thr Ser Val Gln Ala
            690

ccctattcta tagtgtcacc taaatgctag agctcgctga tcagcctcga ctgtgccttc   3076
tagttgccag ccatctgttg tttgcccctc ccccgtgcct tccttgaccc tggaaggtgc   3136
cactcccact gtcctttcct aataaaatga ggaaattgca tcgcattgtc tgagtaggtg   3196
tcattctatt ctgggggggtg gggtggggca ggacagcaag gggaggatt ggggaagacaa  3256
tagcaggcat gctggggatg cggtgggctc tatggcttct gaggcggaaa gaacagctgc   3316
attaatgaat cggccaacgc gcggggagag gcggtttgcg tattgggcgc tcttccgctt   3376
cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagctcact   3436
caaaggcggt aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag   3496
caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata   3556
ggctccgccc cctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc   3616
cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg   3676
ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccttcgggg agcgtggcgc   3736
tttctcaatg ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc tccaagctgg   3796
gctgtgtgca cgaaccccc gttcagcccg accgctgcgc cttatccggt aactatcgtc   3856
ttgagtccaa cccggtaaga cacgacttat cgccactggc agcagccact ggtaacagga   3916
ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtggtgg cctaactacg   3976
gctacactag aaggacagta tttggtatct gcgctctgct gaagccagtt accttcggaa   4036
```

aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggt ggtttttttg 4096
tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt 4156
ctacggggtc tgacgctcag tggaacgaaa actcacgtta agggattttg gtcatgagat 4216
tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct 4276
aaagtatata tgagtaaact tggtctgaca gttaccaatg cttaatcagt gaggcaccta 4336
tctcagcgat ctgtctattt cgttcatcca tagttgcctg actccccgtc gtgtagataa 4396
ctacgatacg ggagggctta ccatctggcc ccagtgctgc aatgataccg cgagacccac 4456
gctcaccggc tccagattta tcagcaataa accagccagc cggaagggcc gagcgcagaa 4516
gtggtcctgc aactttatcc gcctccatcc agtctattaa ttgttgccgg gaagctagag 4576
taagtagttc gccagttaat agtttgcgca acgttgttgc cattgctaca ggcatcgtgg 4636
tgtcacgctc gtcgtttggt atggcttcat tcagctccgg ttcccaacga tcaaggcgag 4696
ttacatgatc ccccatgttg tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg 4756
tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg cataattctc 4816
ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca accaagtcat 4876
tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata cgggataata 4936
ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa 4996
aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaacccact cgtgcaccca 5056
actgatcttc agcatctttt actttcacca gcgtttctgg gtgagcaaaa acaggaaggc 5116
aaaatgccgc aaaaaaggga ataagggcga cacggaaatg ttgaatactc atactcttcc 5176
tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg 5236
aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac 5296
ctgacgtc 5304

<210> 22
<211> 692
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 22
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
1 5 10 15
Val Val Ile Ala Gly Thr Ser Ala Leu Gln Leu Ser Thr Tyr Gln Gly
20 25 30
Lys Val Leu Met Ser Ile Asn Lys Thr Asp Ala Gln Ser Ala Ile Asn
35 40 45
Ile Pro Ser Ala Asn Gly Ala Asn Thr Cys Ile Val Arg Ala Leu Asp
50 55 60
Val Gly Val Met Cys Lys Asp Asp Ile Thr Tyr Leu Cys Pro Val Leu
65 70 75 80
Ser Ala Gly Asn Asp Pro Glu Asp Ile Asp Cys Trp Cys Asp Val Glu
85 90 95
Glu Val Trp Val His Tyr Gly Arg Cys Thr Arg Met Gly His Ser Arg
100 105 110
Arg Ser Arg Arg Ser Ile Ser Val Gln His His Gly Asp Ser Thr Leu
115 120 125
Ala Thr Lys Asn Thr Pro Trp Leu Asp Thr Val Lys Thr Thr Lys Tyr
130 135 140
Leu Thr Lys Val Glu Asn Trp Val Leu Arg Asn Pro Gly Tyr Ala Leu
145 150 155 160
Val Ala Leu Ala Ile Gly Trp Met Leu Gly Ser Asn Asn Thr Gln Arg
165 170 175
Val Val Phe Val Ile Met Leu Met Leu Ile Ala Pro Ala Tyr Ser Phe
180 185 190
Asn Cys Leu Gly Thr Ser Asn Arg Asp Phe Val Glu Gly Ala Ser Gly
195 200 205
Ala Thr Trp Ile Asp Leu Val Leu Glu Gly Gly Ser Cys Val Thr Val
210 215 220

Met Ala Pro Glu Lys Pro Thr Leu Asp Phe Lys Val Met Lys Met Glu
225 230 235 240
Ala Thr Glu Leu Ala Thr Val Arg Glu Tyr Cys Tyr Glu Ala Thr Leu
245 250 255
Asp Thr Leu Ser Thr Val Ala Arg Cys Pro Thr Thr Gly Glu Ala His
260 265 270
Asn Thr Lys Arg Ser Asp Pro Thr Phe Val Cys Lys Arg Asp Val Val
275 280 285
Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile
290 295 300
Asp Thr Cys Ala Lys Phe Thr Cys Lys Asn Lys Ala Thr Gly Lys Thr
305 310 315 320
Ile Leu Arg Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe Val His Gly
325 330 335
Ser Thr Asp Ser Thr Ser His Gly Asn Tyr Ser Glu Gln Ile Gly Lys
340 345 350
Asn Gln Ala Ala Arg Phe Thr Ile Ser Pro Gln Ala Pro Ser Phe Thr
355 360 365
Ala Asn Met Gly Glu Tyr Gly Thr Val Thr Ile Asp Cys Glu Ala Arg
370 375 380
Ser Gly Ile Asn Thr Glu Asp Tyr Tyr Val Phe Thr Val Lys Glu Lys
385 390 395 400
Ser Trp Leu Val Asn Arg Asp Trp Phe His Asp Leu Asn Leu Pro Trp
405 410 415
Thr Ser Pro Ala Thr Thr Asp Trp Arg Asn Arg Glu Thr Leu Val Glu
420 425 430
Phe Glu Glu Pro His Ala Thr Lys Gln Thr Val Val Ala Leu Gly Ser
435 440 445
Gln Glu Gly Ala Leu His Thr Ala Leu Ala Gly Ala Ile Pro Ala Thr
450 455 460
Val Ser Ser Ser Thr Leu Thr Leu Gln Ser Gly His Leu Lys Cys Arg
465 470 475 480
Ala Lys Leu Asp Lys Val Lys Ile Lys Gly Thr Thr Tyr Gly Met Cys
485 490 495
Asp Ser Ala Phe Thr Phe Ser Lys Asn Pro Thr Asp Thr Gly His Gly
500 505 510
Thr Val Ile Val Glu Leu Gln Tyr Thr Gly Ser Asn Gly Pro Cys Arg
515 520 525
Val Pro Ile Ser Val Thr Ala Asn Leu Met Asp Leu Thr Pro Val Gly
530 535 540
Arg Leu Val Thr Val Asn Pro Phe Ile Ser Thr Gly Gly Ala Asn Asn
545 550 555 560
Lys Val Met Ile Glu Val Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val
565 570 575
Val Gly Arg Gly Thr Thr Gln Ile Asn Tyr His Trp His Lys Glu Gly
580 585 590
Ser Ser Ile Gly Lys Ala Leu Ala Thr Thr Trp Lys Gly Ala Gln Arg
595 600 605
Leu Ala Val Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Ile Gly Gly
610 615 620
Val Phe Asn Ser Ile Gly Lys Ala Val His Gln Val Phe Gly Gly Ala
625 630 635 640
Phe Arg Thr Leu Phe Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Leu
645 650 655
Gly Ala Leu Leu Leu Trp Met Gly Leu Gln Ala Arg Asp Arg Ser Ile
660 665 670
Ser Leu Thr Leu Leu Ala Val Gly Gly Ile Leu Ile Phe Leu Ala Thr
675 680 685
Ser Val Gln Ala
690

<210> 23
<211> 5271
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (910)...(2953)

<400> 23
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc 900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc 951
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser
1 5 10

ttg gca gtt gtc ata gct ggt aca agc gct gtg acc ttg gtg cgg aaa 999
Leu Ala Val Val Ile Ala Gly Thr Ser Ala Val Thr Leu Val Arg Lys
15 20 25 30

aac aga tgg ttg ctc cta aat gtg aca tct gag gac ctc ggg aaa aca 1047
Asn Arg Trp Leu Leu Leu Asn Val Thr Ser Glu Asp Leu Gly Lys Thr
35 40 45

ttc tct gtg ggc aca ggc aac tgc aca aca aac att ttg gaa gcc aag 1095
Phe Ser Val Gly Thr Gly Asn Cys Thr Thr Asn Ile Leu Glu Ala Lys
50 55 60

tac tgg tgc cca gac tca atg gaa tac aac tgt ccc aat ctc agt cca 1143
Tyr Trp Cys Pro Asp Ser Met Glu Tyr Asn Cys Pro Asn Leu Ser Pro
65 70 75

aga gag gag cca gat gac att gat tgc tgg tgc tat ggg gtg gaa aac 1191
Arg Glu Glu Pro Asp Asp Ile Asp Cys Trp Cys Tyr Gly Val Glu Asn
80 85 90

gtt aga gtc gca tat ggt aag tgt gac tca gca ggc agg tct agg agg 1239
Val Arg Val Ala Tyr Gly Lys Cys Asp Ser Ala Gly Arg Ser Arg Arg
95 100 105 110

tca aga agg gcc att gac ttg cct acg cat gaa aac cat ggt ttg aag 1287
Ser Arg Arg Ala Ile Asp Leu Pro Thr His Glu Asn His Gly Leu Lys
115 120 125

```
acc cgg caa gaa aaa tgg atg act gga aga atg ggt gaa agg caa ctc      1335
Thr Arg Gln Glu Lys Trp Met Thr Gly Arg Met Gly Glu Arg Gln Leu
            130                 135                 140

caa aag att gag aga tgg ttc gtg agg aac ccc ttt ttt gca gtg acg      1383
Gln Lys Ile Glu Arg Trp Phe Val Arg Asn Pro Phe Phe Ala Val Thr
        145                 150                 155

gct ctg acc att gcc tac ctt gtg gga agc aac atg acg caa cga gtc      1431
Ala Leu Thr Ile Ala Tyr Leu Val Gly Ser Asn Met Thr Gln Arg Val
    160                 165                 170

gtg att gcc cta ctg gtc ttg gct gtt ggt ccg gcc tac tca gct cac      1479
Val Ile Ala Leu Leu Val Leu Ala Val Gly Pro Ala Tyr Ser Ala His
175                 180                 185                 190

tgc att gga att act gac agg gat ttc att gag ggg gtg cat gga gga      1527
Cys Ile Gly Ile Thr Asp Arg Asp Phe Ile Glu Gly Val His Gly Gly
                195                 200                 205

act tgg gtt tca gct acc ctg gag caa gac aag tgt gtc act gtt atg      1575
Thr Trp Val Ser Ala Thr Leu Glu Gln Asp Lys Cys Val Thr Val Met
            210                 215                 220

gcc cct gac aag cct tca ttg gac atc tca cta gag aca gta gcc att      1623
Ala Pro Asp Lys Pro Ser Leu Asp Ile Ser Leu Glu Thr Val Ala Ile
        225                 230                 235

gat aga cct gct gag gtg agg aaa gtg tgt tac aat gca gtt ctc act      1671
Asp Arg Pro Ala Glu Val Arg Lys Val Cys Tyr Asn Ala Val Leu Thr
    240                 245                 250

cat gtg aag att aat gac aag tgc ccc agc act gga gag gcc cac cta      1719
His Val Lys Ile Asn Asp Lys Cys Pro Ser Thr Gly Glu Ala His Leu
255                 260                 265                 270

gct gaa gag aac gaa ggg gac aat gcg tgc aag cgc act tat tct gat      1767
Ala Glu Glu Asn Glu Gly Asp Asn Ala Cys Lys Arg Thr Tyr Ser Asp
                275                 280                 285

aga ggc tgg ggc aat ggc tgt ggc cta ttt ggg aaa ggg agc att gtg      1815
Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Val
            290                 295                 300

gca tgc gcc aaa ttc act tgt gcc aaa tcc atg agt ttg ttt gag gtt      1863
Ala Cys Ala Lys Phe Thr Cys Ala Lys Ser Met Ser Leu Phe Glu Val
        305                 310                 315

gat cag acc aaa att cag tat gtc atc aga gca caa ttg cat gta ggg      1911
Asp Gln Thr Lys Ile Gln Tyr Val Ile Arg Ala Gln Leu His Val Gly
    320                 325                 330

gcc aag cag gaa aat tgg act acc gac att aag act ctc aag ttt gat      1959
Ala Lys Gln Glu Asn Trp Thr Thr Asp Ile Lys Thr Leu Lys Phe Asp
335                 340                 345                 350

gcc ctg tca ggc tcc cag gaa gtc gag ttc att ggg tat gga aaa gct      2007
Ala Leu Ser Gly Ser Gln Glu Val Glu Phe Ile Gly Tyr Gly Lys Ala
                355                 360                 365
```

```
aca ctg gaa tgc cag gtg caa act gcg gtg gac ttt ggt aac agt tac      2055
Thr Leu Glu Cys Gln Val Gln Thr Ala Val Asp Phe Gly Asn Ser Tyr
            370                 375                 380

atc gct gag atg gaa aca gag agc tgg ata gtg gac aga cag tgg gcc      2103
Ile Ala Glu Met Glu Thr Glu Ser Trp Ile Val Asp Arg Gln Trp Ala
        385                 390                 395

cag gac ttg acc ctg cca tgg cag agt gga agt ggc ggg gtg tgg aga      2151
Gln Asp Leu Thr Leu Pro Trp Gln Ser Gly Ser Gly Gly Val Trp Arg
    400                 405                 410

gag atg cat cat ctt gtc gaa ttt gaa cct ccg cat gcc gcc act atc      2199
Glu Met His His Leu Val Glu Phe Glu Pro Pro His Ala Ala Thr Ile
415                 420                 425                 430

aga gta ctg gcc ctg gga aac cag gaa ggc tcc ttg aaa aca gct ctt      2247
Arg Val Leu Ala Leu Gly Asn Gln Glu Gly Ser Leu Lys Thr Ala Leu
                435                 440                 445

act ggc gca atg agg gtt aca aag gac aca aat gac aac aac ctt tac      2295
Thr Gly Ala Met Arg Val Thr Lys Asp Thr Asn Asp Asn Asn Leu Tyr
            450                 455                 460

aaa cta cat ggt gga cat gtt tct tgc aga gtg aaa ttg tca gct ttg      2343
Lys Leu His Gly Gly His Val Ser Cys Arg Val Lys Leu Ser Ala Leu
        465                 470                 475

aca ctc aag ggg aca tcc tac aaa ata tgc act gac aaa atg ttt ttt      2391
Thr Leu Lys Gly Thr Ser Tyr Lys Ile Cys Thr Asp Lys Met Phe Phe
    480                 485                 490

gtc aag aac cca act gac act ggc cat ggc act gtt gtg atg cag gtg      2439
Val Lys Asn Pro Thr Asp Thr Gly His Gly Thr Val Val Met Gln Val
495                 500                 505                 510

aaa gtg tca aaa gga gcc ccc tgc agg att cca gtg ata gta gct gat      2487
Lys Val Ser Lys Gly Ala Pro Cys Arg Ile Pro Val Ile Val Ala Asp
                515                 520                 525

gat ctt aca gcg gca atc aat aaa ggc att ttg gtt aca gtt aac ccc      2535
Asp Leu Thr Ala Ala Ile Asn Lys Gly Ile Leu Val Thr Val Asn Pro
            530                 535                 540

atc gcc tca acc aat gat gat gaa gtg ctg att gag gtg aac cca cct      2583
Ile Ala Ser Thr Asn Asp Asp Glu Val Leu Ile Glu Val Asn Pro Pro
        545                 550                 555

ttt gga gac agc tac att atc gtt ggg aga gga gat tca cgt ctc act      2631
Phe Gly Asp Ser Tyr Ile Ile Val Gly Arg Gly Asp Ser Arg Leu Thr
    560                 565                 570

tac cag tgg cac aaa gag gga agc tca ata gga aag ttg ttc act cag      2679
Tyr Gln Trp His Lys Glu Gly Ser Ser Ile Gly Lys Leu Phe Thr Gln
575                 580                 585                 590

acc atg aaa ggc gtg gaa cgc ctg gcc gtc atg gga gac acc gcc tgg      2727
Thr Met Lys Gly Val Glu Arg Leu Ala Val Met Gly Asp Thr Ala Trp
                595                 600                 605
```

```
gat ttc agc tcc gct gga ggg ttc ttc act tcg gtt ggg aaa gga att     2775
Asp Phe Ser Ser Ala Gly Gly Phe Phe Thr Ser Val Gly Lys Gly Ile
            610                 615                 620

cat acg gtg ttt ggc tct gcc ttt cag ggg cta ttt ggc ggc ttg aac     2823
His Thr Val Phe Gly Ser Ala Phe Gln Gly Leu Phe Gly Gly Leu Asn
        625                 630                 635

tgg ata aca aag gtc atc atg ggg gcg gta ctt ata tgg gtt ggc atc     2871
Trp Ile Thr Lys Val Ile Met Gly Ala Val Leu Ile Trp Val Gly Ile
    640                 645                 650

aac aca aga aac atg aca atg tcc atg agc atg atc ttg gta gga gtg     2919
Asn Thr Arg Asn Met Thr Met Ser Met Ser Met Ile Leu Val Gly Val
655                 660                 665                 670

atc atg atg ttt ttg tct cta gga gtt ggg gcg t gagcggccgc           2963
Ile Met Met Phe Leu Ser Leu Gly Val Gly Ala
                675                 680

tcgagcatgc atctagaggg ccctatttcta tagtgtcacc taaatgctag agctcgctga   3023
tcaggctgga ctgtgcgttg tagttgccag ccatctgttg tttgcccctc ccccgtgcct   3083
tccttgaccc tggaaggtgc cactcccact gtcctttcct aataaaatga ggaaattgca   3143
tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg gggtgggggca ggacagcaag  3203
ggggaggatt gggaagacaa tagcaggcat gctgggggatg cggtgggctc tatggcttct  3263
gaggcggaaa gaacagctgc attaatgaat cggccaacgc gcggggagag gcggtttgcg   3323
tattgggcgc tcttccgctt cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg   3383
gcgagcggta tcagctcact caaaggcggt aatacggtta tccacagaat caggggataa   3443
cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc   3503
gttgctggcg tttttccata ggctccgccc cctgacgag catcacaaaa atcgacgctc   3563
aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag   3623
ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt ccgcctttct   3683
cccttcggga agcgtggcgc tttctcaatg ctcacgctgt aggtatctca gttcggtgta   3743
ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccccc gttcagcccg accgctgcgc   3803
cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat cgccactggc   3863
agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt   3923
gaagtggtgg cctaactacg gctacactag aaggacagta tttggtatct gcgctctgct   3983
gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc   4043
tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca   4103
agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta   4163
agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa   4223
atgaagtttt aaatcaatct aaagtatata tgagtaaact tggtctgaca gttaccaatg   4283
cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca tagttgcctg   4343
actccccgtc gtgtagataa ctacgatacg ggagggctta ccatctggcc ccagtgctgc   4403
aatgataccg cgagacccac gctcaccggc tccagattta tcagcaataa accagccagc   4463
cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc agtctattaa   4523
ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca acgttgttgc   4583
cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat tcagctccgg   4643
ttcccaacga tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag cggttagctc   4703
cttcggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat   4763
ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt ctgtgactgg   4823
tgagtactca accaagtcat tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc   4883
ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg   4943
aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat   5003
gtaacccact cgtgcaccca actgatcttc agcatctttt actttcacca gcgtttctgg   5063
gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataagggcga cacggaaatg   5123
ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg gttattgtct   5183
catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg ttccgcgcac   5243
atttccccga aaagtgccac ctgacgtc                                      5271
```

<210> 24

<211> 681
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 24

```
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
 1               5                  10                  15
Val Val Ile Ala Gly Thr Ser Ala Val Thr Leu Val Arg Lys Asn Arg
            20                  25                  30
Trp Leu Leu Leu Asn Val Thr Ser Glu Asp Leu Gly Lys Thr Phe Ser
        35                  40                  45
Val Gly Thr Gly Asn Cys Thr Thr Asn Ile Leu Glu Ala Lys Tyr Trp
    50                  55                  60
Cys Pro Asp Ser Met Glu Tyr Asn Cys Pro Asn Leu Ser Pro Arg Glu
65                  70                  75                  80
Glu Pro Asp Asp Ile Asp Cys Trp Cys Tyr Gly Val Glu Asn Val Arg
                85                  90                  95
Val Ala Tyr Gly Lys Cys Asp Ser Ala Gly Arg Ser Arg Arg Ser Arg
            100                 105                 110
Arg Ala Ile Asp Leu Pro Thr His Glu Asn His Gly Leu Lys Thr Arg
        115                 120                 125
Gln Glu Lys Trp Met Thr Gly Arg Met Gly Glu Arg Gln Leu Gln Lys
    130                 135                 140
Ile Glu Arg Trp Phe Val Arg Asn Pro Phe Phe Ala Val Thr Ala Leu
145                 150                 155                 160
Thr Ile Ala Tyr Leu Val Gly Ser Asn Met Thr Gln Arg Val Val Ile
                165                 170                 175
Ala Leu Leu Val Leu Ala Val Gly Pro Ala Tyr Ser Ala His Cys Ile
            180                 185                 190
Gly Ile Thr Asp Arg Asp Phe Ile Glu Gly Val His Gly Gly Thr Trp
        195                 200                 205
Val Ser Ala Thr Leu Glu Gln Asp Lys Cys Val Thr Val Met Ala Pro
    210                 215                 220
Asp Lys Pro Ser Leu Asp Ile Ser Leu Glu Thr Val Ala Ile Asp Arg
225                 230                 235                 240
Pro Ala Glu Val Arg Lys Val Cys Tyr Asn Ala Val Leu Thr His Val
                245                 250                 255
Lys Ile Asn Asp Lys Cys Pro Ser Thr Gly Glu Ala His Leu Ala Glu
            260                 265                 270
Glu Asn Glu Gly Asp Asn Ala Cys Lys Arg Thr Tyr Ser Asp Arg Gly
        275                 280                 285
Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Val Ala Cys
    290                 295                 300
Ala Lys Phe Thr Cys Ala Lys Ser Met Ser Leu Phe Glu Val Asp Gln
305                 310                 315                 320
Thr Lys Ile Gln Tyr Val Ile Arg Ala Gln Leu His Val Gly Ala Lys
                325                 330                 335
Gln Glu Asn Trp Thr Thr Asp Ile Lys Thr Leu Lys Phe Asp Ala Leu
            340                 345                 350
Ser Gly Ser Gln Glu Val Glu Phe Ile Gly Tyr Gly Lys Ala Thr Leu
        355                 360                 365
Glu Cys Gln Val Gln Thr Ala Val Asp Phe Gly Asn Ser Tyr Ile Ala
    370                 375                 380
Glu Met Glu Thr Glu Ser Trp Ile Val Asp Arg Gln Trp Ala Gln Asp
385                 390                 395                 400
Leu Thr Leu Pro Trp Gln Ser Gly Ser Gly Gly Val Trp Arg Glu Met
                405                 410                 415
```

```
His His Leu Val Glu Phe Glu Pro Pro His Ala Ala Thr Ile Arg Val
            420                 425                 430
Leu Ala Leu Gly Asn Gln Glu Gly Ser Leu Lys Thr Ala Leu Thr Gly
        435                 440                 445
Ala Met Arg Val Thr Lys Asp Thr Asn Asp Asn Asn Leu Tyr Lys Leu
    450                 455                 460
His Gly Gly His Val Ser Cys Arg Val Lys Leu Ser Ala Leu Thr Leu
465                 470                 475                 480
Lys Gly Thr Ser Tyr Lys Ile Cys Thr Asp Lys Met Phe Phe Val Lys
                485                 490                 495
Asn Pro Thr Asp Thr Gly His Gly Thr Val Val Met Gln Val Lys Val
            500                 505                 510
Ser Lys Gly Ala Pro Cys Arg Ile Pro Val Ile Val Ala Asp Asp Leu
        515                 520                 525
Thr Ala Ala Ile Asn Lys Gly Ile Leu Val Thr Val Asn Pro Ile Ala
    530                 535                 540
Ser Thr Asn Asp Asp Glu Val Leu Ile Glu Val Asn Pro Pro Phe Gly
545                 550                 555                 560
Asp Ser Tyr Ile Ile Val Gly Arg Gly Asp Ser Arg Leu Thr Tyr Gln
                565                 570                 575
Trp His Lys Glu Gly Ser Ser Ile Gly Lys Leu Phe Thr Gln Thr Met
            580                 585                 590
Lys Gly Val Glu Arg Leu Ala Val Met Gly Asp Thr Ala Trp Asp Phe
        595                 600                 605
Ser Ser Ala Gly Gly Phe Phe Thr Ser Val Gly Lys Gly Ile His Thr
    610                 615                 620
Val Phe Gly Ser Ala Phe Gln Gly Leu Phe Gly Gly Leu Asn Trp Ile
625                 630                 635                 640
Thr Lys Val Ile Met Gly Ala Val Leu Ile Trp Val Gly Ile Asn Thr
                645                 650                 655
Arg Asn Met Thr Met Ser Met Ser Met Ile Leu Val Gly Val Ile Met
            660                 665                 670
Met Phe Leu Ser Leu Gly Val Gly Ala
        675                 680
```

<210> 25
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-35
<223> POW 454

<400> 25
aaaagaaaaa gcgctaccac catccaccgg gacag 35

<210> 26
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature

<222> 1-41
<223> CPOW 2417

<400> 26
actgttaccc tcaacccegt actcgccggc gaaaaagaaa a 41

<210> 27
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<223> Modified JE Signal

<400> 27
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
1 5 10 15
Val Val Ile Ala Gly Thr Ser Ala
20

<210> 28
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-36
<223> YF 482

<400> 28
aaaagaaaaa gcgctgtgac cttggtgcgg aaaaac 36

<210> 29
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-41
<223> CYF 2433

<400> 29
acagagatcc tcaacccegc actcgccggc gaaaaagaaa a 41

<210> 30
<211> 41
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-41
<223> SLE 463

<400> 30
aaaagaaaaa gcgctttgca gttatcaacc tatcagggga a 41

<210> 31
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-40
<223> CSLE 2477

<400> 31
accgttggtc gcacgttcgg actcgccggc gaaaaagaaa 40

<210> 32
<211> 39
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 32
Leu Asp Thr Ile Asn Arg Arg Pro Ser Lys Lys Arg Gly Gly Thr Arg
1 5 10 15
Ser Leu Leu Gly Leu Ala Ala Leu Ile Gly Leu Ala Ser Ser Leu Gln
20 25 30
Leu Leu Ser Thr Tyr Gln Gly
35

<210> 33
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 33
Met Trp Leu Ala Ser Leu Ala Val Val Ile Ala Cys Ala Gly Ala Met
1 5 10 15
Lys Leu Ser Asn Phe Gln Gly Lys
20

<210> 34
<211> 30
<212> PRT

<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 34
Met Asn Glu Gly Ser Ile Met Trp Leu Ala Ser Leu Ala Val Val Ile
1 5 10 15
Ala Cys Ala Gly Ala Met Lys Leu Ser Asn Phe Gln Gly Lys
20 25 30

<210> 35
<211> 39
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 35
Met Gly Arg Lys Gln Asn Lys Arg Gly Gly Asn Glu Gly Ser Ile Met
1 5 10 15
Trp Leu Ala Ser Leu Ala Val Val Ile Ala Cys Ala Gly Ala Met Lys
20 25 30
Leu Ser Asn Phe Gln Gly Lys
35

<210> 36
<211> 34
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 36
Met Ser Lys Lys Arg Gly Gly Ser Glu Thr Ser Val Leu Met Val Ile
1 5 10 15
Phe Met Leu Ile Gly Phe Ala Ala Ala Leu Lys Leu Ser Asn Phe Gln
20 25 30
Gly Lys

<210> 37
<211> 33
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 37
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
1 5 10 15
Val Val Ile Ala Cys Ala Gly Ala Val Thr Leu Ser Asn Phe Gln Gly
20 25 30
Lys

<210> 38
<211> 46
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 38
Met Asn Val Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn
1 5 10 15
Ile Leu Asn Arg Arg Arg Arg Thr Ala Gly Met Ile Ile Met Leu Ile
20 25 30
Pro Thr Val Met Ala Phe His Leu Thr Thr Arg Asn Gly Glu
35 40 45

<210> 39
<211> 40
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 39
Met Val Gly Leu Gln Lys Arg Gly Lys Arg Arg Ser Ala Thr Asp Trp
1 5 10 15
Met Ser Trp Leu Leu Val Ile Thr Leu Leu Gly Met Thr Leu Ala Ala
20 25 30
Thr Val Arg Lys Glu Arg Gly Asp
35 40

<210> 40
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 40
Met Gly Trp Leu Leu Val Val Val Leu Leu Gly Val Thr Leu Ala Ala
1 5 10 15
Thr Val Arg Lys Glu Arg Gly Asp
20

<210> 41
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 41

Met Ser Trp Leu Leu Val Ile Thr Leu Leu Gly Met Thr Ile Ala Ala
1 5 10 15
Thr Val Arg Lys Glu Arg Gly Asp
20

<210> 42
<211> 5292
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (910)...(2964)

<400> 42
gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc 900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc 951
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser
1 5 10

ttg gca gtt gtc ata gct tgt gca ggc gcc ttc cat tta acg acg cgt 999
Leu Ala Val Val Ile Ala Cys Ala Gly Ala Phe His Leu Thr Thr Arg
15 20 25 30

aac gga gaa cca cac atg atc gtc agc aga caa gag aaa ggg aaa agt 1047
Asn Gly Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly Lys Ser
35 40 45

ctt ctg ttt aaa aca gag gat ggc gtg aac atg tgt acc ctc atg gcc 1095
Leu Leu Phe Lys Thr Glu Asp Gly Val Asn Met Cys Thr Leu Met Ala
50 55 60

atg gac ctt ggt gaa ttg tgt gaa gac aca atc acg tac aag tgt ccc 1143
Met Asp Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys Cys Pro
65 70 75

ctt ctc agg cag aat gag cca gaa gac ata gac tgt tgg tgc aac tct 1191
Leu Leu Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Ser
80 85 90

acg tcc acg tgg gta act tat ggg acg tgt acc acc atg gga gaa cat 1239
Thr Ser Thr Trp Val Thr Tyr Gly Thr Cys Thr Thr Met Gly Glu His
95 100 105 110 .

```
aga aga gaa aaa aga tca gtg gca ctc gtt cca cat gtg gga atg gga      1287
Arg Arg Glu Lys Arg Ser Val Ala Leu Val Pro His Val Gly Met Gly
                115                 120                 125

ctg gag aca cga act gaa aca tgg atg tca tca gaa ggg gcc tgg aaa      1335
Leu Glu Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys
            130                 135                 140

cat gtc cag aga att gaa act tgg atc ttg aga cat cca ggc ttc acc      1383
His Val Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr
        145                 150                 155

atg atg gca gca atc ctg gca tac acc ata gga acg aca cat ttc caa      1431
Met Met Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln
    160                 165                 170

aga gcc ctg att ttc atc tta ctg aca gct gtc act cct tca atg aca      1479
Arg Ala Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr
175                 180                 185                 190

atg cgt tgc ata gga atg tca aat aga gac ttt gtg gaa ggg gtt tca      1527
Met Arg Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser
                195                 200                 205

gga gga agc tgg gtt gac ata gtc tta gaa cat gga agc tgt gtg acg      1575
Gly Gly Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr
            210                 215                 220

acg atg gca aaa aac aaa cca aca ttg gat ttt gaa ctg ata aaa aca      1623
Thr Met Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr
        225                 230                 235

gaa gcc aaa cag cct gcc aac cta agg aag tac tgt ata gag gca aag      1671
Glu Ala Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys
    240                 245                 250

cta acc aac aca aca aca gaa tct cgc tgc cca aca caa ggg gaa ccc      1719
Leu Thr Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro
255                 260                 265                 270

agc cta aat gaa gag cag gac aaa agg ttc gtc tgc aaa cac tcc atg      1767
Ser Leu Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met
                275                 280                 285

gta gac aga gga tgg gga aat gga tgt gga cta ttt gga aag gga ggc      1815
Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly
            290                 295                 300

att gtg acc tgt gct atg ttc aga tgc aaa aag aac atg gaa gga aaa      1863
Ile Val Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys
        305                 310                 315

gtt gtg caa cca gaa aac ttg gaa tac acc att gtg ata aca cct cac      1911
Val Val Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His
    320                 325                 330

tca ggg gaa gag cat gca gtc gga aat gac aca gga aaa cat ggc aag      1959
Ser Gly Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys
335                 340                 345                 350
```

```
gaa atc aaa ata aca cca cag agt tcc atc aca gaa gca gaa ttg aca      2007
Glu Ile Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr
                355                 360                 365

ggt tat ggc act gtc aca atg gag tgc tct cca aga acg ggc ctc gac      2055
Gly Tyr Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp
            370                 375                 380

ttc aat gag atg gtg ttg ttg cag atg gaa aat aaa gct tgg ctg gtg      2103
Phe Asn Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val
        385                 390                 395

cac agg caa tgg ttc cta gac ctg ccg tta cca tgg ttg ccc gga gcg      2151
His Arg Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala
    400                 405                 410

gac aca caa ggg tca aat tgg ata cag aaa gag aca ttg gtc act ttc      2199
Asp Thr Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe
415                 420                 425                 430

aaa aat ccc cat gcg aag aaa cag gat gtt gtt gtt tta gga tcc caa      2247
Lys Asn Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln
                435                 440                 445

gaa ggg gcc atg cac aca gca ctt aca ggg gcc aca gaa atc caa atg      2295
Glu Gly Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met
            450                 455                 460

tca tca gga aac tta ctc ttc aca gga cat ctc aag tgc agg ctg aga      2343
Ser Ser Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg
        465                 470                 475

atg gac aag cta cag ctc aaa gga atg tca tac tct atg tgc aca gga      2391
Met Asp Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly
    480                 485                 490

aag ttt aaa gtt gtg aag gaa ata gca gaa aca caa cat gga aca ata      2439
Lys Phe Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile
495                 500                 505                 510

gtt atc aga gtg caa tat gaa ggg gac ggc tct cca tgc aag atc cct      2487
Val Ile Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro
                515                 520                 525

ttt gag ata atg gat ttg gaa aaa agr cat gtc tta ggt cgc ctg att      2535
Phe Glu Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile
            530                 535                 540

aca gtc aac cca att gtg aca gaa aaa gat agc cca gtc aac ata gaa      2583
Thr Val Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu
        545                 550                 555

gca gaa cct cca ttc gga gac agc tac atc atc ata gga gta gag ccg      2631
Ala Glu Pro Pro Phe Gly Asp Ser Tyr Ile Ile Ile Gly Val Glu Pro
    560                 565                 570

gga caa ctg aag ctc aac tgg ttt aag aaa gga agt tct atc ggc caa      2679
Gly Gln Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Ser Ile Gly Gln
575                 580                 585                 590
```

atg ttt gag aca aca atg agg ggg gcg aag aga atg gcc att tta ggt 2727
Met Phe Glu Thr Thr Met Arg Gly Ala Lys Arg Met Ala Ile Leu Gly
595 600 605

gac aca gcc tgg gat ttt gga tcc ttg gga gga gtg ttt aca tct ata 2775
Asp Thr Ala Trp Asp Phe Gly Ser Leu Gly Gly Val Phe Thr Ser Ile
610 615 620

gga aag gct ctc cac caa gtc ttt gga gca atc tat gga gct gcc ttc 2823
Gly Lys Ala Leu His Gln Val Phe Gly Ala Ile Tyr Gly Ala Ala Phe
625 630 635

agt ggg gtt tca tgg act atg aaa atc ctc ata gga gtc att atc aca 2871
Ser Gly Val Ser Trp Thr Met Lys Ile Leu Ile Gly Val Ile Ile Thr
640 645 650

tgg ata gga atg aat tca cgc agc acc tca ctg tct gtg aca cta gta 2919
Trp Ile Gly Met Asn Ser Arg Ser Thr Ser Leu Ser Val Thr Leu Val
655 660 665 670

ttg gtg gga att gtg aca ctg tat ttg gga gtc atg gtg cag gcc 2964
Leu Val Gly Ile Val Thr Leu Tyr Leu Gly Val Met Val Gln Ala
675 680 685

taattagttg agcggccgct cgagcatgca tctagagggc cctattctat agtgtcacct 3024
aaatgctaga gctcgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt 3084
ttgcccctcc cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcctttccta 3144
ataaaatgag gaaattgcat cgcattgtct gagtaggtgt cattctattc tgggggggtgg 3204
ggtggggcag gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc 3264
ggtgggctct atggcttctg aggcggaaag aaccagctgc attaatgaat cggccaacgc 3324
gcggggagag gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg 3384
cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta 3444
tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc 3504
aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc cctgacgag 3564
catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac 3624
caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc 3684
ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt 3744
aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc 3804
gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga 3864
cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 3924
ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag aagaacagta 3984
tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga 4044
tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg 4104
cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag 4164
tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc 4224
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact 4284
tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt 4344
cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta 4404
ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta 4464
tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc 4524
gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat 4584
agtttgcgca acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt 4644
atggcttcat tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg 4704
tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca 4764
gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta 4824
agatgctttt ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg 4884
cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact 4944
ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg 5004
ctgttgagat ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt 5064
actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga 5124
ataagggcga cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc 5184

atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa 5244
caaataggggg ttccgcgcac atttccccga aaagtgccac ctgacgtc 5292

<210> 43
<211> 685
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 43
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
1 5 10 15
Val Val Ile Ala Cys Ala Gly Ala Phe His Leu Thr Thr Arg Asn Gly
20 25 30
Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly Lys Ser Leu Leu
35 40 45
Phe Lys Thr Glu Asp Gly Val Asn Met Cys Thr Leu Met Ala Met Asp
50 55 60
Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys Cys Pro Leu Leu
65 70 75 80
Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Ser Thr Ser
85 90 95
Thr Trp Val Thr Tyr Gly Thr Cys Thr Thr Met Gly Glu His Arg Arg
100 105 110
Glu Lys Arg Ser Val Ala Leu Val Pro His Val Gly Met Gly Leu Glu
115 120 125
Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys His Val
130 135 140
Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr Met Met
145 150 155 160
Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln Arg Ala
165 170 175
Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr Met Arg
180 185 190
Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser Gly Gly
195 200 205
Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr Thr Met
210 215 220
Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr Glu Ala
225 230 235 240
Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys Leu Thr
245 250 255
Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro Ser Leu
260 265 270
Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met Val Asp
275 280 285
Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly Ile Val
290 295 300
Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys Val Val
305 310 315 320
Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His Ser Gly
325 330 335
Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys Glu Ile
340 345 350
Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr
355 360 365
Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn
370 375 380

```
Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val His Arg
385                 390                 395                 400
Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala Asp Thr
                405                 410                 415
Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe Lys Asn
            420                 425                 430
Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln Glu Gly
        435                 440                 445
Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met Ser Ser
    450                 455                 460
Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg Met Asp
465                 470                 475                 480
Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly Lys Phe
                485                 490                 495
Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile Val Ile
            500                 505                 510
Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu
        515                 520                 525
Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile Thr Val
    530                 535                 540
Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu Ala Glu
545                 550                 555                 560
Pro Pro Phe Gly Asp Ser Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln
                565                 570                 575
Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Ser Ile Gly Gln Met Phe
            580                 585                 590
Glu Thr Thr Met Arg Gly Ala Lys Arg Met Ala Ile Leu Gly Asp Thr
        595                 600                 605
Ala Trp Asp Phe Gly Ser Leu Gly Gly Val Phe Thr Ser Ile Gly Lys
    610                 615                 620
Ala Leu His Gln Val Phe Gly Ala Ile Tyr Gly Ala Ala Phe Ser Gly
625                 630                 635                 640
Val Ser Trp Thr Met Lys Ile Leu Ile Gly Val Ile Ile Thr Trp Ile
                645                 650                 655
Gly Met Asn Ser Arg Ser Thr Ser Leu Ser Val Thr Leu Val Leu Val
            660                 665                 670
Gly Ile Val Thr Leu Tyr Leu Gly Val Met Val Gln Ala
        675                 680                 685
```

<210> 44
<211> 5293
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (810)...(2964)

<400> 44

```
gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg     60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg    120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc    180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt    240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata    300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc    360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc    420
attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt    480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt    540
```

```
atgcccagta catgaccttа tgggactttc ctacttggca gtacatctac gtattagtca    600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg    660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc    720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg    780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca    840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc    900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc     951
          Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser
           1               5                  10

ttg gca gtt gtc ata gct tgt gca ggc gcc ttc cat tta acc aca cgt       999
Leu Ala Val Val Ile Ala Cys Ala Gly Ala Phe His Leu Thr Thr Arg
 15                  20                  25                  30

aac gga gaa cca cac atg atc gtc agc aga caa gag aaa ggg aaa agt      1047
Asn Gly Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly Lys Ser
                35                  40                  45

ctt ctg ttt aaa aca gag gat ggc gtg aac atg tgt acc ctc atg gcc      1095
Leu Leu Phe Lys Thr Glu Asp Gly Val Asn Met Cys Thr Leu Met Ala
            50                  55                  60

atg gac ctt ggt gaa ttg tgt gaa gac aca atc acg tac aag tgt ccc      1143
Met Asp Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys Cys Pro
        65                  70                  75

ctt ctc agg cag aat gag cca gaa gac ata gac tgt tgg tgc aac tct      1191
Leu Leu Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Ser
    80                  85                  90

acg tcc acg tgg gta act tat ggg acg tgt acc acc atg gga gaa cat      1239
Thr Ser Thr Trp Val Thr Tyr Gly Thr Cys Thr Thr Met Gly Glu His
 95                  100                 105                 110

aga aga gaa aaa aga tca gtg gca ctc gtt cca cat gtg gga atg gga      1287
Arg Arg Glu Lys Arg Ser Val Ala Leu Val Pro His Val Gly Met Gly
                115                 120                 125

ctg gag aca cga act gaa aca tgg atg tca tca gaa ggg gcc tgg aaa      1335
Leu Glu Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys
            130                 135                 140

cat gtc cag aga att gaa act tgg atc ttg aga cat cca ggc ttc acc      1383
His Val Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr
        145                 150                 155

atg atg gca gca atc ctg gca tac acc ata gga acg aca cat ttc caa      1431
Met Met Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln
    160                 165                 170

aga gcc ctg att ttc atc tta ctg aca gct gtc act cct tca atg aca      1479
Arg Ala Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr
175                 180                 185                 190

atg cgt tgc ata gga atg tca aat aga gac ttt gtg gaa ggg gtt tca      1527
Met Arg Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser
                195                 200                 205

gga gga agc tgg gtt gac ata gtc tta gaa cat gga agc tgt gtg acg      1575
Gly Gly Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr
            210                 215                 220
```

```
acg atg gca aaa aac aaa cca aca ttg gat ttt gaa ctg ata aaa aca      1623
Thr Met Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr
        225                 230                 235

gaa gcc aaa cag cct gcc acc cta agg aag tac tgt ata gag gca aag      1671
Glu Ala Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys
    240                 245                 250

cta acc aac aca aca aca gaa tct cgc tgc cca aca caa ggg gaa ccc      1719
Leu Thr Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro
255                 260                 265                 270

agc cta aat gaa gag cag gac aaa agg ttc gtc tgc aaa cac tcc atg      1767
Ser Leu Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met
                275                 280                 285

gta gac aga gga tgg gga aat gga tgt gga cta ttt gga aag gga ggc      1815
Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly
            290                 295                 300

att gtg acc tgt gct atg ttc aga tgc aaa aag aac atg gaa gga aaa      1863
Ile Val Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys
        305                 310                 315

gtt gtg caa cca gaa aac ttg gaa tac acc att gtg ata aca cct cac      1911
Val Val Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His
    320                 325                 330

tca ggg gaa gag cat gca gtc gga aat gac aca gga aaa cat ggc aag      1959
Ser Gly Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys
335                 340                 345                 350

gaa atc aaa ata aca cca cag agt tcc atc aca gaa gca gaa ttg aca      2007
Glu Ile Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr
                355                 360                 365

ggt tat ggc act gtc aca atg gag tgc tct cca aga acg ggc ctc gac      2055
Gly Tyr Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp
            370                 375                 380

ttc aat gag atg gtg ttg ttg cag atg gaa aat aaa gct tgg ctg gtg      2103
Phe Asn Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val
        385                 390                 395

cac agg caa tgg ttc cta gac ctg ccg tta cca tgg ttg ccc gga gcg      2151
His Arg Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala
    400                 405                 410

gac aca caa ggg tca aat tgg ata cag aaa gag aca ttg gtc act ttc      2199
Asp Thr Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe
415                 420                 425                 430

aaa aat ccc cat gcg aag aaa cag gat gtt gtt gtt tta gga tcc caa      2247
Lys Asn Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln
                435                 440                 445

gaa ggg gcc atg cac aca gca ctt aca ggg gcc aca gaa atc caa atg      2295
Glu Gly Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met
            450                 455                 460
```

```
tca tca gga aac tta ctc ttc aca gga cat ctc aag tgc agg ctg aga      2343
Ser Ser Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg
        465                 470                 475

atg gac aag cta cag ctc aaa gga atg tca tac tct atg tgc aca gga      2391
Met Asp Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly
    480                 485                 490

aag ttt aaa gtt gtg aag gaa ata gca gaa aca caa cat gga aca ata      2439
Lys Phe Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile
495                 500                 505                 510

gtt atc aga gtg caa tat gaa ggg gac ggc tct cca tgc aag atc cct      2487
Val Ile Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro
                515                 520                 525

ttt gag ata atg gat ttg gaa aaa aga cat gtc tta ggt cgc ctg att      2535
Phe Glu Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile
            530                 535                 540

aca gtc aac cca att gtg aca gaa aaa gat agc cca gtc aac ata gaa      2583
Thr Val Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu
        545                 550                 555

gca gaa cct cca ttc gga gac agc cac atc atc ata gga gta gag ccg      2631
Ala Glu Pro Pro Phe Gly Asp Ser His Ile Ile Ile Gly Val Glu Pro
    560                 565                 570

gga caa ctg aag ctc aac tgg ttt aag aaa gga agt tct atc ggc caa      2679
Gly Gln Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Ser Ile Gly Gln
575                 580                 585                 590

atg ttt gag aca aca atg agg ggg gcg aag aga atg gcc att tta ggt      2727
Met Phe Glu Thr Thr Met Arg Gly Ala Lys Arg Met Ala Ile Leu Gly
                595                 600                 605

gac aca gcc tgg gat ttt gga tcc ttg gga gga gtg ttt aca tct ata      2775
Asp Thr Ala Trp Asp Phe Gly Ser Leu Gly Gly Val Phe Thr Ser Ile
            610                 615                 620

gga aag gct ctc cac caa gtg ttt ggt ggt gcc ttc aga aca ctc ttt      2823
Gly Lys Ala Leu His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe
        625                 630                 635

ggg gga atg tct tgg atc aca caa ggg cta atg ggt gcc cta ctg ctc      2871
Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu
    640                 645                 650

tgg atg ggc gtc aac gca cga gac cga tca att gct ttg gcc ttc tta      2919
Trp Met Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu
655                 660                 665                 670

gcc aca ggg ggt gtg ctc gtg ttc tta gcg acc aat gtg cat gct          2964
Ala Thr Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala
                675                 680                 685

taattagttt gggcggcgcgc tcgagcatgc atctagaggg ccctattcta tagtgtcacc   3024
taaatgctag agctcgctga tcagcctcga ctgtgccttc tagttgccag ccatctgttg   3084
tttgcccctc ccccgtgcct tccttgaccc tggaaggtgc cactcccact gtcctttcct   3144
aataaaatga ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg  3204
gggtggggca ggacagcaag ggggaggatt gggaagacaa tagcaggcat gctgggggatg  3264
```

cggtgggctc tatggcttct gaggcggaaa gaaccagctg cattaatgaa tcggccaacg 3324
cgcgggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca ctgactcgct 3384
gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt 3444
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 3504
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga 3564
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 3624
ccaggcgttt cccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac 3684
cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg 3744
taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 3804
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag 3864
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt 3924
aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta gaagaacagt 3984
atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 4044
atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac 4104
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca 4164
gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac 4224
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 4284
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 4344
tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac gggagggctt 4404
accatctggc cccagtgctg caatgatacc gcgagaccca cgctcaccgg ctccagattt 4464
atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc 4524
cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa 4584
tagtttgcgc aacgttgttg ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg 4644
tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat cccccatgtt 4704
gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc 4764
agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt 4824
aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg 4884
gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac 4944
tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc 5004
gctgttgaga tccagttcga tgtaacccac tcgtgcaccc aactgatctt cagcatcttt 5064
tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg 5124
aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat attattgaag 5184
catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa 5244
acaaataggg gttccgcgca catttccccg aaaagtgccc cctgacgtc 5293

<210> 45
<211> 685
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 45

Met	Gly	Lys	Arg	Ser	Ala	Gly	Ser	Ile	Met	Trp	Leu	Ala	Ser	Leu	Ala
1				5				10					15		
Val	Val	Ile	Ala	Cys	Ala	Gly	Ala	Phe	His	Leu	Thr	Thr	Arg	Asn	Gly
			20					25				30			
Glu	Pro	His	Met	Ile	Val	Ser	Arg	Gln	Glu	Lys	Gly	Lys	Ser	Leu	Leu
		35					40					45			
Phe	Lys	Thr	Glu	Asp	Gly	Val	Asn	Met	Cys	Thr	Leu	Met	Ala	Met	Asp
	50					55					60				
Leu	Gly	Glu	Leu	Cys	Glu	Asp	Thr	Ile	Thr	Tyr	Lys	Cys	Pro	Leu	Leu
65					70					75					80
Arg	Gln	Asn	Glu	Pro	Glu	Asp	Ile	Asp	Cys	Trp	Cys	Asn	Ser	Thr	Ser
				85					90					95	
Thr	Trp	Val	Thr	Tyr	Gly	Thr	Cys	Thr	Thr	Met	Gly	Glu	His	Arg	Arg
			100					105					110		
Glu	Lys	Arg	Ser	Val	Ala	Leu	Val	Pro	His	Val	Gly	Met	Gly	Leu	Glu
		115					120					125			

Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys His Val
130 135 140
Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr Met Met
145 150 155 160
Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln Arg Ala
165 170 175
Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr Met Arg
180 185 190
Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser Gly Gly
195 200 205
Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr Thr Met
210 215 220
Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr Glu Ala
225 230 235 240
Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys Leu Thr
245 250 255
Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro Ser Leu
260 265 270
Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met Val Asp
275 280 285
Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly Ile Val
290 295 300
Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys Val Val
305 310 315 320
Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His Ser Gly
325 330 335
Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys Glu Ile
340 345 350
Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr
355 360 365
Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn
370 375 380
Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val His Arg
385 390 395 400
Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala Asp Thr
405 410 415
Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe Lys Asn
420 425 430
Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln Glu Gly
435 440 445
Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met Ser Ser
450 455 460
Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg Met Asp
465 470 475 480
Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly Lys Phe
485 490 495
Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile Val Ile
500 505 510
Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu
515 520 525
Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile Thr Val
530 535 540
Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu Ala Glu
545 550 555 560
Pro Pro Phe Gly Asp Ser His Ile Ile Ile Gly Val Glu Pro Gly Gln
565 570 575
Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Ser Ile Gly Gln Met Phe
580 585 590
Glu Thr Thr Met Arg Gly Ala Lys Arg Met Ala Ile Leu Gly Asp Thr
595 600 605

```
Ala Trp Asp Phe Gly Ser Leu Gly Gly Val Phe Thr Ser Ile Gly Lys
    610                 615                 620
Ala Leu His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly
625                 630                 635                 640
Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met
                645                 650                 655
Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr
            660                 665                 670
Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala
        675                 680                 685
```

<210> 46
<211> 8293
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (910)...(2964)

<400> 46

```
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg      60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg     120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc     180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt     240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata     300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc     360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc     420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt     480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt     540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca     600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg     660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc     720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg     780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca     840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc     900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc     951
          Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser
           1               5                  10

ttg gca gtt gtc ata gct tgt gca ggc gcc ttc cat tta acc aca cgt      999
Leu Ala Val Val Ile Ala Cys Ala Gly Ala Phe His Leu Thr Thr Arg
 15                  20                  25                  30

aac gga gaa cca cac atg atc gtc agc aga caa gag aaa ggg aaa agt     1047
Asn Gly Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly Lys Ser
                35                  40                  45

ctt ctg ttt aaa aca gag gat ggc gtg aac atg tgt acc ctc atg gcc     1095
Leu Leu Phe Lys Thr Glu Asp Gly Val Asn Met Cys Thr Leu Met Ala
            50                  55                  60

atg gac ctt ggt gaa ttg tgt gaa gac aca atc acg tac aag tgt ccc     1143
Met Asp Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys Cys Pro
        65                  70                  75
```

```
ctt ctc agg cag aat gag cca gaa gac ata gac tgt tgg tgc aac tct     1191
Leu Leu Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Ser
     80                  85                  90

acg tcc acg tgg gta act tat ggg acg tgt acc acc atg gga gaa cat     1239
Thr Ser Thr Trp Val Thr Tyr Gly Thr Cys Thr Thr Met Gly Glu His
 95                 100                 105                 110

aga aga gaa aaa aga tca gtg gca ctc gtt cca cat gtg gga atg gga     1287
Arg Arg Glu Lys Arg Ser Val Ala Leu Val Pro His Val Gly Met Gly
                115                 120                 125

ctg gag aca cga act gaa aca tgg atg tca tca gaa ggg gcc tgg aaa     1335
Leu Glu Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys
            130                 135                 140

cat gtc cag aga att gaa act tgg atc ttg aga cat cca ggc ttc acc     1383
His Val Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr
        145                 150                 155

atg atg gca gca atc ctg gca tac acc ata gga acg aca cat ttc caa     1431
Met Met Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln
    160                 165                 170

aga gcc atg att ttc atc tta ctg aca gct gtc act cct tca atg aca     1479
Arg Ala Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr
175                 180                 185                 190

atg cgt tgc ata gga atg tca aat aga gac ttt gtg gaa ggg gtt tca     1527
Met Arg Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser
                195                 200                 205

gga gga agc tgg gtt gac ata gtc tta gaa cat ggg agc tgt gtg acg     1575
Gly Gly Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr
            210                 215                 220

acg atg gca aaa aac aaa cca aca ttg gat ttt gaa ctg ata aaa aca     1623
Thr Met Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr
        225                 230                 235

gaa gct aaa cag cct gcc acc cta agg aag tac tgt ata gag gca aag     1671
Glu Ala Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys
    240                 245                 250

cta acc aac aca aca aca gaa tct cgc tgc cca aca caa ggg gaa ccc     1719
Leu Thr Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro
255                 260                 265                 270

agc cta aat gaa gag cag gac aaa agg ttc gtc tgc aaa cac tcc atg     1767
Ser Leu Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met
                275                 280                 285

gta gac aga gga tgg gga aat gga tgt gga cta ttt gga aag gga ggc     1815
Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly
            290                 295                 300

att gtg acc tgt gct atg ttc aga tgc aaa aag aac atg gaa gga aaa     1863
Ile Val Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys
        305                 310                 315
```

```
gtt gtg caa cca gaa aac ttg gaa tac acc att gtg ata aca cct cac      1911
Val Val Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His
    320                 325                 330

tca ggg gaa gag cat gca gtc gga aat gac aca gga aaa cat ggc aag      1959
Ser Gly Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys
335                 340                 345                 350

gaa atc aaa ata aca cca cag agt tcc atc aca gaa gca gaa ttg aca      2007
Glu Ile Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr
                355                 360                 365

ggt tat ggc act gtc aca atg gag tgc tct cca aga acg ggc ctc gac      2055
Gly Tyr Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp
            370                 375                 380

ttc aat gag atg gtg ttg ttg cag atg gaa aat aaa gct tgg ctg gtg      2103
Phe Asn Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val
        385                 390                 395

cac agg caa tgg ttc cta gac ctg ccg tta cca tgg ttg ccc gga gcg      2151
His Arg Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala
    400                 405                 410

gac aca caa ggg tca aat tgg ata cag aaa gag aca ttg gtc act ttc      2199
Asp Thr Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe
415                 420                 425                 430

aaa aat ccc cat gcg aag aaa cag gat gtt gtt gtt tta gga tcc caa      2247
Lys Asn Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln
                435                 440                 445

gaa ggg gcc atg cac aca gca ctt aca ggg gcc aca gaa atc caa atg      2295
Glu Gly Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met
            450                 455                 460

tca tca gga aac tta ctc ttc aca gga cat ctc aag tgc agg ctg aga      2343
Ser Ser Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg
        465                 470                 475

atg gac aag cta cag ctc aaa gga atg tca tac tct atg tgc aca gga      2391
Met Asp Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly
    480                 485                 490

aag ttt aaa gtt gtg aag gaa ata gca gaa aca caa cat gga aca ata      2439
Lys Phe Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile
495                 500                 505                 510

gtt atc aga gtg caa tat gaa ggg gac ggc tct cca tgc aag atc cct      2487
Val Ile Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro
                515                 520                 525

ttt gag ata atg gat ttg gaa aaa aga cat gtc tta ggt cgc ctg att      2535
Phe Glu Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile
            530                 535                 540

aca gtc aac cca att gtg aca gaa aaa gat agc cca gtc aac ata gaa      2583
Thr Val Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu
        545                 550                 555
```

```
gca gaa cct cca ttc gga gac agc tac atc atc ata gga gta gag ccg     2631
Ala Glu Pro Pro Phe Gly Asp Ser Tyr Ile Ile Ile Gly Val Glu Pro
    560                 565                 570

gga caa ctg aag ctc aac tgg ttt aag aaa gga agc acg ctg ggc aag     2679
Gly Gln Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Thr Leu Gly Lys
575                 580                 585                 590

gcc ttt tca aca act ttg aag gga gct caa aga ctg gca gcg ttg ggc     2727
Ala Phe Ser Thr Thr Leu Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly
                595                 600                 605

gac aca gcc tgg gac ttt ggc tct att gga ggg gtc ttc aac tcc ata     2775
Asp Thr Ala Trp Asp Phe Gly Ser Ile Gly Gly Val Phe Asn Ser Ile
            610                 615                 620

gga aaa gcc gtt caa caa gtg ttt ggt ggt gcc ttc aga aca ctc ttt     2823
Gly Lys Ala Val His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe
        625                 630                 635

ggg gga atg tct tgg atc aca caa ggg cta atg ggt gcc cta ctg ctc     2871
Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu
    640                 645                 650

tgg atg ggc gtc aac gca cga gac cga tca att gct ttg gcc ttc tta     2919
Trp Met Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu
655                 660                 665                 670

gcc aca ggg ggt gtg ctc gtg ttc tta gcg acc aat gtg cat gct         2964
Ala Thr Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala
                675                 680                 685

taattagttt gagcggccgc tcgagcatgc atctagaggg cccattctta tagtgtcacc   3024
taaatgctag agctcgctga tcagcctcga ctgtgccttc tagttgccag ccatctgttg   3084
tttgcccctc ccccgtgcct tccttgaccc tggaaggtgc cactcccact gtcctttcct   3144
aataaaatga ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg   3204
gggtggggca ggacagcaag ggggaggatt gggaagacaa tagcaggcat gctggggatg   3264
cggtgggctc tatggcttct gaggcggaaa gaaccagctg cattaatgaa tcggccaacg   3324
cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca ctgactcgct   3384
gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt   3444
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc   3504
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga   3564
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata   3624
ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac   3684
cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcaat gctcacgctg   3744
taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc   3804
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag   3864
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt   3924
aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta gaaggacagt   3984
atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg   4044
atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac   4104
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca   4164
gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac   4224
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac   4284
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt   4344
tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac gggagggctt   4404
accatctggc cccagtgctg caatgatacc gcgagaccca cgctcaccgg ctccagattt   4464
atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc   4524
cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa   4584
tagtttgcgc aacgttgttg ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg   4644
tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat cccccatgtt   4704
```

gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc 4764
agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt 4824
aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg 4884
gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac 4944
tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc 5004
gctgttgaga tccagttcga tgtaacccac tcgtgcaccc aactgatctt cagcatcttt 5064
tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg 5124
aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat attattgaag 5184
catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa 5244
acaaatagggg gttccgcgca catttccccg aaaagtgcca cctgacgtc 5293

<210> 47
<211> 685
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 47
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
1 5 10 15
Val Val Ile Ala Cys Ala Gly Ala Phe His Leu Thr Thr Arg Asn Gly
20 25 30
Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly Lys Ser Leu Leu
35 40 45
Phe Lys Thr Glu Asp Gly Val Asn Met Cys Thr Leu Met Ala Met Asp
50 55 60
Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys Cys Pro Leu Leu
65 70 75 80
Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Ser Thr Ser
85 90 95
Thr Trp Val Thr Tyr Gly Thr Cys Thr Thr Met Gly Glu His Arg Arg
100 105 110
Glu Lys Arg Ser Val Ala Leu Val Pro His Val Gly Met Gly Leu Glu
115 120 125
Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys His Val
130 135 140
Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr Met Met
145 150 155 160
Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln Arg Ala
165 170 175
Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr Met Arg
180 185 190
Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser Gly Gly
195 200 205
Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr Thr Met
210 215 220
Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr Glu Ala
225 230 235 240
Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys Leu Thr
245 250 255
Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro Ser Leu
260 265 270
Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met Val Asp
275 280 285
Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly Ile Val
290 295 300
Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys Val Val
305 310 315 320

Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His Ser Gly
325 330 335
Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys Glu Ile
340 345 350
Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr
355 360 365
Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn
370 375 380
Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val His Arg
385 390 395 400
Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala Asp Thr
405 410 415
Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe Lys Asn
420 425 430
Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln Glu Gly
435 440 445
Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met Ser Ser
450 455 460
Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg Met Asp
465 470 475 480
Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly Lys Phe
485 490 495
Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile Val Ile
500 505 510
Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu
515 520 525
Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile Thr Val
530 535 540
Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu Ala Glu
545 550 555 560
Pro Pro Phe Gly Asp Ser Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln
565 570 575
Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Thr Leu Gly Lys Ala Phe
580 585 590
Ser Thr Thr Leu Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr
595 600 605
Ala Trp Asp Phe Gly Ser Ile Gly Gly Val Phe Asn Ser Ile Gly Lys
610 615 620
Ala Val His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly
625 630 635 640
Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met
645 650 655
Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr
660 665 670
Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala
675 680 685

<210> 48
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 48
tgtgcaggcg ccttccattt aaccacacgt aacg 34

<210> 49
<211> 40

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 49
tcgagcggcc gctcaactaa ttaggcctgc accatgactc 40

<210> 50
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 50
cttatcgaaa ttaatacgac tcactatagg 30

<210> 51
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 51
atagattgct ccaaacactt ggtgg 25

<210> 52
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 52
actccatagg aaaagccgtt cacc 24

<210> 53
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 53
gcgagctcta gcatttaggt gacactatag 30

<210> 54
<211> 33
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 54
ctccaccaag tgtttggtgg tgccttcaga aca 33

<210> 55
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 55
Leu His Gln Val Phe Gly Gly Ala Phe Arg Thr
1 5 10

<210> 56
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 56
cttatcgaaa ttaatacgac tcactatagg 30

<210> 57
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 57
gaattcgtct cacttccttt cttaaaccag ttgagcttc 39

<210> 58
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 58
ggaattcgtc tcggaagcac gctgggcaag g 31

<210> 59
<211> 30
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 59
gcgagctcta gcatttaggt gacactatag 30

<210> 60
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 60
aactggttta agaaaggaag cacgctgggc gcc 33

<210> 61
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 61
Asn Trp Lys Lys Gly Ser Thr Leu Gly Lys Ala
1 5 10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/10764

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :Please See Extra Sheet.
US CL :435/5, 7.1, 69.1; 530/387.1; 424/818.1; 536/23.72
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)

U.S. : 435/5, 7.1, 69.1; 530/387.1; 424/818.1; 536/23.72

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 practicable, search terms used)

MEDLINE, BIOSIS, WPIDS, USPATFUL

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,494,671 A (LAI, C., et al.) 27 February 1996, see entire document.	1-43
Y	US 6,074,865 A (KELLY, E. P. and A. D. KING) 13 June 2000, see entire document.	1-43
Y	US 6,165,477 A (IVY, J., et al.) 26 December 2000, see entire document.	1-43

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"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
"A"	document defining the general state of the art which is not considered to be of particular relevance	"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
"E"	earlier document published on or after the international filing date	"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
"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)	"&"	document member of the same patent family
"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		

Date of the actual completion of the international search 18 JUNE 2002	Date of mailing of the international search report 18 JUL 2002
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer J. S. PARKIN Telephone No. (703) 308-1234

Form PCT/ISA/210 (second sheet) (July 1998)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/10764

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (7):

C12Q 1/70; G01N 33/53; C12P 21/06; C07K 16/00; C07H 21/04; A61K 39/12

VACUNAS DE ÁCIDO NUCLÉICO PARA PREVENCIÓN DE LA INFECCIÓN DE FLAVIVIRUS

Esta solicitud es una continuación parcial de la solicitud de patente norteamericana Serie No. 09/826.115 presentada el 4 de abril de 2001, y la cual reivindica el beneficio de la misma, cuyo estatutos se encuentra pendiente y la cual es una continuación en parte de la solicitud de patente norteamericana Serie No. 09/701.536, presentada el 29 de noviembre de 2000, y reivindica el beneficio de la misma, cuyo estatutos se encuentra pendiente y la cual es una solicitud de etapa nacional de la solicitud internacional Serie No. PCT/US99/12298 presentada el 3 de junio de 1999, que reivindica el beneficio de la solicitud provisional norteamericana Serie No. 60/087.908 presentada el 4 de junio de 1998, cuyas solicitudes se incorporan a este documento en su totalidad por referencia.

Campo del Invento

Este invento se relaciona con vacunas novedosas, diagnósticos y métodos para utilizar tanto en el tratamiento como en la prevención de enfermedades causadas por flavivirus. En particular, las vacunas son ácidos nucléicos recombinantes que contienen genes para proteínas estructurales de flavivirus, tales como el

María del Pilar Iglesias
C. C. 41.339.183 de Bogotá
Intérprete Oficial según Resolución No. 819 del
1971 del Ministerio de Justicia

000189

virus de encefalitis japonesa (JEV), el virus West Nile (WNV) o flavivirus relacionados. Estas vacunas sirven como una unidad transcripcional para la biosíntesis de los antígenos de la proteína del virus cuando se administran *in vivo*. Los diagnósticos son composiciones que contienen antígenos producidos de los ácidos nucléicos recombinantes y se pueden utilizar para detectar la infección del flavivirus.

Antecedes del Invento

Los flavivirus son miembros del género *Flavivirus*, que se clasifica dentro de la familia Flaviviridae. Los flavivirus son ampliamente patogénicos para los humanos y otros mamíferos. Los flavivirus que causan enfermedades en los humanos y en los animales incluyen Alfuy, Apoi, Aroa, Bagaza, Banzi, Batu · Cave, Bouboui, murciélago Bukalasa, Bussuquara, Cacipacore, Isla de Carey, Cowbone Ridge, murciélago Dakar, Dengue (serotipos 1, 2, 3 y 4), Edge Hill, murciélago Entebbe, Gadgets Gully, Iguape, Ilheus, meningoencefalitis de pavo de Israel, encefalisitis Japonesa, Jugra, Jutiapa, Kadam, Karshi, Kedougou, Kokobera, Koutango, Kunjin, enfermedad de la selva Kyasanur, Langat, Meaban, Modoc, Montana myotis leucoencefalitis, encefalitis Murray Valley, Naranjal, Negishi, Ntaya, fiebre hemorrágica Omsk, murciélago de Phnom Penh, Potiskum, Powassan, Río Bravo, Rocío, Royal Farm, encefalitis de verano primaveral Ruso, Saboya, Sal

Vieja, San Perlita, Saumarez Reef, Sepik, Sokuluk, Spondweni, encefalitis de St. Louis, Stratford, encefalitis transmitida por garrapata - subtipo de Europa central, encefalitis transmitida por garrapata - subtipo del lejano oriente, Tembusu, THCAr, Tyuleniy, Uganda S, Usutu, West Nile, Yaounde, Fiebre amarilla, Yokose, Ziki, Agente que funde las células y otros Flavivirus relacionados, tal como se enumeran en Kuno y otros. (*J. Virol.* 72: 73-83 (1998)).

Los flavivirus contienen las siguientes tres proteínas estructurales: prM/M, la proteína de premembrana y de membrana; E, la proteína envolvente; y C, la proteína cápsida. (Monath, en Virología (ed. Fields.), Imprenta Raven, Nueva York, 1990, pp. 763-814; Heinz y Roehrig, en Inmunoquímica de los Virus II: La Base para Serodiagnóstico y Vacunas (eds. van Regenmortel y Neurath.), Elsevier, Ámsterdam, 1990, pp. 289-305). M tiene un peso molecular (MW) de alrededor de 7-8 kilodaltones (kDa) y E tiene un MW de alrededor de 55-60 kDa. M se sintetiza como un precursor más grande denominado prM. La porción pr de prM se remueve cuando prM se procesa para formar la proteína M en viriones maduros. M y E están localizados en la membrana de la partícula del flavivirus, y por consiguiente hace tiempo se considera que constituyen componentes inmunogénicos importantes de los virus.

990132

Los flavivirus son virus que comprenden RNA de un solo ramal que tiene una longitud, entre las diversas especies, de alrededor de 10 kilobases (kb). La proteína C, con un peso molecular de 12-14 kDa se compleja con el RNA para formar un complejo nucleocápsido. Varias proteínas no estructurales están también codificadas por el genoma RNA, las cuales de denominan NS1, NS2A, NS2B, NS3, NS4A, NS4B y NS5. El genoma se traduce dentro de la célula huésped como una poliproteína, luego se procesa co o post-translacionalmente hacia los productos de gen individual mediante proteasas virales o de huésped específico (Figura 1).

Se conocen las secuencias nucleótidas de los genomas de varios Flavivirus, tal como se resume en la Patente Norteamericana No. 5.494.671. Para el JEV esto lo suministró Sumiyoshi y otros. (*Virología* 161: 497-510 (1987)) y Hashimoto y otros. (*Genes de Virus* 1: 305-317 (1988)). Las secuencias nucleótidas de la cepa virulenta SA-14 del JEV y de la cepa atenuada SA-14-14-2, que se utilizan como vacuna en la República Popular China, se comparan en el trabajo de Nitayaphan y otros. (*Virología* 177: 541-552 (1990)).

También se conocen las secuencias nucleótidas que codifican las proteínas estructurales de otras especies de flavivirus. En muchos casos se han reportado las

secuencias de los genomas completos. Las secuencias disponibles incluyen el virus dengue serotipo 1, el virus dengue serotipo 2 (Deubel y otros, *Virología* 155: 365-377 (1986); Gruenberg y otros, *J. Gen. Virol.* 69: 1391-1398 (1988); Hahn y otros. *Virología* 162: 167-180 (1988)), virus del dengue serotipo 3 (Osatomi y otros, *Virus Genes* 2: 99-108 (1988)), virus del dengue serotipo 4 (Mackow et al., *Virology* 159: 217-228 (1987), Zhao et al., *Virology* 155: 77-88 (1986)), virus West Nile (Lanciotti et al., *Science* 286: 2331-2333 (1999)), virus Powassan (Mandl et al., *Virology* 194: 173-184 (1993)) y virus de fiebre amarilla (YFV) (Rice et al., *Science* 229: 726-733 (1985)).

Muchos flavivirus, incluyendo el virus de encefalitis St. Louis (SLEV), WNV y JEV, se transmiten a los humanos y a otros animales huéspedes por los mosquitos. Por consiguiente ocurren en áreas ampliamente esparcidas y su transmisión no se puede interrumpir o prevenir fácilmente.

La fiebre del Nilo Occidental (West Nile fever) es una infección flaviviral transmitida por un mosquito, que se transmite a los vertebrados primordialmente por diversas especies de los mosquitos *Culex*. Al igual que otros miembros de la encefalitis japonesa (JE) el complejo antigénico de los flavivirus, incluyendo los virus JE, SLE y el de encefalitis Murray Valley (MVE), el

WNV se mantiene en un *ciclo* natural entre los vectores artrópodos y los pájaros. El virus se aisló primero de un humano febril en el distrito del Nilo Occidental de Uganda en 1937 (Smithburn et al., *Am. J. Trop. Med. Hyg.* 20: 471-492 (1940)). Pronto se reconoció como uno de los flavivirus más ampliamente distribuidos, y su rango geográfico incluía África, el Mediano Oriente, Asia Occidental, Europa y Australia (Hubalek et al., *Emerg. Infect. Dis.* 5: 643-50 (1999)). Clínicamente, la fiebre del Nilo Occidental en los humanos es una enfermedad febril aguda autolimitada acompañada por dolor de cabeza, mialgia, poliartropatía, erupción y linfadenopatía (Monath y Tsai, en Clinical Virology, (Richman, Whitley y Hayden eds.), Churchill-Livingtone, New York, 1997, pp. 1133-1186). La hepatitis aguda o la pancreatitis se han reportado ocasionalmente y algunas veces los casos de infección de WNV en pacientes mayores se complican con encefalitis o meningitis (Asnis et al., *Clin. Infect. Dis.* 30: 413-418 (2000)). Por consiguiente, la infección por WNV constituye una seria preocupación de salud en muchas regiones del mundo.

La difusión geográfica de la enfermedad, particularmente la introducción del WNV en los Estados Unidos desde 1999 ha aumentado enormemente la concientización de los problemas de salud humana y animal causados por esta enfermedad. Entre finales de agosto y

comienzos de septiembre de 1999, la Ciudad de Nueva York y las áreas circundantes experimentaron una irrupción de encefalitis viral, con 62 casos confirmados, de los cuales siete resultaron en muertes. Concurrentemente con esta irrupción, los funcionarios locales de salud observaron un aumento de la mortalidad entre los pájaros (especialmente los cuervos) y los caballos. Posteriormente se mostró que la irrupción fue causada por WNV, basandose en el mapeo y detección del anticuerpo monoclonal (Mab) de las secuencias genómicas en especimenes humanos, avianos y de mosquitos (Anderson et al., *Science* 286: 2331-2333 (1999); Jia et al., *Lancet* 354: 1971-1972 (1999); Lanciotti et al., *Science* 286: 2333-2337 (1999)). La actividad viral detectada durante los meses siguientes del invierno indicó que el virus se había establecido en norteamérica (*Morb. Mortal. Wkly. Rep.* 49: 178-179 (2000); Asnis et al., *Clin. Infect. Dis.* 30: 413-418 (2000); Garmendia et al., *J. Clin. Micro.* 38: 3110-3111 (2000)). Los datos de supervigilancia reportados de los estados nororientales y del Atlántico medio durante el año 2000 confirmaron una transmisión epidémica/epizoótica intensificada y una expansión geográfica del virus con documentación de diversos casos de infección en pájaros, mosquitos y caballos, así como casos en humanos (*Morb. Mortal. Wkly. Rep.* 49: 820-822 (2000)).

003186

Actualmente, no existe una vacuna humana o veterinaria para prevenir la infección WNV y el control del mosquito constituye la única estratégica práctica para combatir la difusión de la enfermedad.

El virus de la encefalitis japonesa (JEV) infecta a los adultos y a los niños y existe una alta tasa de mortalidad entre los infantes, los niños y los mayores en áreas de Asia tropical y subtropical (Tsai et al., en Vaccines (ed. Plotkin) W.B. Saunders, Philadelphia, Pa, 1999, pp. 672-710). Entre los sobrevivientes existen serias consecuencias neurológicas relacionadas con los síntomas de la encefalitis que persisten después de la infección. En países más desarrollados de esta región, tales como el Japón, la República de China (Taiwán) y Corea, el JEV se ha controlado ampliamente mediante la utilización de una vacuna de JEV inactivado. No obstante, sigue siendo prevalente en otros países de la región.

Las vacunas disponibles para uso contra la infección JEV incluyen virus vivos inactivados por métodos tales como tratamiento con formalina, así como con virus atenuados (Tsai et al., en Vaccines (ed. Plotkin) W.B. Saunders, Philadelphia, Pa, 1994, pp. 671-713). Las vacunas de virus totales, aún cuando son efectivas, también tienen algunos problemas y/o desventajas. Los virus se cultivan en cerebro de ratones o en cultivo de células utilizando células de mamíferos como huésped.

Esos métodos de cultivo son engorrosos y costosos. Además, existe el riesgo del operador de incorporar antígenos provenientes de las células huésped, v.g. del cerebro u otro huésped, en el producto de la vacuna final, lo que potencialmente conduce a respuestas alérgicas no pretendidas y no deseadas en los receptores de la vacuna. Existe también el riesgo de infección inadvertida entre los trabajadores involucrados en la producción de la vacuna. Finalmente, existe el riesgo de que el virus no pueda ser total o completamente inactivado o atenuado y, por consiguiente, la vacuna puede realmente causar la enfermedad.

La fiebre del dengue y la fiebre hemorrágica del dengue (DF/DHF) son causadas por virus dengue, que también es un flavivirus transmitido por mosquito. Existen cuatro estereotipos de virus dengue relacionados antigenicalmente, pero distintos, (DEN-1, DEN-2, DEN-3 y DEN-4), todos los cuales pueden causar DF/DHF. Los síntomas del DF, la enfermedad relacionada con el dengue en forma suave, incluyen fiebre, erupción, dolor de cabeza severo y dolor en las articulaciones. La mortalidad entre los sujetos que sufren de DF es baja; sin embargo, entre aquellos sujetos que sufren de DHF, la mortalidad puede llegar al 5%. De las pruebas disponibles, más de 3 millones de casos de DHF y de 58.000 muertes se han atribuido al DHF durante los

últimos 40 años, haciendo del DHF una enfermedad principal emergente (Halstead, en Dengue y Fiebre Hemorrágica de Dengue (eds. Gubler y Kuno.) CAB International, Nueva York, NY, (1997) pp 23-44). No obstante, a pesar de décadas de esfuerzo, aún no se dispone de vacunas seguras y eficaces para protegerse contra la infección del virus del dengue.

La fiebre amarilla prevalece en las regiones tropicales de Sur América y en el **Africa del sub-Sahara** y es transmitida por mosquitos. La infección produce fiebre, escalofrío, fuertes dolores de cabeza y otros dolores, anorexia, náusea y vómito, con aparición de ictericia. Una vacuna del virus vivo, -17D-, cultivada en embriones de pollo infectados, se considera segura y eficaz. No obstante, sigue existiendo la necesidad de una vacuna que sea estable bajo condiciones adversas, tales como las que comúnmente se encuentran en las regiones tropicales de África y de las Américas donde más se necesita la vacuna.

Un flavivirus recombinante que es una quimera entre dos flavivirus se revela en la publicación PCT WO 93/06214. La quimera es una proteína no estructural de estructura que se funde a partir de un "tipo" o "serotipo" del virus dengue o de un flavivirus con proteínas estructurales de un "tipo" o serotipo diferente del virus dengue u otro flavivirus.

Varias vacunas virales y de subunidad recombinante se han diseñado en los años recientes. La Patente Norteamericana No. 4.810.492 describe la producción de la glicoproteína E del JEV para su uso como antígeno en una vacuna. El DNA correspondiente se clona en un sistema de expresión con el fin de expresar la proteína de antígeno en una célula huésped idónea, tal como *E. coli*, levadura, o un cultivo de célula de organismo superior. La Patente Norteamericana No. 5.229.293 revela baculovirus recombinantes que hospedan al gen para la Proteína E JEV. El virus se utiliza para infectar células de insectos en cultivos tales que la Proteína E se produzca y recubra para utilización como una vacuna.

La Patente Norteamericana 5.021.347 revela un genoma de virus vaccinia recombinante en el que se ha incorporado el gen para la Proteína E JEV. El virus de vaccinia recombinante vivo se utiliza como la vacuna para inmunización frente al JEV. Los virus de vaccinia recombinante y los baculovirus en los que los virus incorporan un gen para el truncamiento de un terminal-C de la proteína E del dengue serotipo 2, dengue serotipo 4 y JEV se revelan en la Patente Norteamericana 5.494.671. La Patente Norteamericana 5.514.375 revela diversos virus de vaccinia recombinante que expresan porciones del marco de lectura abierta del JEV que se prolongan desde el prM hasta el NS2B. Estos poxvirus indujeron la

000191

formación de partículas extracelulares que contienen la proteína M procesada y la proteína E. Dos virus recombinantes que codifican estas proteínas JEV produjeron altas titulaciones de anticuerpos neutralizantes e inhibidores de hemaglutinina, e inmunidad protectora, en ratones. El alcance de estos efectos fue mayor después de dos tratamientos de inmunización que solo después de uno. El virus vaccinia recombinante que contiene genes para las proteínas prM/M y E de JEV confirieron inmunidad protectora cuando se administraron a ratones (Konishi y otros, *Virology* 180: 401-410 (1991)). Se mostró que las células HeLa infectadas con virus de vaccinia recombinante que portaban genes de prM y E de JEV producían partículas subvirales (Konishi et al., *Virology* 188: 714-720 (1992)). Dmitriev y otros reportaron inmunización de ratones con un virus de vaccinia recombinante que codifica proteínas estructurales y determinadas proteínas no estructúrales de virus de encefalitis transmitido por garrapatas (*J. Biotechnology* 44: 97-103 (1996)).

Los vectores de virus recombinantes también se han preparado para servir como vacunas de virus para fiebre dengue. Zhao y otros (*J. Virol.* 61: 4019-4022 (1987)) prepararon virus de vaccinia recombinante que lleva proteínas estructurales y NS1 del dengue serotipo 4 y lograron expresión después de infectar células de

mamíferos con el virus recombinante. Se obtuvo expresión similar utilizando baculovirus recombinante para infectar células de insectos meta (Zhang et al., *J. Virol.* 62: 3027-3031(1988)). Bray et al. (*J. Virol.* 63: 2853-2856 (1989)) también reportaron una vacuna de dengue de vaccinia recombinante basada en el gen de proteína E que confiere inmunidad protectora a ratones contra encefalitis de dengue cuando se ponen a prueba. Falgout et al. (*J. Virol* 63: 1852-1860 (1989)) y Falgout et al. (*J. Virol.* 64: 4356-4363 (1990)) reportaron resultados similares. Zhang et al. (*J. Virol* 62: 3027-3031 (1988)) mostraron que el baculovirus recombinante que codifica dengue de proteínas E y NS1 igualmente protegió a los ratones contra la encefalitis de dengue cuando se pusieron a prueba. Otras combinaciones en las que se incorporaron genes estructurales y no estructurales en las vacunas de virus recombinante no produjeron inmunidad significativa (Bray et al., *J. Virol.* 63: 2853-2856 (1989)). Igualmente, los monos no desarrollaron inmunidad totalmente protectora para el reto del virus de dengue cuando se inmunizaron con baculovirus recombinante que expresaba la proteína E (Lai y otros (1990) pp. 119-124 en F. Brown, R. M. Chancock, H. S. Ginsberg y R. Lerner (eds.) Vaccines 90: Modern approaches to new vaccines including prevention of AIDS, Laboratorio Cold Spring Harbor, Cold Spring Harbor, NY).

La inmunización que utiliza preparaciones de DNA recombinante se ha reportado para virus SLEV y dengue-2, utilizando ratones destetados como el modelo (Phillpotts et al., *Arch. Virol.* 141: 743-749 (1996); Kochel et al., *Vaccine* 15: 547-552 (1997)). El DNA plasmida que codifica los genes prM y E del SLEV proporcionó protección parcial contra el reto SLEV con una dosis simple o doble de inmunización de DNA. En estos experimentos, los ratones de control exhibieron alrededor del 25% de supervivencia y no se detectó ningún anticuerpo protector en los ratones inmunizados con DNA (Phillpotts et al., *Arch. Virol.* 141: 743-749 (1996)). En ratones que recibieron tres inyecciones intradérmicas de plasmida de DNA de dengue-2 recombinante que contenían prM, el 100% desarrolló anticuerpos neutralizantes anti-dengue-2 y el 92% de aquellos que recibieron el correspondiente gen E desarrollaron también anticuerpos neutralizantes (Kochel et al., *Vaccine* 15: 547-552 (1997)). Sin embargo, los experimentos retantes que utilizaron un programa de dosis no protegieron a los ratones contra el reto letal del virus del dengue-2.

Las vacunas desarrolladas hasta la fecha para inmunización frente a infección por JEV, SLEV, virus dengue y otros flavivirus tienen una serie de desventajas y problemas que se deben a su uso. La vacuna inactivada es costosa e inconveniente de preparar. Además,

000193

cualquiera de esas vacunas conlleva el riesgo de reacciones alérgicas que provienen de las proteínas de la célula huésped utilizada en la preparación del virus. Además, esas vacunas presentan un riesgo considerable para los trabajadores que se dedican a su producción. Las vacunas candidatas JEV atenuadas están siendo objeto de pruebas clínicas, pero desde 1996 no han encontrado una amplia aceptación fuera de la República Popular China (Hennessy et al., *Lancet* 347: 1583-1586 (1996)).

Las vacunas recombinantes basadas en la utilización de solamente ciertas proteínas de flavivirus, tales como el JEV, producidas por expresión biosintética en el cultivo de células con purificación posterior o tratamiento de antígenos, no inducen altas titulaciones de anticuerpos. Igualmente, tal como las preparaciones de virus enteros, estas vacunas comportan el riesgo de reacción alérgica adversa a los antígenos desde el huésped o hasta el vector. El desarrollo de vacunas contra el virus del dengue y el WNV es menos avanzado y las vacunas basadas en el virus o basadas en proteínas recombinantes presentan problemas similares a los que se aludió arriba.

Por consiguiente, existe la necesidad de vacunas mejoradas o de vacunas encaminadas hacia los flavivirus, tal como el virus de fiebre amarilla, el virus de dengue, el JEV, el SLEV y el WNV que no sean costosas de

preparar, presenten poco riesgo para los trabajadores involucrados en su manufactura, comporten riesgo mínimo de reacciones inmunológicas adversas debido a impurezas o a componentes inmunogénicos adventicios y sean altamente eficaces en producir anticuerpos neutralizantes e inmunidad protectora. Existe además la necesidad de una vacuna contra el JEV, el WNV y los flavivirus relacionados que minimice el número de dosis inmunizantes requeridas

Muchas de las desventajas de la técnica actual, tal como se describió en detalle para la producción de las vacunas, también se aplican a la producción de antígenos y anticuerpos que se utilizan para la producción de los inmunodiagósticos. Particularmente, los riesgos concurrentes y los costos involucrados en la producción de antígenos a partir de los virus y la falla de la mayoría de los antígenos expresados recombinantemente actualmente disponibles para producir respuestas inmunes eficaces, tienen su paralelo en el campo del inmunodiagnóstico por los mismos riesgos, altos costos y una correspondiente carencia de sensibilidad. Así pues, debido a los altos costos, al riesgo de la infección accidental con los virus vivos y a los niveles inferiores a los deseados de sensibilidad de las pruebas disponibles anteriormente, existe una necesidad de pruebas de diagnóstico rápidas, simples y

altamente sensibles para detectar la infección y/o contaminación del flavivirus.

El presente invento satisface todas estas necesidades proporcionando antígenos recombinantes altamente inmunogénicos para utilización en ensayos de diagnóstico con el objeto de detectar anticuerpos para los flavivirus seleccionados. El presente invento proporciona además la utilización de antígenos recombinantes derivados de flavivirus, genes de flavivirus o imitaciones de los mismos en ensayos inmunodiagnósticos para la detección de anticuerpos con respecto a las proteínas de flavivirus.

Resumen del Invento

El presente invento proporciona una molécula de ácido nucléico que contiene una unidad transcripcional (TU) para un antígeno de flavivirus inmunogénico. La TU dirige una célula huésped después de haber sido incorporada en la misma, para sintetizar el antígeno. En un aspecto importante del invento, el flavivirus puede ser un virus de fiebre amarilla (YFV), virus de dengue serotipo 1 (DEN-1), virus de dengue serotipo 2 (DEN-2), virus de dengue serotipo 3 (DEN-3), virus de dengue serotipo 4 (DEN-4), virus de encefalitis St. Louis (SLEV), virus de encefalitis japonesa (JEV), virus West Nile (WNV), virus Powassan o cualquier otro flavivirus. En incorporaciones importantes del presente invento, el

antígeno puede ser el flavivirus de proteína prM/M, la proteína E, o ambos. En incorporaciones importantes del presente invento, el antígeno puede ser una proteína quimérica de flavivirus. En particular cuando la unidad transcripcional (TU) incluye las proteínas tanto de prM/M como de E, la célula huésped secreta partículas subvirales que contienen los antígenos prM/M y E. En otros aspecto importante del invento, el ácido nucléico es una molécula de DNA. En incorporaciones adicionales significativas, la TU de ácido nucléico incluye una secuencia de control dispuesta apropiadamente, en forma tal que controla operablemente la expresión de los antígenos prM/M y E y esta secuencia de control puede ser el promotor temprano inmediato de citomegalovirus. En otra incorporación adicional, la secuencia nucleótida de la TU está diseñada para optimizar la translación eucariótica minimizando amplias estructuras de orquilla en la región no translacionada del extremo 5' de un mRNA producido por la TU y/o la inclusión de una secuencia de consenso Kozak en el lugar del inicio translacional de un mRNA producido por la TU. En una incorporación adicional, la unidad transcripcional también incluye un terminador poli-A.

El presente invento involucra además una célula huésped que comprende una molécula de ácido nucléico que incluye una unidad transcripcional para un antígeno de

flavivirus inmunogénico que dirige la célula huésped para sintetizar el antígeno inmunogénico. El flavivirus puede ser YFV, DEN-1, DEN-2, DEN-3, DEN-4, SLEV, JEV, WNV, virus Powassan u otro flavivirus. En incorporaciones importantes, el antígeno puede ser la proteína prM/M, la proteína E, o ambas proteínas. En el último caso, la célula secreta partículas subvirales que contienen antígenos prM/M y E.

Además, el invento proporciona una composición para vacunar a un sujeto contra un flavivirus que contiene una molécula de ácido nucléico que incluye una unidad transcripcional para un antígeno flaviviral inmunogénico. La unidad transcripcional dirige una célula dentro del cuerpo del sujeto, después de haber sido incorporada al mismo, para sintetizar el antígeno inmunogénico. La composición incluye además, un portador aceptable desde el punto de vista farmacéutico. En incorporaciones significativas, el flavivirus puede ser YFV, DEN-1, DEN-2, DEN-3, DEN-4, SLEV, JEV, WNV, virus Powassan y otros flavivirus. Además, el antígeno puede ser la proteína prM/M, la proteína E, o ambas proteínas. En el último caso, la célula secreta partículas subvirales que comprenden antígenos prM/M y E del flavivirus. Estas partículas subvirales también se denominan antígeno recombinante no infeccioso (NRA). En incorporaciones importantes, la molécula de ácido nucléico es una

molécula de DNA. En otras incorporaciones significativas, la unidad transcripcional adicionalmente contiene una secuencia de control dispuesta apropiadamente de manera que controle en forma operable la síntesis de los antígenos prM/M y E cuando el ácido nucléico se introduce en la célula del sujeto. Esta secuencia de control puede ser el promotor precoz inmediato del citomegalovirus. Inclusive en otra incorporación, la unidad transcripcional también puede incluir un terminador poli-A.

Las composiciones que proporciona el presente invento para vacunar a un sujeto contra un flavivirus pueden incluir una molécula o moléculas de ácido nucléico que incluyen unidades transcripcionales para más de un antígeno flaviviral inmunogénico. El más de un antígeno flaviviral inmunogénico puede ser de diferentes especies, cepas o aislados de flavivirus, en cualquier combinación. En incorporaciones significativas, los flavivirus incluidos pueden ser dos o más, tres o más, cuatro o más, cinco o más o siete o más flavivirus. Los ejemplos de esos flavivirus incluyen YFV, DEN-1, DEN-2, DEN-3, DEN-4, SLEV, JEV, WNV, virus Powassan y otros flavivirus, sin que estos impliquen limitación. Se pueden formular vacunas de combinación para conferir inmunidad a la enfermedad de flavivirus común a regiones geográficas particulares. En una incorporación particular dirigida hacia el Asia

tropical y subtropical, se pueden seleccionar los virus DEN-1, DEN-2, DEN-3, DEN-4,WNV y JE. En una incorporación particular dirigida al África, se pueden seleccionar DEN-1, DEN-2, DEN-3, DEN-4, WN y YF. En una incorporación particular dirigida a América Latina se pueden seleccionar los virus DEN-1, DEN-2, DEN-3, DEN-4, Rocío y YF.

El invento proporciona además un método para inmunizar a un sujeto contra la infección causada por un flavivirus. El método involucra administrar al sujeto una cantidad efectiva de una composición de vacuna que contenga una molécula de ácido nucléico que incluya una unidad transcripcional para un antígeno de flavivirus inmunogénico. La unidad transcripcional dirige una célula dentro del cuerpo del sujeto, después de haber sido tomada por la célula, para sintetizar el antígeno inmunogénico. La composición adicionalmente incluye un portador aceptable desde el punto de vista farmacéutico. En incorporaciones significativas del método, el flavivirus puede ser YFV, DEN-1, DEN-2, DEN-3, DEN-4, SLEV, JEV, WNV, virus Powassan u otros flavivirus. En otros aspectos importantes del método, el antígeno puede ser la proteína prM/M, la proteína E, o ambas proteínas. Cuando el antígeno es tanto la proteína E como la proteína prM/M, la célula dentro del cuerpo del sujeto, después de incorporar el ácido nucléico dentro de él,

secreta partículas subvirales que comprenden los antígenos E y prM/M flavivirales. Adicionalmente, en incorporaciones significativas del método, la composición de vacunación se administra al sujeto en una sola dosis, vía parenteral. En otro aspecto adicional del método, el ácido nucléico es una molécula de DNA. En otras incorporaciones adicionales del método, la unidad transcripcional además incluye una secuencia de control dispuesta apropiadamente de manera tal que ésta controla en forma operable la síntesis de los antígenos prM/M y E y en un aspecto significativo de esta incorporación, la secuencia de control es el promotor precoz inmediato del citomegalovirus. Además, la unidad transcripcional puede incluir un terminador poli-A

Estos aspectos e incorporaciones del invento constituyen la base para sus claros atributos y ventajas. Siendo una estructura de ácido nucléico que involucra solamente porciones del genoma de flavivirus más bien que la secuencia que abarca el genoma completo, la vacuna que contiene TU de ácido nucléico es completamente no viable. Por consiguiente no presenta peligro para infección por el flavivirus a aquellos involucrados en su manufactura o para los sujetos que reciben la vacuna. La vacuna de ácido nucléico es fácil de preparar y fácil de administrar y es estable en el almacenamiento anterior a su uso. Inesperadamente se ha encontrado que la vacuna de

ácido nucléico del invento es esencialmente 100% exitosa en conferir inmunidad protectora a mamíferos después de administrar solamente una dosis simple. Otro resultado adicional inesperado es que la TU de ácido nucléico puede engendrar inmunidad a un flavivirus en un mamífero hembra que se puede transmitir a su progenie a través de la leche. Sin desear quedar limitados por la teoría, el inventor considera que un posible mecanismo para el éxito del ácido nucléico en conferir inmunidad protectora es que la célula huésped que alberga el ácido nucléico, tal como la célula de un sujeto a quien se le administre la vacuna, produce partículas subvirales que contienen los antígenos prM/M y E flavivirales. Estas partículas limitan los atributos inmunogénicos de los viriones de flaviviris nativo.

El presente invento también proporciona polipéptidos antigénicos no infecciosos, fragmentos polipéptidos antigénicos y NRA que comprende las proteínas prM/M y/o E de los flavivirus, en donde la secuencia de la señal transmembrana se deriva de un primer flavivirus y las proteínas M y/o E se derivan de un segundo flavivirus. Además, la proteína prM/M puede comprender secuencias de aminoácidos tanto del primero como del segundo flavivirus. Además, la proteína E puede comprender secuencias de aminoácidos tanto del primero como del segundo flavivirus. "Quimérico" tal como se utiliza en

000202

este documento significa cualquier proteína o ácido nucléico que comprende la secuencia de más de un flavivirus. Tal como se utiliza en este invento, "no virulento" significa que el antígeno o la vacuna de este invento es incapaz de causar enfermedad. Más particularmente, los antígenos proteínicos recombinantes están libres de material genómico contaminante de los flavivirus que es necesario para la infección, replicación y patogénesis del flavivirus.

Los polipéptidos del presente invento pueden comprender las secuencias del aminoácido definido en este documento, o de los que se conocen en la técnica, de las proteínas prM, M y/o E de los flavivirus seleccionados. Los ácidos nucléicos de este invento pueden comprender la secuencia nucleótida que codifica las proteínas prM, M y/o E de los flavivirus seleccionados.

Los antígenos del presente invento pueden estar conjugados, o se pueden conjugar a una molécula portadora que facilita la colocación del antígeno en una fase sólida. Una molécula portadora es aquella con la cual se pueden conjugar los antígenos y que no reaccionará con los anticuerpos en el suero humano. Un ejemplo de esta portadora es la albúmina de suero bovino (BSA, por sus iniciales en inglés).

Los antígenos del presente invento también pueden ser proteínas recombinantes obtenidas mediante la expresión de ácidos nucléicos que codifican el antígeno en un sistema de expresión capaz de producirlo.

Las secuencias de aminoácido de los antígenos presentes pueden contener una porción inmuno-reactiva del antígeno prM, M y/o E. Estos antígenos pueden además estar adjuntos a secuencias diseñadas para proporcionar alguna propiedad adicional, tal como remover/añadir aminoácidos capaces de ligamento con disulfuro para aumentar la reactividad de un epítope proporcionando una estructura secundaria más rígida, para aumentar su biolongevidad o alterar su citotoxicidad o impedir infección. En cualquier caso, el antígeno debe poseer inmuno-reactividad y/o inmunogenicidad.

Breve Descripción de los Dibujos

La Figura 1 es una representación esquemática del procesamiento de la poliproteína flaviviral. La región horizontal central proporciona una representación esquemática del genoma viral. Las líneas denotan las regiones no transferidas 5' y 3' y las regiones encerradas representan el marco de lectura abierto para las proteínas estructurales (izquierda y superior) y no estructurales (derecha e inferior). La separación por parte de la signalasa de la célula huésped ocurre

000204

simultáneamente con la translación en la proteína E terminal-C, dividiendo las regiones estructurales y no estructurales. Una enzima celular parecida a la subtilasa, la furina, puede ser responsable por la división del prM. Los dominios de la transmembrana potencial de la poliproteína viral se indican en las áreas sombreadas.

La Figura 2 es un mapa del genoma JEV (superior) y de la secuencia del DNA de los oligonucleótidos utilizados en una reacción inversa de cadena de polimerasa-transcriptasa (RT-PCR) (centro) para construir la unidad de la trascripción para la expresión de las regiones de codificación de la proteína prM-E (inferior). Los dominios de la transmembrana potencial de la poliproteína viral se indican en las áreas sombreadas.

La Figura 3 muestra una representación esquemática de los vectores plasmidas, pCDNA3, pCBamp, y pCIBamp, y la relación entre ellos. Estos plasmidas incluyen el elemento promotor/realzardor CMV (citomegalovirus), el BGHp(A) (señal de poliadenilación de la hormona de crecimiento bovino y secuencia de terminación de transcripción), el gen de la resistencia a la ampicilina y el origen ColE1 de la replicación para selección y mantenimiento en *E. coli*. El origen f1 de la replicación para el rescate de una sola ramificación en células *E. coli*, el origen SV40 de la replicación (SV40 ORI), la región de

codificación de la resistencia a la neomicina y las secuencias SV40p(A) que se suprimieron del pCDNA3 para generar pCBamp. Se insertó una secuencia intrón en el sitio NcoI-KpnI de pCBamp para generar plasmida pCIBamp.

La Figura 4 muestra análisis inmunosecantes (imnunoblot) SDS-PAGE de las partículas subvirales purificadas del gradiente de sucrosa a partir del fluido del cultivo JE-4B COS-1 (4B, línea derecha de cada par). Se utilizó el virión JE purificado del gradiente de densidad del cultivo de la célula C6/36 infectada con JEV como un control positivo (JEV, línea izquierda de cada par). JE HIAF (fluido ascítico hiperinmune); 4G2, anticuerpo monoclonal anti-E; JM01, anticuerpo monoclonal anti-M; NMAF (fluido ascítico normal de ratones).

La Figura 5 muestra un perfil del antígeno E en un análisis de rata de gradiente de sucrosa zonal preparado a partir del precipitado PEG del medio de cultivo de la célula JE-4B con o sin tratamiento Tritón X-100.

La Figura 6 muestra la probabilidad de la señal péptida del pCBJE1-14 (pCBJE) predicado por el programa de la Señal P-HMM (A). La probabilidad de la señal péptida se mejora alternado la secuencia de la región-c en las posiciones -4 y -2 (C-4G y G-2S) (panel B, JE-LSS-M), acortando la región-n (panel C, JE-SS-ORI), o

mediante una combinación de ambas modificaciones (panel D, JE-SS-M).

La Figura 7 muestra representaciones esquemáticas de los vectores plasmidas pCBD2-14-16 (100% DEN-2 E), pCBD-1J-4-3- (90% DEN-2 E: 10% JEV E), y pCB8d2-2J-2-9-1 (80% DEN-2 E: 20% JEV E). Estos plasmidas incluyen el promotor del gen temprano del citomegalovirus humano (CMV); la secuencia de la señal del virus JE; la región del gen prM y E del virus DEN-2 (aminoterminal 100%, 90% u 80%, respectivamente); la región del gen E del virus JE (ninguno, 10% o 20%, respectivamente); y la señal poli A de la hormona de crecimiento bovino (BHG).

La Figura 8 muestra comparación de la proteína recombinante secretada y ligada con membrana mediante secante occidental (western blot). (A) Análisis del antígeno recombinante secretado después de la precipitación PEG y de la extracción de etanol del fluido de cultivo para los plasmidas DEN-2 pCB8D2-2J-2-9-1, pCB9D2-1J-4-3, pCBD2-14-16, y la plasmida de control pEGFP. La Linea 1 (V), virus DEN-2 purificado manchado por Gold Blot (Owl Separation Systems, Portsmouth, NH).

La reactividad del antígeno secretado, recombinante de cada plasmida con a, antienvolvente (E) Mab 1A6A-8 específico; b, una mezcla del Mab 1A6A-8, anticápsido (C) Mab 1A2A-1 específico, antisuero específico para la

proteína (prM) de la premembrana del virus DEN-2; y **c**, ascitis normal de ratón. (B) Análisis de proteínad de membrana hidrofóbica de célula transformada-plasmida recombinante. Línea 1 (V), virus DEN-2 purificado manchado con Gold Blot; Línea 2 (V), reactividad del virus purificado DEN-2 con una mezcla de Mab 1A6A-8, Mab 1A2A-1, antisuero específico para la proteína M del virus DEN-2, y antisuero para la proteína prM del virus DEN-2. La reactividad de las proteínas de la membrana hidrofóbica aislada de cada línea de célula transformada-plasmida con **a**, Mab 1A6A-8; **b**, una mezcla de Mab 1A6A-8, Mab 1A2A-1, antisuero específico para la proteína M del virus DEM-2, y antisuero para la proteína prM del virus DEM-2, y **c**, ascitis normal de ratón.

Descripción Detallada del Invento

El invento abarca unidades transcripcionales de ácido nucléico que codifican proteínas antigénicas flavirales, tales como los antígenos de proteína prM/M y E. Los ácidos nucléicos funcionan para expresar los antígenos de proteína prM/M y E cuando el ácido nucléico es tomado por una célula apropiada, especialmente cuando la célula es la célula de un sujeto. El invento también abarca una vacuna cuyo agente activo es la unidad transcripcional de ácido nucléico (TU). El invento abarca además células que contienen una TU. El invento abarca también un método de inmunizar un sujeto contra la

208

infección flaviviral administrándole una cantidad eficaz de una vacuna que contiene las moléculas de la TU del ácido nucléico.

El invento proporciona un ácido nucléico aislado que comprende una unidad transcripcional que codifica una secuencia de señal de una proteína estructural de un primer flavivirus y un antígeno de flavivirus inmunogénico de un segundo flavivirus, en donde la unidad transcripcional dirige la síntesis del antígeno. El invento abarca además la utilización de la unidad transcripcional (TU) del ácido nucléico para generar antígenos flavivirales y los antígenos flavivirales producidos por la TU del ácido nucléico. Los antígenos flavivirales abarcados por el presente invento incluyen antígenos de flavivirus quimérico que incorporan la secuencia del aminoácido de un primer flavivirus y por lo menos un flavivirus adicional. El invento abarca también la utilización de los antígenos flavivirales codificados por la TU del invento para producir anticuerpos específicos de flavivirus y para detectar la presencia de anticuerpos específicos de flavivirus.

En una incorporación, el ácido nucléico aislado de este invento puede comprender una unidad transcripcional que codifica una secuencia de señal de virus de encefalitis japonesa.

flavivirus. En una incorporación preferida, el ácido nucléico aislado codifica tanto la proteína M como la proteína E del flavivirus. Además, el ácido nucléico del invento puede ser DNA y puede comprender una secuencia nucleótida SEQ ID NO:15, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23 o SEQ ID No. 42.

En otra incorporación particular, el ácido nucléico de este invento puede codificar una secuencia de señal del virus de encefalitis japonesa, una proteína M de un segundo virus y una proteína E quimérica formada por sustitución de una porción del ácido nucléico que codifica la proteína E del segundo virus con ácido nucléico que codifica la porción correspondiente de la proteína E del JEV. Alternativamente, la porción de la secuencia que corresponde a la parte suprimida de la proteína E del segundo virus se puede sustituir con otra secuencia seleccionada de un tercer virus, o puede ser una secuencia no viral. La segunda proteína puede ser virus West Nile, SLEV, YFV, virus Powassan y/o un serotipo del virus dengue. Las proteínas E quiméricas pueden incluir aquellas en donde la porción terminal carboxi puede ser de un flavivirus y la parte restante de la proteína E quimérica es de otro flavivirus. La porción terminal carboxi puede ser, por ejemplo, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60 o 75% de la proteína E quimérica. El ácido nucléico del invento puede ser DNA o

puede comprender la secuencia de proteína codificada a partir de la secuencia nucleótida SEQ ID NO: 44 o SEQ ID NO: 46. El ácido nucléico del invento puede comprender la secuencia nucleótida SEQ ID NO:44 o SEQ ID NO: 46.

La unidad transcripcional de este invento también puede comprender una secuencia de control dispuesta apropiadamente de manera que ésta controle operablemente la síntesis del antígeno. La secuencia de control puede ser, por ejemplo, el promotor temprano inmediato del citomegalovirus. El ácido nucléico de este invento también puede comprender una secuencia de consenso Kozak localizada en el sitio de iniciación translacional para un polipéptido que comprende el antígeno codificado por la unidad transcripcional. La unidad transcripcional de este invento también puede comprender un terminador poli-A.

El presente invento proporciona además una célula que comprende el ácido nucléico de este invento.

También se suministra una composición que comprende un portador aceptable desde el punto de vista farmacéutico y ácido nucléico o célula o antígeno de este invento. El invento adicionalmente proporciona un método para inmunizar a un sujeto contra infección por un flavivirus, que comprende administrarle al sujeto una cantidad eficaz de una composición de este invento. En una incorporación particular, la composición utilizada

para inmunizar a un sujeto dirige la síntesis tanto de la proteína M como de la proteína E de un flavivirus y de una célula dentro del cuerpo del sujeto, después de incorporar el ácido nucléico dentro de él, secreta partículas subvirales que comprenden la proteína M y la proteína E. Alternativamente, la composición puede comprender una proteína M y/o una proteína E de un flavivirus o partículas subvirales que comprenden la proteína M y la proteína E. En los métodos de este invento, la composición inmunizante se puede administrar a un sujeto en una sola dosis y se puede administrar vía parenteral.

Este invento proporciona además los antígenos producidos a partir de los ácidos nucléicos aislados del invento. Como un ejemplo, el antígeno del segundo flavivirus codificado por la secuencia nucleótida de la TU puede ser la proteína M, que puede ser por ejemplo del virus West Nile. El antígeno también puede ser proteína del virus dengue, el virus de encefalitis St. Louis, el virus de encefalitis japonesa, el virus Powassan y/o el virus de fiebre amarilla. En otra incorporación, el antígeno comprende una proteína prM/M que comprende la secuencia de la señal de la transmembrana de un primer flavivirus y además la secuencia del aminoácido que comprende el resto de la proteína prM/M de un segundo flavivirus, que puede ser del virus SLEV, JEV, YFV, WNV

y/o del virus Powassan. La secuencia de la señal de la transmembrana de un primer flavivirus puede ser una secuencia de señal mejorada o modificada en donde la secuencia de señal imparte las características deseadas, tal como una probabilidad de secuencia de señal alta. Para lograr estas metas mediante el diseño o la selección se pueden utilizar programas de computador de aprendizaje con máquina incluyendo aquellos que utilizan el modelo oculto Markov, sin que éste implique limitación.

El antígeno codificado por la secuencia nucléotida de la TU puede ser antígeno de virus West Nile, antígeno de virus dengue, antígeno de virus de encefalitis St. Louis, antígeno de virus de encefalitis japonesa, antígeno de virus Powassan y/o antígeno de virus de fiebre amarilla.

El antígeno codificado por la secuencia nucléotida de la TU también puede ser la proteína E, que puede ser la proteína E del virus West Nile, el virus dengue, el virus de la encefalitis St. Louis, el virus de la encefalitis japonesa, el virus Powassan y/o el virus de la fiebre amarilla. El antígeno codificado también puede ser una proteína E quimérica que comprende la secuencia del aminoácido seleccionada de más de un flavivirus.

Además, el antígeno codificado por la secuencia nucleótida de la TU puede ser la proteína M y la proteína

E, que pueden provenir del virus West Nile, el virus dengue, el virus de la encefalitis St. Louis, el virus de la encefalitis japonesa, el virus Powassan y/o el virus de la fiebre amarilla.

Tal como aquí se utilizan, "proteína M" o "proteína pr/M" o "proteína prM/M" significa una proteína M de flavivirus o una proteína prM de flavivirus. Los ejemplos incluyen proteínas prM que comprenden una secuencia de aminoácido de una o más proteínas prM de flavivirus, proteínas M que no comprenden una secuencia adicional de aminoácido, y proteínas que comprenden secuencias adicionales de aminoácido que se procesan *in vitro* o *in vivo* para generar la proteína M madura, sin que lo enunciado constituya limitación.

Tal como aquí se utilizan, "unidad transcripcional de ácido nucléico" o molécula de la unidad transcripcional de ácido nucléico" significan un ácido nucléico que codifica una o más proteínas especificadas. La TU tiene actividad biológica de manera que, después de haber sido introducida en una célula adecuada, el ácido nucléico induce la síntesis de uno o más productos de genes especificados codificados por el ácido nucléico. El(los) producto(s) de genes es(son) otras macromoléculas biológicas, tales como proteínas, no químicamente relacionadas con la unidad transcripcional (TU). La TU del ácido nucléico induce a la célula a emplear sus

componentes celulares para producir el producto o productos de gen específico codificados por el ácido nucléico de la TU. Aún cuando cualquier ácido nucléico puede servir como una TU, en una incorporación preferida, la TU es el DNA de una plasmida o vector similar, en donde la plasmida o vector comprende codificar secuencias de genes marcadores u otras construcciones de secuencia que faciliten la utilización de la TU para experimentación y biosíntesis.

Tal como se utiliza en este documento, una "secuencia de control" es una secuencia nucleótida regulatoria incorporada dentro de una TU que interactúa con los componentes celulares apropiados de la célula y conduce a biosíntesis mejorada o activada de los productos del gen codificado por la TU. Por consiguiente, una secuencia de control adecuada es aquella con la que los componentes de la célula tienen la capacidad de interactuar, originando la síntesis del producto de gen. Cuando está dispuesta en forma que se puede operar en un ácido nucléico con respecto a una secuencia de codificación especificada, una secuencia de control controla eficazmente la expresión del ácido nucléico especificado para producir el producto de gen.

Tal como se utiliza en este documento, un "promotor" es una secuencia nucleótida en una TU que sirve como una secuencia de control.

Tal como se utiliza en este documento, una "secuencia Kozak" o una "secuencia de consenso Kozak" es una secuencia nucleótida en el sitio de iniciación translacional que optimiza la translación de los mRNAs eucarióticos (Kozak, *Mol. Cell. Biology* 9: 5134-5142 (1989)) .

Tal como se utiliza en este documento, un "terminador" es una secuencia nucleótida extendida que actúa para inducir la poliadenilación en el extremo 3' del mRNA maduro. Una secuencia de terminador se encuentra después, o corriente abajo de una secuencia de codificación particular.

Tal como se utiliza en este documento, una "célula" es una célula eucariótica o procariótica que comprende una unidad transcripcional (TU) que codifica uno o más productos de genes, o en la cual se ha introducido la TU. Por consiguiente, una célula alberga una sustancia heteróloga o foránea, la TU, que no se encuentra endógena o naturalmente en la célula como un componente. Una célula adecuada es aquella que tiene la capacidad para la biosíntesis de los productos del gen como consecuencia de la introducción de la TU. En particular, la célula idónea es aquella que responde a una secuencia de control y a una secuencia de terminador, si la hubiere, que puede estar incluida dentro de la TU. En incorporaciones

importantes del presente invento, la célula es una célula de un mamífero. En incorporaciones particularmente importantes del invento, la célula es una célula que ocurre naturalmente en el cuerpo de un humano o de un sujeto no humano al que se le ha administrado la TU como un componente de una vacuna. Alternativamente, en aplicaciones analíticas o diagnósticas, incluyendo la preparación de antígeno para utilización como vacuna o en ensayos inmunodiagnósticos, o para propósitos demostrativos, la célula puede ser una célula humana o no humana cultivada *in vitro*.

Tal como se utiliza en este documento, una "vacuna" o una "composición para vacunar a un sujeto" específico para un patógeno particular significa una preparación que, al administrarse al sujeto, conduce a una respuesta inmunogénica en el mismo. Tal como se utiliza en este documento, una respuesta "inmunogénica" es aquella que confiere al sujeto una inmunidad protectora contra el patógeno. Sin desear quedar atados a esta teoría, se considera que una respuesta inmunogénica puede surgir de la generación de anticuerpos neutralizantes (v.g. una respuesta inmune humoral) o de células citotóxicas del sistema inmune (v.g. una respuesta inmune celular) o ambas. Tal como se utiliza en este documento, un "antígeno inmunogénico" es un antígeno que induce una respuesta inmunogénica cuando se introduce en un sujeto,

o cuando se sintetiza dentro de las células de un anfitrión o de un sujeto. Tal como se utiliza en este documento, una "cantidad eficaz" de una vacuna o de una composición de vacuna es una cantidad que, al administrarse a un sujeto, es suficiente para conferir inmunidad protectora al mismo. Históricamente, se ha entendido que una vacuna contiene como principio activo uno o más componentes moleculares específicos o estructuras que comprenden el patógeno, especialmente su superficie. Esas estructuras pueden incluir componentes de superficie tales como proteínas, carbohidratos complejos y/o lípidos complejos que comúnmente se encuentran en los organismos patogénicos.

Tal como se utiliza en este documento, sin embargo, se debe hacer énfasis en que los términos "vacuna" o "composición para vacunar a un sujeto" prolongan el significado convencional resumido en el parágrafo precedente. Tal como aquí se utilizan, estos términos también se relacionan con la TU del presente invento o con composiciones que contienen la TU. La TU induce la biosíntesis de uno o más productos de genes especificados codificados por la TU dentro de las células del sujeto, en donde los productos de genes son antígenos especificados de un patógeno. Los antígenos biosintéticos sirven entonces como un inmunogen. Tal como ya se anotó, la TU, y por consiguiente la vacuna, puede ser cualquier

ácido nucléico que codifique los antígenos inmunogénicos especificados. En una incorporación preferida de este invento, la TU de la vacuna es DNA. La TU puede incluir una plasmida o vector que incorpore genes adicionales o secuencias particulares para la conveniencia del trabajador que conoce la técnica en los campos de biología molecular, biología de células e inmunología viral. (Ver Molecular Cloning: A Laboratory Manual, 2ª Edición, Sambrook, Fritsch y Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; y Current Protocols in Molecular Biology, Ausubel et al., John Wiley y Sons, New York 1987 (actualizados trimestralmente), los cuales se incorporan a este documento por referencia).

Las moléculas de la TU del presente invento comprenden ácidos nucléicos o derivados de ácidos nucléicos que tienen secuencias nucleótidas que codifican productos de genes específicos relacionados con antígenos de flavivirus tales como WNV, JEV, virus dengue, virus de fiebre amarilla y SLEV, sin que estos impliquen limitación. Aún cuando cualquier ácido nucléico puede servir como una TU, en una incorporación importante, la TU es DNA. Alternativamente, los ácidos nucléicos pueden ser moléculas RNA. También pueden ser cualquiera de los muchos derivados de DNA o de RNA que tengan una espina dorsal de enlaces de fosfodiester que hayan sido

químicamente modificados para aumentar la estabilidad de la TU como agente farmacéutico. Las modificaciones así contempladas incluyen derivados de fosforotioato o derivados de fosfonato, sin que estos impliquen limitación. Estos y otros ejemplos de derivados son bien conocidos por las personas que conocen la técnica del campo de la química del ácido nucléico.

El genoma del JEV se ha caracterizado y secuenciado (Figura 1 y 2). La proteína estructural M se expresa como una porción de la poliproteína que incluye una secuencia pre-M (pr). Esta secuencia pr, inmediatamente aminoterminal a la secuencia de la proteína M, previene los problemas conformacionales en el procesamiento de la poliproteína. En particular, la presencia de la secuencia pr es importante para impedir un mal enlace de la proteína E. Por consiguiente, la presencia del prM permite que se ensamblen las partículas JEV. Una vez que se forma el virión o partícula, la secuencia pr se puede separar de la proteína prM para producir partículas de virus maduras que contengan proteínas M, aún cuando la separación de la proteína prM para producir proteína M no es necesaria para producir partículas infecciosas. Las secuencias prM de muchos flavivirus diferentes relacionados se dividen solo en un bajo grado, pero los flavivirus mismos no obstante son infecciosos. Ejemplos de los flavivirus relacionados con estructuras y

funciones genómicas similares incluyen WNV, YFV, virus dengue y SLEV, sin que estos impliquen limitación.

En una incorporación, la TU que codifica las proteínas M y E flavivirales en el presente invento es DNA. De acuerdo con la discusión del parágrafo precedente, esta DNA comprende una secuencia nucleótida que codifica la proteína M, comprende la secuencia pre-M, y una secuencia nucleótida que codifica la proteína E. De esta forma, los productos de gen pretendidos se capacitan para formar partículas subvirales dentro de la célula. La secuencia pre-M se puede luego dividir en una forma análoga a la que ocurre con respecto a los viriones repletos.

Para poder funcionar efectivamente *in vivo* como vacuna, es ventajoso incluir dentro de la TU una secuencia de control que tenga el efecto de mejorar o promover la transcripción de las secuencias nucleótidas que codifican los antígenos. La utilización de esos promotores es bien conocida para los técnicos en el campo de la biología molecular, la biología celular y la inmunología viral (Ver Molecular Cloning: A Laboratory Manual, 2ª Edición, Sambrook, Fritsch y Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989; y Current Protocols in Molecular Biology, Ausubel et al., John Wiley and Sons, New York 1987 (actualizados trimestralmente)). Cuando la TU se utiliza como una

vacuna en un huésped mamífero, el promotor a ser empleado es preferiblemente uno que opere eficazmente en las células mamíferas. Tal promotor está dispuesto con respecto a las secuencias de codificación desde donde la transcripción de va a promover, en una posición en la que éste puede promover operablemente esa transcripción. En una incorporación significativa del presente invento, este promotor es el promotor precoz de citomegalovirus. Además, en otra incorporación preferida del invento, las secuencias de codificación son seguidas, en el ácido nucléico de la TU, por una secuencia de terminador, (Sambrook et al.). Incorporaciones particulares del invento se relacionan con células tanto procarióticas como eucarióticas. Muchas secuencias del promotor se conocen como útiles en cualquiera de las células procarióticas o eucarióticas. (Ver Sambrook et al.)

Los ácidos nucléicos del invento pueden además incluir secuencias DNA conocidas como elementos inmunoestimulatorios para los expertos en la técnica. Ejemplos de estos elementos incluyen determinados motivos CpG en DNA bacteriano, sin que éste implique limitación (Sato et al., *Science* 273: 352-354 (1996); Klinman et al., *Vaccine* 17: 19-25 (1998)).

La preparación de la TU del invento se logra rápidamente mediante métodos bien conocidos para los expertos en la técnica del campo de la biología

000225

molecular. Los procedimientos involucrados se establecen por ejemplo en Molecular Cloning: A Laboratory Manual, 2ª Edición, Sambrook, Fritsch y Maniatis, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989 y Current Protocols in Molecular Biology, Ausubel et al., John Wiley y Sons, New York 1987 (actualizados trimestralmente). La molécula RNA flaviviral se puede aislar de una muestra de virus vivo mediante métodos ampliamente conocidos entre los virólogos familiarizados con flavivirus, por ejemplo, y con otros grupos de virus también. Los métodos utilizados con JEV se resumen en Kuno et al. (*J. Virol.* 72: 73-83 (1998)). El RNA se utiliza como una plantilla para la síntesis del cDNA utilizando transcriptasa inversa. Del cDNA, un fragmento que contiene la región codificada pre-M hasta E (Figura 2) se obtiene mediante digestión con nucleasas de restricción que se saben dividen el cDNA apropiadamente para proporcionar esos fragmentos. Ejemplos de digestión de restricción de JEV se proporcionan en Nitayaphan et al. (1990) y Konishi et al. (1991). La incorporación de promotores, tales como el promotor de citomegalovirus, las secuencias para promover translación eficiente, tal como la secuencia Kozak, y la de la señal de poliadenilación también son bien conocidas para los técnicos en biología molecular y la ingeniería recombinante DNA (Kozak, *Mol. Cell. Biology* 9: 5134-5142 (1989); Azevedo et al., *Braz. J. Med. Biol. Res.* 32: 147-

153 (1999)). Cuando se prepara un ácido nucléico que comprende una TU que contiene las secuencias de codificación deseada y las secuencias de control, éste se puede obtener en mayores cantidades con métodos que amplifican los ácidos nucléicos. Estos métodos son ampliamente conocidos para los técnicos en biología molecular y en ingeniería recombinante de DNA. Los ejemplos de estos métodos incluyen la incorporación del ácido nucléico en una plasmida para replicación mediante cultivo en una célula tal como una célula procariótica y cosechar la plasmida después de completar el cultivo, así como la amplificación del ácido nucléico mediante métodos tales como el PCR y otros protocolos de amplificación, bien conocidos en la técnica. Estos ejemplos no pretenden limitar las formas en que se obtiene el ácido nucléico que contiene la TU.

La TU que contiene moléculas de ácido nucléico del presente invento se puede introducir en células apropiadas en muchas formas bien conocidas por aquellos que conocen la técnica del campo de la biología molecular y la inmunología viral. A manera de ejemplo, estos incluyen la incorporación en un vector de ácido nucléico de plasmida o similar que es tomado por las células o la encapsulación dentro de las estructuras lípidas vesiculares tales como liposomas, especialmente liposomas que contiene lípidos catiónicos o adsorción a partículas

que se incorporan en la célula por endocitosis, sin que esto implique limitación.

En general, una célula de este invento es una célula procariótica o eucariótica que comprende una TU, o en la cual se ha introducido una TU. La TU del presente invento incluye la biosíntesis intracelular de los antígenos prM/M y E codificados. Una célula adecuada es aquella que tiene la capacidad para la biosíntesis de los productos de genes como consecuencia de la introducción del ácido nucléico. En incorporaciones particulares del invento, una célula adecuada es aquella que responde a una secuencia de control y a una secuencia de terminador, si la hubiere, que se puede incluir dentro de la TU. Con el fin de responder de esta forma, tal célula contiene dentro de ella componentes que interactúan con una secuencia de control y con un terminador y actúan para llevar a cabo las respectivas funciones de terminación y promoción. Cuando la célula se cultiva *in vitro*, ésta puede ser una procariota, una eucariota de una sola célula o una célula eucariota multicelular. En incorporaciones particulares del presente invento, la célula es una célula de un mamífero. En estos casos, los productos del gen de la proteína prM/M y E sintetizada están disponibles para ser utilizados en aplicaciones analíticas o diagnósticas, incluida la preparación del

antígeno para utilizarlo como una vacuna o en ensayos de inmunodiagnóstico, o para propósitos demostrativos.

En algunas circunstancias, tal como cuando la célula es una célula de mamífero cultivada, los antígenos prM/M y E se secretan en forma de partículas subvirales. Estos son los agregados de las proteína prM/M y E que se asemejan a virus vivos en las propiedades inmunogénicas y la morfología ultraestructural de la superficie. Sin embargo, como la TU del invento no incluye la parte restante del genoma flaviviral, no hay un cápsido incorporado, y lo que es más importante, no hay RNA viral infeccioso.

En otra incorporación importante de este invento, la célula es un componente celular natural del sujeto a quien se ha administrado la TU como una vacuna. La TU, al administrarse al sujeto, es tomada por las células del mismo. Las células del sujeto tienen la capacidad de responder a cualesquiera secuencias del promotor, y del terminador, si están presentes. En cualquier caso, la TU induce a las células del sujeto a sintetizar los productos de gen prM/M y E flaviviral. Sin desear estar constreñido por las consideraciones teóricas, se considera que las células del sujeto producen partículas subvirales *in vivo* que consisten de antígenos prM/M y E, tal como se ha encontrado que ocurre con las células mamíferas cultivadas *in vivo*, estipulando el sistema

000227

inmune del sujeto para generar repuestas inmunológicas que confieren inmunidad protectora al sujeto. Nuevamente, sin desear quedar limitados por la teoría, la inmunidad protectora resultante puede surgir vía humoral o a través de inmunidad celular, v.g., bien sea a través de un mecanismo restringido MHC clase-II, o clase I-, respectivamente, o con ambos mecanismos.

De acuerdo con el invento, los sujetos se inmunizan contra infección por flavivirus, tales como JEV, YFV, virus dengue, SLEV, WNV u otros flavivirus administrándoles una cantidad eficaz de una TU comprende ácido nucléico que codifica los antígenos prM y/o E. el ácido nucléico, después de incorporarse a las células del sujeto, conduce a la síntesis de los antígenos flavivirales prM/M y/o E.

Con el fin de administrar la TU al sujeto, se incorpora en una composición que comprende un portador aceptable desde el punto de vista farmacéutico. El término "aceptable desde el punto de vista farmacéutico" significa un material que no sea indeseable biológicamente o de cualquier otro punto de vista, v.g. el material se puede administrar a un sujeto junto con el material inmunógenico (v.g. los antígenos de la proteína de flavivirus recombinante o porciones de los mismos) sin causar ningún efecto biológico indeseable o interactuar en forma perjudicial con cualquier otro componente de la

vacuna en la cual está contenido. Los ejemplos de los portadores aceptables desde el punto de vista farmacéutico, o los componentes de los mismos, incluyen agua, salina fisiológica y tampones fisiológicos comunes (para ejemplos adicionales ver Arnon, R. (Ed.) Synthetic Vaccines I: pp. 83-92, CRC Press, Inc., Boca Ratón, Florida, 1987).

Quienes conocen la técnica entienden que el valor crítico para describir una dosis de vacuna es la cantidad total de inmunógeno necesario para producir una respuesta protectora en un huésped que esté sujeto a enfermedad infecciosa causada por flavivirus virulento o tipo silvestre. El número y volumen de dosis utilizadas puede variar y se determina por parte del médico basándose en parámetros tales como la edad, peso, género, especie, tipo de vacuna que se va a administrar, modo de administración, estado general del sujeto, etc., así como en otros factores importantes reconocidos por los expertos en la técnica.

La TU se puede administrar a un sujeto oralmente, parenteralmente (v.g. intravenosamente), mediante inyección intramuscular, inyección intraperitoneal, transdérmicamente, extracorpóreamente, intranasalmente, tópicamente, o en forma similar. La entrega también puede ser directamente a cualquier área del sistema respiratorio (v.g. los pulmones) mediante entubación. La

cantidad exacta de la TU que se requiere variará de sujeto a sujeto, dependiendo de la especie, edad, peso y estado general del mismo, la inmunogenicidad de la vacuna utilizada, la cepa o especie del flavivirus contra el cual se está inmunizando el sujeto, el modo de administración, y otros factores. Por consiguiente, no es posible especificar una cantidad exacta para cada incorporación del presente invento. Sin embargo, quien conozca la técnica puede determinar una cantidad apropiada utilizando solamente experimentación de rutina dadas las enseñanzas de este documento y lo que está disponible en la técnica.

La administración parenteral de la vacuna del presente invento, si se utiliza, generalmente se caracteriza por inyección. Los inyectables se pueden preparar en formas convencionales, bien sea como soluciones o suspensiones líquidas, formas sólidas adecuadas para solución o suspensión en líquido antes de la inyección, o como emulsiones. Un enfoque más recientemente revisado para administración parenteral involucra la utilización de un sistema de liberación sostenida o de liberación lenta, de manera tal que se mantenga una dosis constante. Ver por ejemplo la Patente Norteamericana No. 3.610.795, que se incorpora a este documento por referencia.

Para composiciones sólidas, los portadores sólidos no tóxicos convencionales incluyen, por ejemplo, grados farmacéuticos de manitol, lactosa, almidón, estearato de magnesio, sacarina sódica, talco, celulosa, glucosa, sucrosa, carbonato de magnesio y similares. Las composiciones líquidas que se pueden administrar farmacéuticamente pueden, por ejemplo, prepararse disolviendo, dispersando, etc. un componente activo, tal como se describe en este documento y adyuvantes farmacéuticos opcionales en un excipiente, tales como por ejemplo, agua, salina, dextrosa acuosa, glicerol, etanol, y similares, para formar así una solución o suspensión. Si se desea, la composición farmacéutica que se va a administrar también puede contener cantidades menores de sustancias auxiliares no tóxicas tales como agentes emulsionantes o humectantes, agentes de tampón de pH y similares, por ejemplo, acetato sódico, monolaurato de sorbitan, acetato sódico de trietanolamina, oleato de trietanolamina, etc. Los métodos actuales para preparar tales formas de dosis son conocidos, o serán aparentes para quienes conocen la técnica; por ejemplo ver Remington's Pharmaceutical Sciences (Martin, E.W. (ed.), última edición, Mack Publishing Co., Easton, PA).

En una incorporación la TU de este invento se puede administrar a un sujeto mediante la utilización de entrega de genes *in vivo* mediada por

electrotransferencia, en donde inmediatamente después de la administración de la TU al sujeto, se aplican al mismo pulsaciones eléctricas transcutáneas, que proporcionan una mayor eficiencia y reproducibilidad de la transferencia del ácido nucléico *in vivo* al tejido del sujeto (Mir et al., *Proc. Nat. Acad. Sci USA* 96: 4262-4267 (1999)).

En los métodos del presente invento que describen la inmunización de un sujeto administrando una vacuna del invento al mismo, la eficacia de la inmunización se puede monitorear de acuerdo con protocolos clínicos bien conocidos en la técnica para monitorear el estado inmune del sujeto.

Una cantidad efectiva de una composición de vacuna ya está determinada por quienes conocen la técnica como una cantidad que, al administrarse a un sujeto, confiere inmunidad protectora al mismo. Con el fin de llevar a cabo esa determinación, el técnico puede evaluar la capacidad para inducir anticuerpos específicos prM/M y E flavivirales y/o linfocitos T citotóxicos específicos prM/M- y E flavivirales presentes en la sangre de un sujeto a quien se le haya administrado la vacuna. También se puede determinar el nivel de la inmunidad protectora conferida a un sujeto experimental interrogando al flavivirus vivo correspondiente a la composición antigénica utiliza para inmunizar al sujeto experimental.

Estos experimentos cuestionantes son bien conocidos para quienes conocen la técnica.

En general, con el fin de inmunizar a un sujeto contra infección causada por WNV, JEV, YFV, virus dengue, SLEV, u otros flavivirus de acuerdo con el presente invento, y reconociendo que las TUs empleadas en esos métodos pueden tener tamaños generales divergentes, se pueden utilizar las dosis que fluctúan entre alrededor de 0.1 µg/kg por peso corporal hasta alrededor de 50 µg/kg por peso corporal.

Inesperadamente se han encontrado que una TU del presente invento que es una DNA confiere inmunidad protectora a un nivel de eficacia que se aproxima al 100% después de la administración de solamente una dosis efectiva simple de la TU mediante una inyección i.m. o mediante electrotransferencia. Esto es contraste con muchos métodos de inmunización llevados a cabo utilizando vacunas convencionales (tales como se describe arriba), que requieren una o más vacunas de refuerzo y que pueden no conferir inmunidad protectora a una eficacia cercana al 100%.

Adicionalmente se ha encontrado inesperadamente que la inmunidad protectora puede ser transmitida por un sujeto femenino vacunado a la descendencia de ese sujeto.

Se mostró que una porción significativa de ratones neonatales estaba protegida contra el reto viral después de que las madres se vacunaron utilizando la TU DNA del invento. Sin desear quedar limitados por la teoría, se sabe que la inmunidad pasiva se puede conferir a mamíferos neonatales debido a la presencia en la leche materna de anticuerpos neutralizantes específicos para los diversos patógenos. Es posible que la inmunidad protectora contra el JEV encontrado dentro de los neonatos se les haya transmitido de esta forma.

En otra incorporación del invento, la TU codifica una secuencia de señal de una proteína estructural de un primer flavivirus y un antígeno de flavivirus inmunogénico de un segundo flavivirus. Por consiguiente, en una incorporación, por ejemplo, la secuencia de la señal de la proteína estructural de un primer flavivirus se reemplaza mediante una secuencia de señal de proteína estructural de un segundo flavivirus, que origina un plegamiento adecuado del polipéptido incipiente, un procesamiento adecuado en un huésped y/o un plegamiento adecuado de la proteína procesada.

En otra incorporación del invento, la TU puede codificar un antígeno de flavivirus inmunógenico en donde el antígeno comprende una secuencia de uno o más de un flavivirus. La secuencia de señal puede ser un péptido de señal mejorada. El mejoramiento de las secuencias de

señal, o la selección de secuencias de señal más óptimas se puede lograr mediante la aplicación de los principios y las técnicas enseñadas en el Ejemplo 18 y en las referencias allí citadas, cada una de las cuales se incorpora a este documento por referencia para las enseñanzas expresas de cada una relacionadas con la selección, identificación y diseño de secuencias de señal con propiedades y funciones deseadas. Por lo general, estas propiedades y funciones deseadas incluirán una probabilidad de secuencia de señal alta.

En otra incorporación del invento, más de una TU o una TU que codifica un antígeno de flavirirus inmunogénico de más de un flavivirus se incluyen en una sola composición. Por consiguiente, en una incorporación, por ejemplo, una TU puede codificar un polipéptido polipéptidos incipientes que se procesan en proteínas de más de un flavivirus, y preferiblemente, de las proteínas procesadas de partículas subvirales que producen una respuesta inmunológica contra las proteínas. Las partículas subvirales se pueden formar a partir de proteínas procesadas derivadas de la secuencia del mismo flavivirus, o de una combinación de flavivirus, o de proteínas de flavivirus quimérico. Las vacunas de combinación que comprenden más de una TU o una TU que codifica un antígeno de flavivirus inmunogénico de más de un flavivirus se pueden diseñar para ser utilizadas en

regiones geográficas particulares mediante la inclusión de proteínas de flavivirus endémicos a la región o que se puedan encontrar similarmente de otra manera. Por ejemplo, una vacuna para Asia tropical y subtropical puede incluir TUs que codifiquen proteínas de los cuatro serotipos de vacunas de virus DEN, WN y JE. Las vacunas similarmente útiles para África y América Latina podrían incluir TUs que codifiquen proteínas de los cuatro serotipos de virus DEN, WN y YF y de los cuatro serotipos de virus DEN, Rocío y YF, respectivamente.

En otra incorporación, la TU codifica una secuencia de señal de una proteína estructural de un primer flavivirus y un antígeno de flavivirus quimérico inmunogénico que incluye una secuencia de aminoácido de más de un flavivirus. La secuencia de señal puede ser una secuencia de señal de virus de encefalitis japonesa. El antígeno de flavivirus quimérico puede incluir una secuencia de un antígeno de virus de encefalitis japonesa. En determinadas incorporaciones, el antígeno quimérico es una proteína E. La porción terminal carboxi de la proteína E puede ser una secuencia de proteína E del virus de encefalitis japonesa. La porción terminal carboxi puede ser, por ejemplo, el 5, 10, 15, 20, 25, 30, 40, 50 o 75% de la proteína E quimérica. En una incorporación preferida, la TU codifica una secuencia de señal de una proteína estructural de virus de encefalitis

japonesa, una proteína prM de un virus dengue y una proteína E quimérica que contiene una secuencia tanto del virus de encefalitis japonés como del virus dengue. La proteína quimérica puede ser una proteína E en donde la porción terminal carboxi comprende la secuencia del virus de encefalitis japonés. Los ejemplos de las TUs incluyen secuencias de ácido nucléico que se muestran en SEQ ID NO:44 y SEQ ID NO: 46 que pueden dirigirse la síntesis de los antígenos de flavivirus, tales como aquellas que se muestran en SEQ ID NO: 45 y SEQ ID NO: 47.

El presente invento proporciona además composiciones inmunogénicas que comprenden los polipéptidos de este invento en un portador farmacéuticamente aceptable para uso como vacuna de proteína. Los antígenos producidos a partir de las unidades transcripcionales del presente invento se pueden utilizar para producir respuestas inmunes eficaces en el sujeto. Los antígenos para ese propósito pueden comprender proteína prM de flavivirus, proteína M de flavivirus, proteína E de flavivirus o cualquier combinación de las mismas, incluyendo fragmentos inmunogénicos de las proteínas. Una incorporación particularmente preferida es la utilización del NRA descrito en este documento. Otra incorporación preferida es una proteína quimérica que comprende la secuencia de señal de un flavivirus y la(s) proteína(s) estructural(es) de uno o más flavivirus diferentes. En

000237

una incorporación preferida particularmente, la secuencia de señal del antígeno es la secuencia de la señal de encefalitis japonesa. En otras incorporaciones preferidas, la secuencia de señal es un péptido de señal mejorada. El mejoramiento de las secuencias de señal o la selección de secuencias de señal más óptimas se puede lograr mediante la aplicación de los principios y técnicas enseñadas en el Ejemplo 18 y en las referencias citadas en el mismo, cada una de las cuales se encuentra incorporada a este documento por referencia en cuanto respecta a las enseñanzas expresas en cada una relacionadas con selección, identificación y diseño de las secuencias de señal con propiedades y funciones deseadas. Por lo general estas propiedades y funciones deseadas incluirán una probabilidad de secuencia de señal alta.

En otras incorporaciones, la vacuna de proteína de este invento comprende además un adyuvante adecuado. Tal como se utiliza en este documento, un "adyuvante" es un potenciador o mejorador de la respuesta inmune. El término "adecuado" significa que incluye cualquier sustancia que pueda ser utilizada en combinación con el inmunógeno de la vacuna (v.g. la proteína prM del flavivirus, la proteína M del flavivirus, la proteína E del flavivirus o cualquier combinación de las mismas) para aumentar la respuesta inmune, sin producir

reacciones adversas en el sujeto vacunado. Las cantidades eficaces de un adyuvante específico se pueden determinar rápidamente de manera que se optimice el efecto de potenciación del adyuvante en la respuesta inmune de un sujeto vacunado. En una incorporación preferida, el adyuvante de las vacunas de este invento es un proceso de 2 etapas, que utiliza primero una solución de hidróxido de aluminio al 2% y luego un aceite mineral. En incorporaciones especificas, se pueden elegir adyuvantes adecuados a partir del siguiente grupo: emulsiones en agua con aceite mineral, vegetal o de pescado, adyuvante incompleto de Freund, *E. coli* J5, sulfato de dextran, óxido de hierro, alginato de sodio, Bacto-Adyuvante, algunos polímeros sintéticos, tales como Carbopol (BF Goodrich Company, Cleveland, Ohio), poli-aminoácidos y copolímeros de aminoácidos, saponina, carraginina, REGRESSIN (Vetrepharm, Athens, GA), AVRIDINE (N, N-dioctadecil-N',N'-bis(2-hidroxietil)-propanediamina), polímeros mamíferos ligados ß (1.4) polidispersados de cadena larga esparcidos con grupos O-acetilados (v.g. ACEMANNAN), extractos de pared celular purificada altamente desproteinizados derivados de la cepa no patogénica de especies Micobacterianas (v.g. EQUIMUNE, Vetrepharm Research Inc., Atenas GA), monooleato de Manita, aceite de parafina y dipéptido de muramilo.

En otro aspecto, este invento proporciona un método para inmunizar sujetos con cantidades inmunogénicas de la vacuna de proteína del invento para producir una respuesta inmune eficaz en el sujeto. La inmunización se puede lograr oralmente, parenteralmente, intranasalmente, intratraquealmente, intramuscularmente, intramamariamente, subcutáneamente, intravenosamente y/o intradérmicamente. La vacuna que contiene la proteína prM del flavivirus, la proteína M del flavivirus y/o la proteína E del flavivirus se puede administrar mediante inyección, inhalación, ingestión o infusión. Una sola dosis se puede dar y/o dosis repetidas de las preparaciones de la vacuna, v.g. "refuerzos" se pueden administrar a intervalos periódicos para mejorar la respuesta inmune inicial, o después de un largo período de tiempo desde la última dosis. El intervalo de tiempo entre las vacunas puede variar, dependiendo de la edad y del estado del sujeto.

El término "cantidad inmunogénica" significa una cantidad de un inmunógeno, o de una parte del mismo, que sea suficiente para inducir una respuesta inmune en un sujeto vacunado y que proteja al sujeto frente a la enfermedad causada por infecciones de flavivirus virulento o de tipo silvestre al estar expuesto a las mismas o que tenga un efecto benéfico terapéutico o

comercial que disminuya el efecto de la infección del flavivirus en el sujeto vacunado.

El invento proporciona además un anticuerpo producido en respuesta a una inmunización mediante el antígeno del invento. Los anticuerpos del presente invento pueden incluir anticuerpos policlonales y monoclonales que pueden ser moléculas de inmunoglobulina intacta, moléculas de inmunoglobulina quimérica, "anticuerpos humanizados" o fragmentos Fab o $F(ab')_2$. Estos anticuerpos y fragmentos de anticuerpos pueden ser producidos por técnicas bien conocidas que incluyen aquellas descritas en Harlow y Lane (Anticuerpos: Un Manual de Laboratorio. Laboratorio Cold Spring Harbor, Cold Spring Harbor, NY, 1989) y Kohler et al. (*Nature* 256:495-97, 1975) y las Patentes Norteamericanas 5.545.806, 5.569.825 y 5.625.126, incorporadas a este documento por referencia. Los anticuerpos pueden ser de cualquier isotipo IgG, IgA, IgD, IgE e IgM.

El presente invento también puede incluir anticuerpos de cadena sencilla (ScFv), que comprenden dominios ligados V_H y V_L y que retienen la conformación y actividad aglutinante específica del idiotipo nativo del anticuerpo. Estos anticuerpos de cadena simple son bien conocidos en la técnica y pueden ser producidos mediante métodos corrientes. (Ver como por ejemplo Álvarez et al., *Hum. Gene Ther.* 8: 229-242 (1997)).

Se pueden producir anticuerpos contra los antígenos de este invento que se sintetizan a partir de secuencias de ácido nucléico codificando secuencias de aminoácido inmunogénico de los antígenos prM, M y/o E de uno o más flavivirus y la secuencia de señal de un flavivirus diferente (como por ejemplo JEV). Los péptidos inmungénicos sintetizados a partir de la utilización de estas estructuras quiméricas se pueden identificar fácilmente con la utilización de métodos bien conocidos en la técnica para identificar regiones inmunogénicas en una secuencia de aminoácido y se pueden utilizar para producir los anticuerpos de este invento.

Son corrientes en la técnica las condiciones por medio de las cuales un complejo antígeno/anticuerpo se puede formar, así como los ensayos para la detección de la formación de un complejo antígeno/anticuerpo y la cuantificación de la proteína detectada. Esos ensayos pueden incluir secante Western, inmunoprecipitación, inmunofluorescencia, inmunocitoquímica, inmunohistoquímica, clasificación de células activadas por fluorescencia (FACS), hibridización por fluorescencia *in situ* (FISH), ensayos immunomagnéticos, ELISA, ELISPOT (Coligan et al., eds. 1995. Current Protocols in Immunology. Wiley, New York.), ensayos de aglutinación, ensayos de floculación, planificación de células, etc., que son bien conocidos para los expertos en la técnica.

Tal como se utiliza en este documento el término "aglutinar" significa la aglutinación bien caracterizada del anticuerpo al antígeno así como otra asociación no aleatoria con un antígeno. "Específicamente ligado" tal como se utiliza en este documento describe un anticuerpo u otro ligando que sustancialmente no tiene una reacción cruzada con cualquier antígeno distinto del especificado, que en este caso es un antígeno de este invento.

El anticuerpo o ligando del invento se puede ligar a un subestrato (como por ejemplo, cuentas, tubos, platinas, placas, hojas de nitrocelulosa, etc.) o conjugar con una mitad detectable, o ligar y conjugar. Las mitades detectables contempladas para el presente invento pueden incluir una mitad inmunofluorescente (v.g. fluoresceína, rodamina), una mitad radioactiva (v.g., ^{32}P, ^{125}I, ^{35}S), una mitad de enzima (v.g., peroxidasa de rábano rusticano, fosfatasa alcalina), una mitad de oro coloidal y una mitad de biotina. Estas técnicas de conjugación son corrientes en este arte (por ejemplo, Harlow y Lane, Anticuerpos: Un Manual de Laboratorio. Laboratorio Cold Spring Harbor, Cold Spring Harbor, NY (1989); Yang et al., *Nature* 382: 319-324 (1996)).

El presente invento proporciona además un método para detectar un anticuerpo de un flavivirus en una muestra, que comprende contactar la muestra con el

antígeno del flavivirus del presente invento bajo condiciones en las cuales se pueda formar un complejo de antígeno/anticuerpo; y detectar la formación del complejo, detectando así el anticuerpo del flavivirus en la muestra.

El presente invento proporciona además un método para detectar un antígeno de flavivirus en una muestra, que comprende contactar la muestra con un anticuerpo de este invento bajo condiciones en las que se pueda formar un antígeno/anticuerpo; y detectar la formación del complejo, para detectar así el antígeno del flavivirus en la muestra.

El método de detectar el antígeno del flavivirus en la muestra se puede realizar, por ejemplo, poniendo en contacto un fluido o tejido de muestra de un sujeto con un anticuerpo de este invento y detectar la aglutinación del anticuerpo al antígeno. Se contempla que el antígeno estará en un virión de flavivirus intacto, será una proteína codificada del flavivirus exhibida en la superficie de una célula infectada con flavivirus expresando el antígeno, o será un fragmento del antígeno. Una muestra del fluido de este método puede comprender cualquier fluido biológico que contendría el antígeno o una célula que contiene el antígeno, tal como fluido cerebroespinal, sangre, bilis, plasma, suero, saliva y

orina. Otros posibles ejemplos de fluidos corporales incluyen esputo, moco y similares.

El método de detectar el anticuerpo del flavivirus en una muestra se puede realizar, por ejemplo, poniendo en contacto un fluido o tejido muestra de un sujeto con una antígeno de este invento y detectar la aglutinación del antígeno al anticuerpo. Una muestra de fluido de este método puede comprender cualquier fluido biológico que pueda contener el anticuerpo, tal como un fluido cerebroespinal, sangre, bilis, plasma, suero, saliva y orina. Otros posibles ejemplos de fluidos corporales incluyen esputo, moco y similares.

Los inmunoensayos de enzimas, tales como ensayos de inmunofluorescencia (IFA), ensayos inmunosorbentes ligados a la enzima (ELISA) e inmunosecantes se pueden adaptar rápidamente para lograr la detección de los anticuerpos de flavivirus de acuerdo con los métodos de este invento. Un método ELISA eficaz para la detección de los anticuerpos puede por ejemplo, ser el siguiente: (1) ligar el antígeno a un substrato; (2) contactar el antígeno ligado con una muestra de fluido o tejido que contenga el anticuerpo; (3) contactar lo antedicho con un anticuerpo secundario ligado a una mitad detectable que sea reactiva con el anticuerpo ligado (v.g. enzima de peroxidasa de rábano rusticano o enzima de fosfatasa alcalina); (4) contactar lo antedicho con el substrato

para la enzima; (5) contactar lo antedicho con un reactivo a color; y (6) observar/medir el cambio o desarrollo del color.

Otra técnica inmunológica que puede ser útil para detectar los anticuerpos de flavivirus utiliza anticuerpos monoclonales (MAbs) para detectar los anticuerpos específicamente reactivos con los antígenos de flavivirus en un ensayo de inhibición competitiva. En resumen, la muestra se pone en contacto con una antígeno de este invento que se liga a un substrato (v.g. una placa ELISA 96-well). El exceso de muestra se lava totalmente. Una anticuerpo monoclonal rotulado (como por ejemplo ligado con enzima, fluorescente, radioactivo, etc.) se pone entonces en contacto con cualesquiera complejos de anticuerpo de antígeno formado previamente y se mide la cantidad de aglutinación del anticuerpo monoclonal. La cantidad de inhibición del aglutinamiento del anticuerpo monoclonal se mide en relación con un control (sin anticuerpo), permitiendo la detección y medición del anticuerpo en la muestra. El grado de inhibición del anticuerpo monoclonal puede ser un ensayo muy específico para detectar una variedad o cepa de flavivirus particular cuando se basa en la especificidad de aglutinación del anticuerpo monoclonal para una variedad o cepa en particular del flavivirus. También se pueden utilizar MAbs para la detección directa de

antígenos del flavivirus en las células mediante, por ejemplo un ensayo de inmunofluorescencia (IFA) de acuerdo con los métodos estándar.

Como un ejemplo adicional, se puede utilizar una prueba de micro-aglutinación para detectar la presencia de anticuerpos de flavivirus en una muestra, en resumen, las cuentas de látex, los glóbulos rojos u otras partículas aglutinables se revisten con el antígeno de este invento y se mezclan con una muestra de manera que los anticuerpos de la muestra que sean específicamente reactivos con el antígeno se entrelacen intermolecularmente con el mismo, causando aglutinación. Los complejos de antígeno-anticuerpo aglutinados forman un precipitado, visible a simple vista o medible mediante espectrofotómetro. En una modificación de la prueba antedicha, los anticuerpos de este invento se pueden ligar a las partículas del aglutinado y del antígeno en la muestra así detectada.

El presente invento proporciona además un método de diagnosticar una infección de flavivirus en un sujeto, que comprende contactar una muestra del sujeto con el antígeno de este invento bajo condiciones en las que se pueda formar un complejo antígeno/anticuerpo; y detectar la formación del complejo antígeno/anticuerpo, diagnosticando así una infección de flavivirus en un sujeto.

reaccionar con cada anticuerpo primario, haciendo que el anticuerpo primario sea más detectable.

La mitad detectable permite la detección visual de un precipitado o un cambio de color, la detección visual mediante microscopio, o la detección automatizada mediante espectrometría, la medición radiométrica o similares. Los ejemplos de mitades detectables incluyen fluoresceína y rodamina (para microscopio de fluorescencia), peroxidasa de rábano radicular (para microscopio de luz o de electrones y para detección bioquímica), biotina-estreptavidina (para microscopio de luz o de electrones) y fosfatasa alcalina (para detección bioquímica mediante cambio de color).

Las incorporaciones particulares del presente invento se estipulan en los ejemplos siguientes. Estos ejemplos no pretenden limitar el alcance del invento, tal como se revela en estas especificaciones.

Ejemplos

Los métodos generales que utilizan biología molecular y técnicas de DNA recombinantes relacionadas para preparar y expresar las moléculas TU de ácido nucléico del invento se establecen, por ejemplo en Current Protocols in Molecular Biology, Ausubel et al., John Wiley y Sons, New York 1987 (actualizado trimestralmente), y en Molecular Cloning: A Laboratory

000449
248

Manual 2ª Edición, Sambrook, Fritsch y Maniatis, Laboratorio de Cold Spring Harbor, Cold Spring Harbor, NY, 1989.

Ejemplo 1. Preparación de plasmidas recombinantes que contienen la unidad transcripcional que codifica los antígenos prM y E del JEV.- Se extrajo RNA genómico de 150 µL de semilla de virus SA 14 de cepa JEV cultivado de cerebro de ratón utilizando un Kit Viral RNA de la marca registrada QIAamp™ (Qiagen, Santa Clarita, CA). El RNA adsorbido en una membrana de sílice se levigó en 80 µL de agua libre de nucleasa y se utilizó como una plantilla para la amplificación de las secuencias que codifican el gen prM y E del JEV. Se obtuvieron las secuencias del imprimador a partir del trabajo de Nitayaphan et al. (*Virology* 177: 541-552 (1990)). Se amplificó un fragmento cDNA simple que contenía la región nucleótida genómica 389-2478 mediante la reacción inversa en cadena transcripteasa-polimerasa (RT-PCR). Los sitios de restricción KpnI y XbaI, la secuencia de consenso de aglutinación ribosomal Kozak, y el sitio de iniciación de la translación se diseñaron en el terminal 5' del cDNA mediante el amplímero 14DV389 (secuencia nuecleótida, SEQ ID NO:1; secuencia de aminoácido, SEQ ID NO:2). Un codon de terminación de translación en trama, seguido por un sitio de restricción NotI se introdujo en el terminal 3' del cDNA por el amplímero c14DV2453 (SEQ ID NO:3) (Figura

093243

El presente invento proporciona además un método de diagnosticar una infección de flavivirus en un sujeto, que comprende poner en contacto una muestra del sujeto con el anticuerpo de este invento bajo condiciones por medio de las cuales se pueda formar un complejo antígeno/anticuerpo; y detectar la formación del complejo antígeno/anticuerpo, diagnosticando así la infección del flavivirus en un sujeto.

En los métodos de diagnóstico que se enseñan en este documento, el antígeno de este invento puede estar ligado a un substrato y ponerse en contacto con una muestra de fluido tal como sangre, suero, orina o saliva. Esta muestra se puede tomar directamente del paciente o en una forma parcialmente purificada. De esta manera, los anticuerpos específicos para el antígeno (el anticuerpo primario) reaccionarán específicamente con el antígeno ligado. Posteriormente, un anticuerpo secundario ligado a una mitad detectable o rotulado con la misma se puede añadir para mejorar la detección del anticuerpo primario. Por lo general, el anticuerpo secundario o el otro ligando, que es reactivo, bien sea específicamente con un epítope diferente del antígeno o no específicamente con el ligando o el anticuerpo reactivado, se seleccionará por su capacidad para reaccionar con múltiples sitios en el anticuerpo primario. Por consiguiente, por ejemplo, varias moléculas del anticuerpo secundario pueden

2). Se realizó RT-PCR de un solo tubo utilizando un Kit Titan RT-PCR (Boehringer Mannheim, Indianapolis, IN). Se mezclaron 10 µL de RNA viral con un 1 µL de 14DV389 (50 µM) y de c14DV2453 (50 µM) con 18 µL de agua libre de nucleasa y la mezcla se dejó calentar a 85°C por 5 minutos y luego enfriar a 4°C. Se añadieron 75 µL de mezcla de la reacción [tampón 20 µL 5x, 2 µL de mezcla dNTP (10 mM cada uno), 5 µL de ditiotreitol (0.1 mM), 0.5 µL de Rnasin™ (40 U/µL, Boehringer Mannheim), 2 µL de mezcla de polimerasa, y 45.5 µL de agua libre de nucleasa] y el RT-PCR se desempeñó así: 1 ciclo (50°C por 30 minutos, 94°C por 3 minutos, 50°C por 30 s, 68°C por 2.5 min), 9 ciclos (94°C por 30 s, 50°C por 30 s, 68°C por 2.5 min), 20 ciclo (94°C por 30 s, 50°C por 30 s, 68°C por 2.5 min en el primer ciclo, con un incremento de 5 s por ciclo de allí en adelante), y una extensión final a 68°C por 15 min. El producto RT-PCR se purificó mediante un Kit de Purificación PCR QIAquick™ (Qiagen) y se levigó con 50 µL de 1 mM Tris-HCl, pH 7.5.

Todas las construcciones y análisis de vectores se llevaran a cabo utilizando técnicas estándar (Sambrook et al., 1989). El cDNA amplificado con RT-PCR, digerido con nucleasas KpnI y NotI se insertó en el sitio KpnI-NotI del vector plasmida de la expresión eucariótica (pCDNA3, Invitrogen, Carlsbad, CA). Las células Azules *Escherichia coli* XL1 de electroporación competente (Stratagene, La

000251

Jolla, CA) se transformaron mediante electroporación (Gene Pulser™, Bio-Rad, Hercules, CA) y se colocaron en placas de agar LB que contenían 100 µg/mL de carbenicilina (Sigma Chemical Co., St. Louis, MO). Se recogieron las clonaciones y se inocularon en 3 mL de caldo LB que contenía 100 µg/mL de carbenicilina. El DNA de plasmida se extrajo de un cultivo de 14 horas utilizando un Kit Spin Miniprep QIAprep™ (Qiagen). Se realizó una secuencia de DNA automatizada, tal como se recomienda (Biosistemas Aplicados/Perkin Elmer, Foster City, CA). Se secuenciaron ambas cepas del cDNA y demostraron ser idénticas a la secuencia de la cepa SA14 original (Nitayaphan et al., 1990).

El fragmento de plasmida pCDNA3 (Invitrogen, Carlsbad, CA) del nucleótido (nt) 1289 a nt 3455, que contiene f1 ori, SV40 ori, el gen de resistencia a la neomicina, y los elementos SV40 poli(A) se suprimieron mediante digestión PvuII y luego se ligaron para generar la plasmida pCBamp. El vector pCIBamp, que contiene una inserción intron quimérica en el sitio NcoI/KpnI del pCBamp se construyó escindiendo la secuencia intron del pCI (Promega, Madison, WI) mediante digestión con NcoI y KpnI. El fragmento resultante 566-bp se clonó en pCBamp digiriendo con NcoI-KpnI para reemplazar su fragmento 289-bp. La Figura 3 presenta la relación entre las plasmidas pCDA3, pCBamp, y pCIBamp.

Las plasmidas que contienen la unidad transcripcional que codifica las proteínas prM y E del JEV se prepararon a partir de estas plasmidas. El fragmento cDNA que contiene las regiones codificadas prM y E del JEV en la plasmida recombinante pCDJE2-7 (secuencia nucleótida, SEQ ID NO:10; la secuencia del aminoácido, SEQ ID NO:11), derivada del vector pCDNA3 se escindió mediante digestión con NotI y KpnI o XbaI y se clonó en el sitio KpnI-NotI del pCBamp, pCIBamp, pCEP4 (Invitrogen, Carlsbad, CA), o pREP4 (Invitrogen, Carlsbad, CA), o en el sitio SpeI-NotI del pRc/RSV (Invitrogen, Carlsbad, CA) del vector de expresión para crear pCBJE1-14 (secuencia nucleótida, SEQ ID NO:17; secuencia del aminoácido, SEQ ID NO:18), pCIBJES14, pCEJE, pREFE, y pRCJE, respectivamente. Ambas cepas del cDNA de los clones de cada plasmida se secuenciaron y se identificaron los clones recombinantes con la secuencia nucleótida correcta. Se purificó la DNA de plasmida para utilizarla en la transformación in vitro de las células mamíferas o en los experimentos de inmunización de ratones mediante cromatografía de intercambio de aniones utilizando un Kit Plasmid Maxi EndoFree™ (Qiagen).

Ejemplo 2. Evaluación de las proteínas prM y E del JEV expresadas mediante diversos plasmidas recombinantes utilizando un ensayo de anticuerpos inmunofluorescente indirecto.- La expresión de los productos de gen

específico JEV por parte de los diversos plasmidas de expresión recombinante se evaluó en líneas de células transfectadas transitoriamente de COS-1, COS-7 y SV-T2 (ATCC, Rockville MD; 1650-CRL, 1651-CRL, y 163.1-CCL, respectivamente) mediante ensayo de anticuerpo inmunofluorescente indirecto (IFA). La línea de la célula SV-T2 se excluyó de pruebas posteriores puesto que un resultado preliminar mostró que solamente 1-2% de las células SV-T2 transformadas fueron positivas al antígeno JEV. Para transformación, las células se cultivaron a una confluencia del 75% en matraces de cultivo de 150 cm^2 y se resuspendieron a 4°C en salina tamponada de fosfato (PBS) para un recuento celular final de 5 x 10^6 por mL. 10 µg de DNA plasmida se electroporaron en 300 µL de suspensión de células utilizando un BioRad Gene Pulse™ (Bio-Rad) graduado a 150 V, 960 µF y resistencia 100 Ω. Cinco minutos después de la electroporación, las células se diluyeron con 25 mL de medio fresco y se sembraron en un matraz de 75 cm^2. 48 horas después de la transformación se removió el medio de las células y éstas se tripsinizaron y resuspendieron en 5 mL de PBS con suero de cabra normal al 3%. Alícuotas de 10 µL se mancharon en placas, se secaron con aire y se dejaron con acetona a -20°C por 20 minutos. Se realizó el IFA con células transformadas de plasmidas fijadas con acetona utilizando inmunoglobina G de cabra anti-ratón conjugada

con isotiocianato de fluoresceína (Sigma Chemical Co.) y HIAF de JEV.

Para determinar la influencia de diversos elementos promotores y poli(A) en la expresión de proteína prM y E de JEV, se transformaron transitoriamente líneas de células COS-1 y COS-7 mediante una cantidad igual de pCDJE2-7 (SEQ ID NO:10), DNA de plasmida pCEJE, pREJE, o pRCJE. Los antígenos JEV se expresaron en ambas líneas de células transformadas por todos los cuatro plasmidas recombinantes, confirmando así que el promotor CMV o RSV (virus rous sarcoma) y los elementos BGH o SV40 poli(A) eran funcionalmente activos. Sin embargo, el porcentaje de células transformadas y el nivel de antígenos JEV expresados, según lo determinado por el número de células positivas IFA y por la intensidad IFA, respectivamente, difirió enormemente entre las diversas plasmidas (Tabla 1). Un porcentaje altamente significativo de células COS-1 transformadas por pCDJE2-7 (SEQ ID NO:10), pCBJE1-14 (SEQ ID NO:17) y pCIBJES14 expresaron los antígenos JEV y el nivel de las proteínas expresadas fue compatible con las células infectadas JEV. Las células transfectadas con vectores pCEJE, pREJE, o pRCJE por otra parte, tuvieron un bajo porcentaje de células que expresan antígeno, así como una baja intensidad de fluorescencia, indicando expresión débil de los antígenos.

Con el fin de determinar si la expresión mejorada de las proteínas JEV mediante pCDJE2-7 (SEQ ID NO:10) fue influenciada por el origen de replicación eucariótica codificada SV40, la plasmida pCBJE1-14 (SEQ ID NO:17) se interpretó de manera que un fragmento 2166-bp, que contenía f1 ori, SV40 ori, el gen de resistencia a la neomicina y los elementos poli(a) SV40 del pCDJE2-7, se eliminó. Se introdujeron un intron quimérico en el pCBJE1-14 para generar pCIBJES14. La plasmida pCIBJES14 se utilizó para determinar si la expresión de las proteínas JEV podía mejorarse mediante la secuencia intron. Después de la transformación, las células que albergan los vectores pCBJE1-14 y pCIBJES14 expresaron un nivel de antígenos JEV similar al observado con pCDJE2-7 (Tabla 1). Este resultado indica que la expresión de antígenos prM y E de JEV mediante los vectores recombinantes está influenciada solamente por los elementos regulatorios transcripcionales. Ni el origen eucariótico de la replicación ni la secuencia del intrón mejoraron la expresión del antígeno JEV en las células utilizadas. Se seleccionaron vectores que contienen el promotor CMV y el BGH poli(A) (Figura 3) para análisis posterior.

Ejemplo 3. Selección de una línea de célula estable transformada in vitro que constitutivamente expresa productos del gen específico JEV.- Las células COS-1 se

256

transformaron con 10 µg de pCDJE2-7 DNA mediante electroporación, tal como describió en el ejemplo precedente. Después de 24 horas de incubación en un medio de cultivo no selectivo las células se trataron con neomicina (0.5 mg/mL, Sigma Chemical Co.). Las colonias resistentes a la neomicina, que fueron visibles después de 2-3 semanas, se clonaron mediante dilución limitada en un medio que contenía neomicina. La expresión de los productos del gen JEV codificados con vector fue inicialmente examinada mediante IFA utilizando JEV HIAF. Un clon positivo JEV-IFA (JE-4B) y un clon negativo (JE-5A) se seleccionaron para análisis posterior y se mantuvieron en el medio que contenía 200 µg/mL de neomicina.

La autenticidad del la proteína E de JEV expresada por el clon JE-4B fue demostrada mediante diagramación epítope por IFA utilizando un panel de anticuerpos monoclonales de murina específica JEV E (Mab) (Kimura-Kuroda et al., *J. Virol.* 45: 124-132 (1983); Kimura-Kuroda et al., *J. Gen. Virol.* 67: 2663-2672 (1986); Zhang et al., *J. Med. Virol.* 29: 133-138 (1989); y Roehrig et al., *Virol.* 128: 118-126 (1983)). El JEV HIAF y el suero de ratón normal se utilizaron como controles de anticuerpo positivo y negativo, respectivamente. Cuatro JEV específicos, seis subgrupos de flavivirus específicos, y dos Mabs reactivos del grupo flavivirus

reaccionaron similarmente con el clon 4B o con las células JEV-infectadas COS-1 (Tabla 2).

Ejemplo 4. Propiedades antigénicas y detección inmunológica de las partículas subvirales secretadas por la línea de la célula JE-4B COS-1.

a. Preparación de las partículas subvirales. Las células JE-4B COS-1 crecieron y se mantuvieron en un medio que contenía 200 µg/mL de neomicina. El medio del cultivo se cosechó rutinariamente y almacenó a 4°C y se reabasteció dos veces a la semana, y las células se dividieron 1:5 cada 7-10 días. El medio de cultivo se aclaró mediante centrifugación a 10.000 rpm por 30 minutos en un rotor Sorvall F16/250 a 4°C y se centrífugo además por 4 horas a 39.000 rpm en un rotor Sorvall TH641 a 4°C a través de un cojín de sucrosa al 5% (p/p, preparado con 10 mM Tris HCl, pH 7.5, 100 mM NaCl (tampón TN)). El pelet que contiene las partículas subvirales se resuspendió en el tampón TN y se almacenó a 4°C. Alternativamente, se añadió PEG-8000 (p/v) al 7% o al 10% al medio de cultivo aclarado. La mezcla se agitó a 4°C durante por lo menos 2 horas y las partículas precipitadas se recopilaron mediante centrifugación a 10.000 rpm por 30 minutos. El precipitado se resuspendió en el tampón TN y se almacenó a 4°C. Las partículas subvirales se purificaron tanto en las preparaciones peletizadas como en las precipitadas-PEG mediante rata de centrifugación zonal

000253

en un gradiente de sucrosa continua al 5-25% en TN a 38.000 rpm a 4°C por 90 minutos. Se recopilaron 1-mL fracciones desde la parte superior del gradiente, se probaron mediante captación del antígeno ELISA (ver a continuación), y las fracciones positivas se cargaron en un gradiente de sucrosa al 25-50% en TN. Esto se centrífugo de un día para otro en una centrifugación de densidad de equilibrio a 35.000 rpm, a 4°C. Se recopilaron del fondo fracciones de 0.9-mL de los gradientes de equilibrio. Se probaron mediante la captura de antígeno ELISA y se evaluaron para la actividad de hemaglutinación (HA) a un pH 6.6. Se pesó una alícuota de 100 µL de cada fracción precisamente para determinar su densidad. Las fracciones ELISA-positivas se unieron y nodulizaron a 39.000 rpm a 4°C por 3-4 horas y el pelet se resuspendió en tampón TN. El ELISA de captura de antígenos y las titulaciones de HA se determinaron en las muestras nodulizadas. El licor flotante de la célula COS-1 infectadà con JEV también se sometió a protocolos de purificación similar, según se detalla arriba y se utilizó como un control positivo para el análisis de gradiente. Los viriones JE también se purificaron a partir de las células C6/36 infectadas 5-6 días postinfección mediante sedimentación en un gradiente de equilibrio glicerol/tartrato

b. Secantes Western (Western blots) de partículas subvirales.- Las muestras de gradiente purificado de las partículas subvirales se mezclaron con tampón de muestra de electroforesis y se corrieron sobre sulfato de dodecilo sódico al 10 o 1.25% que contenía gels de poliacrilamida (SDS-PAGE) según lo descrito por Laemmli (*Nature* 277: 680-685 (1970)). Las proteínas se transfirieron a una membrana de nitrocelulosa y se detectaron inmunoquímicamente con JEV HIAF policlonal, anti-E Mab 4G1 de reactivo cruzado de flavivirus (Henchal et al., *Amer. J. Trop. Med. Hyg.* 31: 830-836 (1982)), o con suero de ratón hiperinmune péptido anti-prM (JM01). La Figura 4 muestra una comparación de las proteínas M y E producidas por células C6/36 infectadas JEV JE-4B COS-1. Se observó alguna reactividad no específica a la proteína E en el fluido ascítico de ratón normal y suero anti-péptido Jmo1. Las proteínas idénticas en tamaño a la M y E fueron secretadas en las partículas subvirales y se pudieron detectar mediante Mab 4G2 E-específico y antisuero JM01 prM-específico, respectivamente.

c. Detección del gradiente de densidad de las partículas subvirales JEV en medio de cultivo.- Para el ELISA, el anticuerpo con captura de antígeno (4G2) se diluyó en 0.1 M de tampón de carbonato sódico, pH 9.6, y se utilizó para revestir placas de microtitulación 96-well (Immulon II, Dynatech. Chantilly, VA) mediante

incubación de un día para otro a 4°C. Después de bloquear con suero de cabra normal al 3% en PBS, se añadieron muestras diluidas en serie con doble propósito a la placa revestida 4G2 y se incubaron 1.5 horas a 37°C. El antígeno capturado se detectó mediante 6B6C-1 Mag conjugado con peroxidasa de rábano radicular, y se incubó por 1 hora a 37°C. La actividad de la enzima en la fase sólida se detectó luego con TMB (3,3',5,5'-tetrametilbencidina)-ELISA (Life Technologies, Grand Island, NY).

Aproximadamente 500 mL de medio de cultivo de célula de matraces de 15 X 150 cm^2 de células JE-4B se recopilaron cuatro días después de que se sembraron las células. Las partículas subvirales precipitadas PEG se resuspendieron en 2 mL de tampón TN, pH 7.5; una alícuota de 0.7 mL de este pelet resuspendido se cargó en un gradiente de sucrosa al 5-25%. Se añadió Tritón X-100, que interrumpe las partículas subvirales a otra alícuota de 0.7 mL a un concentración final de 0.1% y ésta se cargó a un gradiente de sucrosa al 5-25% preparado en tampón TN que contenía Tritón X-100 al 0.1%. Una banda opaca definida se observó aproximadamente a 2.5 cm de la parte superior del gradiente que contenía Tritón X-100, pero no en el gradiente sin detergente. Se recopilaron fracciones (1 mL) desde arriba hasta abajo para cada gradiente (Figura 5). Cada fracción recopilada se analizó mediante

el ELISA con captura de antígeno. El antígeno se detectó en fracciones 4-6, indicando una característica de sedimentación relativamente rápida de las partículas subvirales. El tratamiento del precipitado PEG a partir del medio de cultivo JE-4B con Tritón X-100 cambió la posición del material reactivo-ELISA a la parte superior del gradiente. Por consiguiente, el tratamiento con Tritón X-100 produce solo moléculas de sedimentación lenta. Se reportó un hallazgo similar por parte de Konishi et al. (*Virol.* 188: 714-720 (1992)). Estos resultados muestran que las partículas subvirales que se sedimentan rápidamente con contenido de prM/M y E pueden interrumpirse mediante tratamiento con detergente.

La actividad de la hemaglutinación (HA) se determinó en el rango de pH entre 6.1 y 7.0 con el método de Clarke y Casals (*Amer. J. Trop. Med. Hyg.* 7: 561-573 (1958)). La partícula subviral secretada por las células JE-4B y la partícula de virion producido por las células COS-1 infectadas con JEV tuvo un perfil HA similar con el pH óptimo determinado en 6.6.

Ejemplo 5. Comparación de la respuesta inmune en ratones vacunados con vacuna de ácido nucléico pCDJE2-7 del invento y la vacuna comercial JEV.- Grupos de cinco ratones hembras multiplicados con mezcla de razas ICR, de 3 semanas fueron inyectados intramuscularmente en los cuádriceps izquierdo y derecho con 100 µg de plasmida

pCDJE2-7 en 100 µL de dH_2O o se les administraron dosis de JE-VAX (fabricado por la Fundación de Investigación para Enfermedades Microbiológicas de la Universidad de Osaka y distribuido por Laboratorios Connaught, Swiftwater, PA.) subcutáneamente equivalente a una quinta parte de la dosis que se les da a los humanos. El plasmida pCDNA3/CAT (Invitrogen), que codifica y expresa una proteína no relacionada se utilizó como el control de vacunación negativa. Salvo por un grupo de ratones vacunados pCDJE2-7, todos los animales fueron reforzados 3 semanas más tarde con una dosis adicional de plasmida o de JE-VAX. Los ratones se sangraron del seno retroorbital a las 3, 6, 9, 23, 40 y 60 semanas después de la inoculación. Las titulaciones de anticuerpo JEV se determinaron mediante ensayo inmunosorbente ligado a la enzima (ELISA) frente al JEV purificado o mediante pruebas de neutralización de reducción de placa (PRNT) (Roehrig et al., *Virol.* 171: 49-60 (1989); y Hunt y Calisher, *Amer. J. Trop. Med. Hyg.* 28: 740-749 (1979)).

La vacuna de ácido nucléico pCDJE2-7 y el VAX-JE proporcionaron seroconversión del 100% dentro de las tres semanas siguientes a la primera vacunación en todos los tres grupos de ratones (Tabla 3). Las titulaciones de anticuerpo PRNT y JEV ELISA alcanzaron el nivel más alto en la semana 6 y en la semana 9, respectivamente, después de la inmunización. Los ratones que recibieron una dosis

de la vacuna DNA tuvieron respuestas de anticuerpos similares a los que recibieron 2 dosis. Se mantuvieron titulaciones de anticuerpo ELISA comparables en grupos vacunados con DNA hasta 60 semanas, después de lo cual se terminó el experimento. Sin embargo, solo uno de los cuatro ratones en el grupo JEV-VAX fue positivo en anticuerpo JEV y 60 semanas después de la inoculación. El grupo de control pCDNA3/CAT no tuvo ningún anticuerpo JEV medible. Estos resultados demuestran que una sola dosis de vacuna de ácido nucléico JEV-específica es más eficaz para mantener el anticuerpo JEV en ratones que la vacuna JE-VAX comercial aprobada por la FDA-

Ejemplo 6. Comparación de diversas estructuras de vacuna de ácido nucléico del invento con la vacuna comercial JEV para eficacia de la vacunación a diferentes edades.- Se expresó un nivel similar de proteína JEV por las células COS-1 transformadas bien sea por pCDJE2-7, pCBJE1-14, o pCIBJES14. La inducción del anticuerpo JEV mediante estas estructuras de ácido nucléico se comparó con la vacuna comercial JE-VAX en dos edades diferentes de la vacunación. Ratones sometidos a exogamia ICR de tres días de nacidos (de ambos sexos) o de 3 semanas (hembras), 10 por grupo, se vacunaron intramuscularmente con 50 o 100 µg de plasmida DNA, o subcutáneamente con dosis de JE-VAX que son una décima parte o una quinta parte de la dosis que se le da a los humanos. Se

recopilaron especimenes de suero 3 y 7 semanas después de la inmunización y se probaron a una dilución de 1:1600 mediante ELISA utilizando JEV purificado como antígeno. Los resultados se muestran en la Tabla 4.

La plasmida pCBJE1-14 proporcionó el grado más alto de seroconversión, v.g. una titulación de anticuerpos superior a 1:1600, logrando el 80-100% en ambas edades de la vacunación. La administración de pCDJE2-7 o pCIBJES14 proporcionó una seroconversión moderada a las 7 semanas cuando los ratones de 3 días de nacidos se vacunaron (60% para cada una), pero una seroconversión más débil (40% y 10%, respectivamente) cuando se midieron 3 semanas después de la vacunación. Sin embargo, cuando estos plasmidas se administran a la edad de 3 semanas, las seroconversiones de 90% o de 100% se lograron a las 3 y 7 semanas después de la vacunación. En contraste, la vacuna comercial, VAX-JE, no confirió seroconversión cuando se administró a los 3 días de edad, y de 100% cuando se administró a la edad de 3 semanas. Por consiguiente, las TUs del ácido nucléico para el prM y E del JEV proporcionaron un alcance de seroconversión mejor que una dosis muy alta de la vacuna comercial, e inesperadamente una seroconversión alta tanto en animales jóvenes como en los más maduros.

Ejemplo 7. Inmunidad protectora conferida por la vacuna de ácido nucléico del invento.- Los grupos

vacunados de tres días de nacidos del Ejemplo 6 se retaron 7 semanas después de la vacunación mediante inyección intraperitoneal de 50.000 pfu/100 µL de la cepa SA 14 del JEV adaptado a ratones y se observaron a las 3 semanas. Se logró 100% de protección en los grupos que recibieron diversas estructuras de vacunas que contenían TU de ácido nucléico hasta por 21 días (Tabla 5). En contraste, el 60% de los ratones vacunados con JE-VAX, así como el 70% de los controles negativos vacunados con pCDNA3/CAT no sobrevivieron al reto del virus a los 21 días. Estos resultados indican que las TUs del ácido nucléico del invento confieren protección eficaz inesperada en los ratones vacunados. Esto sugiere la posibilidad de emplear la vacuna de ácido nucléico del invento como una vacuna para humanos de temprana edad. En contraste, JE-VAX, la vacuna humana inactivada actualmente utilizada, no parece ser eficaz en animales jóvenes.

Ejemplo 8. Protección pasiva de ratones neonatales correlacionada con la titulación de anticuerpos maternos.- El ratón ICR hembra a la edad de 3 semanas se vacunó con cualquiera de las dosis o con dos dosis espaciadas por dos días de plasmida DNA pCDJE2-7, a 100 µg/100 µL, o con dos dosis de JEV-VAX que eran una quinta parte de la dosis que se le da a los humanos. El grupo de control negativo recibió dos dosis de 100 µg/100 µL de

000266

pCDNA-3/CAT. Se evaluó la protección pasiva mediante anticuerpo materno en crías resultantes de apareamientos de hembras experimentales con ratones machos no inmunizados que ocurrieron nueve semanas después de la primera vacunación o 6 semanas después de la segunda vacunación. Las crías se retaron entre 3-15 días después del nacimiento mediante administración intraperitoneal de 5.000 pfu/100 µL de virus SA14 adaptado a ratones y se observaron diariamente por 3 semanas (Tabla 6). Las tasas de supervivencia se correlacionaron con las titulaciones de anticuerpo neutralizante materno. El 100% de las crías alimentadas por las madres con un PRNT de 1:80 sobrevivieron la infección viral, mientras ninguna de las crías de la madre de control sobrevivió (Tabla 6). Se observó una protección parcial del 45% y del 75% en las crías mayores que fueron alimentadas por las madres con una titulación PRNT de 1:20 y de 1:40, respectivamente. Las tasas de supervivencia también se correlacionaron con la longitud del tiempo que las crías fueron alimentadas por la madre inmune. Tal como se acaba de indicar, las crías de 13-15 días de nacidas tuvieron altas tasas de supervivencia. Sin embargo, ninguna de las crías de 3-4 días de nacidas sobrevivió al reto del virus cuando la madre tuvo una titulación PRNT de 1:20 o 1:40. Por consiguiente, el anticuerpo materno proporciona inmunidad protectora parcial a completa a la descendencia. Además,

el anticuerpo JEV se detectó mediante ELISA en los sueros del 97% de las crías después del reto (29/30).

Los ratones fueron inoculados intramuscularmente con 1 o 2 dosis de 100 µg de plasmida DNA, o subcutáneamente con dos dosis humanas 1/5 de la vacuna JE-VAX. Se recopilaron los sueros nueve semanas después de la vacunación para probar el PRNT antes del apareamiento con una hembra no inmune.

Ejemplo 9. Preparación de las plasmidas recombinantes que contienen la unidad transcripcional que codifica los antígenos WNV prM y E.- Se extrajo RNA genómico de 150 µL de medio de cultivo de célula Vero infectada con cepa NY 99-6480, una cepa aislada de la irrupción en Nueva York en 1999, utilizando el Kit Viral RNA QIAamp™ (Qiagen, Santa Clarita, CA). El RNA extraído se levigó y suspendió en 80 µl de agua libre de nucleasa y se utilizó como una plantilla para la amplificación de las frecuencias de codificación del gen prM y E WNV. Las secuencias del imprimador se obtuvieron a partir del trabajo de Lanciotti et al. (*Science* 286: 2333-2337 (1999)). Se amplificó un fragmento de cDNA que contenía la región nucleótida genómica mediante la reacción en cadena inversa de transcriptasa-polimerasa (RT-PCR). Se diseñaron los sitios de restricción BsmBI y KasI en el terminal 5' del cDNA mediante la utilización del amplímero WN466 (secuencia nucleótida, SEQ ID NO:12). Se

introdujo un codon de terminación de traslación dentro de la trama, seguido por un sitio de restricción NotI en el terminal 3' del cDNA utilizando amplímero cWN2444 (SEQ ID NO:13). El producto RT-PCR se purificó mediante un Kit de Purificación PCR QIAquick™ (Qiagen).

Se digirió un amplicon de doble ramificación producido mediante la utilización de los dos amplímeros arriba mencionados (SEQ ID NO:12 y SEQ ID NO:13) con enzimas KasI y NotI para generar un fragmento de DNA de 998 bp (nt-1470 a 2468) se insertó en los sitios KasI y NotI de un vector pCBJESS para formar una plasmida intermedia, pCBINT. El pCBJESS se obtuvo de la plasmida pCBamp que contenía el promotor de gen precoz de citomegalovirus y el elemento de control translacional y un elemento de secuencia de señal JE diseñado (Chang et al., *J. Virol.* 74: 4244-4252 (2000)). El elemento de secuencia de señal JE comprende la secuencia de señal JE (SEQ ID NO:14).

El amplicon cDNA fue posteriormente digerido con enzimas BsmBI y Kas y el fragmento restante 1003 bp (nt-466 a 1470) se insertó en el sitio KasI del pCBINT para formar pCBWN (secuencia de ácido nucléico, SEQ ID NO:15; secuencia de aminoácido, SEQ ID NO:16). La secuencia automatizada del DNA utilizando un prisma ABI Consecuenciador 377 (Applied Biosystems/Perkin Elmer, Foster City, CA) se utilizó para confirmar que la

plasmida recombinante tenía una secuencia correcta de prM y E, tal como lo definió Lanciotti y otros (*Science* 286: 2333-2337 (1999)).

La plasmida DNA para utilizar en la transformación in vitro de las células mamíferas o en los experimentos de inmunización de ratones se purificó mediante cromatografía de intercambio de aniones, tal como se describe en el Ejemplo 1.

Ejemplo 10. Caracterización inmunoquímica y evaluación de las proteínas prM y E de WNV expresadas por pCBWN.- Los productos de genes específicos WNV codificados por la plasmida pCBWN se expresaron en células COS-1. Las células se electroporación y transformaron con plasmida pCBWN de acuerdo con Chang et al. (*J. Virol.* 74: 4244-4252 (2000)). Las células electroporadas se sembraron en matraces de cultivo de 75 cm^2 o en un plato de 12 cavidades para cultivo de tejido que contenía una cubierta/pozo (coverslip/well) estéril. Todos los matraces y las placas de 12 cavidades se mantuvieron en una incubadora a 37°C, con CO_2 al 5 %. Cuarenta horas después de la electroporación, se retiraron las células adherentes que contenían las cubiertas de las células, se lavaron brevemente con PBS, se fijaron con acetona por 2 minutos a temperatura ambiente, y se les dejó secar al aire.

000270

Se detectó expresión de proteína utilizando un ensayo de anticuerpos con inmunofluorescencia indirecta (IFA), tal como se describió en el Ejemplo 2. El anticuerpo monoclonal específico del flavivirus de la proteína E (Mab) 4G2, el fluido ascítico hiperinmune de ratón WNV (HIAF) y suero de ratón normal (NMS) a una dilución de 1:200 en PBS se utilizó como el anticuerpo primario para detectar la expresión de proteína (Henchal et al., *Am. J. Trop. Med. Hyg.* 31: 830-836 (1982)).

El medio de cultivo del tejido se cosechó 40 y 80 horas después de la electroporación. Se utilizó la captura del antígeno (captura-Ag) ELISA para detectar el antígeno del virus WN secretado en el medio de cultivo de las células COS-1 transitoriamente transformadas. El Mab 4G2 y el Mab 6B6C-1 de rábano radicular conjugado con peroxidasa se utilizaron para capturar los antígenos del virus WN y detectar el antígeno capturado, respectivamente. (Chang et al., *J. Virol.* 74: 4244-4452 (2000); Henchal et al., *Am. J. Trop. Med. Hyg.* 31: 830-836 (1983); Roehrig et al., *Virology* 128: 118-126 (1983)).

El antígeno del virus WN en el medio se concentró mediante precipatación con polietilenglicol al 10% (PEG)-8000. El precipitante se resuspendió en el tampón TNE (50 mM Tris, 100 mM NaCl, 10 mM EDTA, pH 7.5), se aclaró mediante centrifugación, y se almacenó a 4°C.

Alternativamente, el precipitante se resuspendió en un tampón de liofilización (0.1 M TRIZMA y albúmina de suero bovino al 0.4% en tampón salino borato, pH 9.0), se liofilizó y se almacenó a 4°C. Las preparaciones liofilizadas se utilizaron con antígeno para la evaluación en MAC- y en IgG ELISAs indirecto.

La proteína específica del virus WN se detectó mediante IFA en las células COS-1 transformadas transitoriamente. Las proteínas E, prM y M expresadas en estas células fueron secretadas en el medio de cultivo. El antígeno del virus WN concentrado por la precipitación PEG se extrajo con etanol al 7.0% para remover el PEG residual (Aizawa et al., *Appl. Enviro. Micro.* 39: 54-57 (1980)). Se analizaron los antígenos extraídos con etanol y los viriones WN de gradiente purificada en un NuPAGE, gradiente Bis-Tris Gel al 4-12% en un aparato de Electroforesis Excel Plus (Invitrogen Corp., Carlsbad,CA) y se siguió con electrosecante (electroblotting) hacia las membranas de nitrocelulosa utilizando una Unidad Blot Excel Plus (Invitrogen Corp.). Las proteínas específicas del virus WN producidas por las células COS-1 transformadas transitoriamente fueron detectadas por HIAF de ratón de virus específico WN o por el Mab 4G2 de flavivirus reactivo a la proteína E en un análisis de secante Occidental (Western blot), utilizando NMS como control de suero negativo. Las proteínas exhibieron una

reactividad similar e idénticos pesos moleculares al virion de la proteína E, prM y M del gradiente purificado correspondiente, derivado del virus WN infectado del cerebro del ratón cría (SMB).

En análisis de NRA como un antígeno para el diagnóstico ELISA, se reconstituyó una ampolleta de NRA liofilizado, representando el antígeno cosechado de 40 ml de fluido de cultivo de tejido en 1.0 ml de agua destilada y se comparó con el antígeno reconstituido del cerebro de ratón cría infectado con virus de WN (SMB) provisto como extractos de sucrosa-acetona liofilizados como β-propiolactona (Clarke et al., *Am. J. Trop. Med. Hyg.* 7: 561-573 (1958)). Todas las proteínas recombinantes, prM, M y E, tuvieron una reactividad similar a aquella del virion del gradiente purificado de las proteínas E, prM y M.

Los especimenes humanos codificados se probaron concurrentemente con antígenos en la misma prueba, en la etapa de desarrollo. Los protocolos de MAC- e IgG ELISA empleados fueron idénticos a los métodos publicados (Johnson et al., *J. Clin. Microbiol.* 38: 1827-1831 (2000); Martin et al., *J. Clin. Microbiol.* 38: 1823-1826. (2000)). Se obtuvieron especimenes de suero humano del banco de suero en nuestras instalaciones, que consisten de especimenes enviados al DVBID para prueba de confirmación del virus WN durante la irrupción de 1999.

En estas pruebas, se realizaron exámenes MAC- e IgG ELISA en una disolución de espécimen de 1:400. Los especímenes que produjeron relaciones positivas/negativas (P/N) OD entre 2 y 3 se consideraron positivos sospechosos. Los especímenes de suero sospechoso se sometieron a confirmación como positivos tanto mediante ELISA como mediante titulación de punto final y prueba de neutralización de reducción de placa (PRNT). Todos los especímenes que dieron relaciones P/N OD superiores a 3.0 se consideraron positivos sin necesidad de prueba confirmatoria adicional.

Un ELISA capturado-AG empleando reactivo del grupo flavivirus, anti-E Mab, 4G2 y 6B6C-1, se utilizó para detectar NRA secretado en el fluido del cultivo de células COS-1 pCBWN transformadas. El antígeno se pudo detectar en el medio un día después de la transformación; y la máxima titulación ELISA (1:32-1:64) en el fluido del cultivo sin concentración posterior se observó entre el día dos y el día cuatro. El NRA se concentró mediante precipitación PEG, se resuspendió en un tampón de liofilización y se liofilizó para preservación. Para el desarrollo de la prueba diagnóstica, se reconstituyó una ampolleta de NRA liofilizado con 1.0 ml de agua destilada y se titulo en el MAC- o IgG ELISA indirecto utilizando sueros humanos de referencia positiva y negativa de virus WN (Johnson et al., *J. Clin. Microbiol.* 38: 1827-1831

000274

(2000); Martin et al., *J. Clin. Microbiol.* 38: 1823-1826 (2000)). Se encontró que las diluciones 1:320 y 1:160 del NRA estaban en las concentraciones óptimas para utilización en MAC- e IgG ELISA, respectivamente. Estas diluciones generaron una relación P/N OD_{450} de 4.19 y 4.54, respectivamente para MAC- y la prueba IgG. Los antígenos SMB del virus WN producidos por las cepas NY-6480 y Eg101 se utilizaron a una dilución de 1:320 y 1:640 para MAC-ELISA, y de 1:120 y 1:320 para IgG ELISA, respectivamente. Los antígenos de control negativo, los precipitados PEG del medio de cultivo de células COS-1 normales y antígeno normal SMB, se utilizaron a las mismas diluciones que para el respectivo NRA y antígeno SMB. Los especímenes de suero humano, diluidos a 1:400 se probaron concurrentemente por triplicado con antígenos de control negativo y específicos del virus. Para que el resultado positivo de la prueba fuera válido, el OD_{450} del suero de la prueba que reaccionó con el antígeno viral (P) tenía que ser por lo menos dos veces mayor que el valor correspondiente de densidad óptica del mismo suero reaccionado con antígeno de control negativo (N).

La reactividad del NRA y de los SMBs del virus SLE y Eg101, NY-06480, se compararon por el MAC- y el IgG ELISAs utilizando 21 especimenes de suero humano codificados. De los 21 especimenes, 19 tuvieron resultados similares en todos los tres antígenos (8

negativos y 11 positivos sospechosos o positivos). Diez y ocho especímenes también se probaron separadamente utilizando antígeno SLE SMB. Solamente tres de los especímenes positivos 13 Eg-101-SMB fueron positivos en el SLE MAC-ELISA (Tabla 1). Ninguno de los especimenes negativos del antígeno WN fue positivo con el SLE MAC-ELISA. Este resultado confirmó una observación previa de que el virus anti-WN IgM no reaccionó cruzadamente de forma significativa con otros flavivirus (Tardei et al., *J. Clin. Microbiol.* 38: 2232-2239 (1940)) y fue específico para diagnosticar infección de virus agudo WN independientemente de si se utilizó el antígeno NRA o SMB en la prueba. Todos los especímenes fueron también probados concurrentemente mediante IgG ELISA. Diez de los 21 especimenes fueron positivos utilizando cualesquiera de los tres antígenos.

Los dos especímenes de suero discrepante (7 y 9) ambos del mismo paciente, recopilados en el día 4 y 44 después de que se inició la enfermedad, respectivamente, fueron negativos IgM con NRA y antígeno SMB NY y positivos IgM utilizando antígeno Eg-101 SMB en la prueba inicial. Para investigar aún más estos dos especímenes discordantes, se volvieron a probar seis especímenes recopilados secuencialmente de este paciente mediante el punto final MAC- y los IgG ELISAs. Se pudo demostrar un incremento serial superior a 32 veces

según la titulación MAC-ELISA entre el día 3 y el día 15, con todos los antígenos utilizados. El fluido cerebroespinal recopilado en el día 9 después de la iniciación de la enfermedad también confirmó que este paciente en efecto estaba infectado por WN poco tiempo antes de tomar la muestra. El fluido cerebroespinal tenía una lectura IgM P/N de 13.71 y 2.04 frente a los antígenos Eg-101- y SLE-SMB respectivamente. Los especimenes del día 31 y del día 44 fueron negativos (<1:400) utilizando antígeno NY-SMB, pero positivos utilizando NRA y Eg101-SMB. Se observaron titulaciones compatibles IgG con todos los tres antígenos utilizados en la prueba.

Ejemplo 11. Evaluación de la respuesta inmune en animales vacunados con pCBWN.- En el estudio se utilizaron grupo de diez ratones hembras ICR de tres semanas de nacidos. A los ratones se les inyectó intramuscularmente (i.m.) una sola dosis de pCBWN o una proteína fluorescente verde expresando plasmida (pEGFP) DNA (Clonetech, San Francisco, CA.). El DNA de plasmida pCBWN se purificó a partir de las células azules XL-1 con Kits de Giga Plasmida EndoFree (Qiagen) y se resuspendió en PBS, pH 7.5, a una concentración de 1.0 µg/µl. Los ratones que recibieron 100 µg de pEGFP se utilizaron como controles no vacunados. A los ratones se inyectó la plasmida pCBWN a una dosis de 100, 10, 1.0, o

0.1 µg en un volumen de 100 µl. Los grupos que recibieron 10, 1.0, o 0.1 µg de pCBWN fueron vacunados mediante el protocolo de entrega de gen *in vivo* mediado por electrotransferencia utilizando el electroporador de onda cuadrada EMC-830 (Genetronics Inc. San Diego, CA.). El protocolo de electrotransferencia se basó en el método de Mir et al., (*Proc. Natl. Acad. Sci. USA* 96: 4262-4267.(1999)). Inmediatamente después de la inyección de DNA, se aplicaron pulsaciones eléctricas transcutáneas mediante dos electrodos de placa de acero inoxidable, colocados a 4.5-5.5 mm de distancia, a cada lado de la pata. El contacto eléctrico con la piel de la pata se aseguró humedeciendo completamente la pata con PBS. Se aplicaron dos juegos de cuatro pulsos de 40 voltios/mm de 25 msec de duración con un intervalo de 200 msec entre pulsos. La polaridad del electrodo se reversó entre el conjunto de pulsaciones para mejorar la eficiencia de la electrotransferencia.

Los ratones se sangraron cada 3 semanas después de la inyección. La respuesta del anticuerpo específico del virus WN se evaluó mediante ELISA de captura-AG y prueba de neutralización de reducción de placa (PRNT). Se probaron los sueros individuales con IgG ELISA, y los sueros agrupados de 10 ratones de cada grupo se ensayaron con PRNT. Todos los ratones vacunados con pCBWN tenían titulaciones de IgG ELISA que fluctuaban entre 1:640 y

1:1280 tres semanas después de la vacunación. Los sueros agrupados recolectados a las semanas tres y seis tenían una titulación de anticuerpo Nt de 1:80. Ninguno de los especímenes de suero de los ratones de control pEGFP exhibieron ningún ELISA o titulación Nt al virus WN.

Para determinar si la vacunación i.m. simple de pCBWN podía proteger a los ratones contra la infección del virus WN, los ratones se retaron con virus NY-6480 bien sea mediante inyección intraperitoneal o mediante exposición a la picada de mosquitos infectados con el virus *Culex*. La mitad de los grupos de lo ratones fueron retados intraperitonealmente (ip) 6 semanas después de la vacunación con virus 1.000 LD_{50} (1,025 PFU/100 µl) de NY99-6480. Los ratones restantes se expusieron a picadas de tres mosquitos *tritaeniorhynchus* Culex que habían sido infectados con virus NY99-6480 7 días antes del experimento del reto. A los mosquitos se les permitió alimentarse en los ratones hasta que estuvieron totalmente llenos. Se observaron los ratones dos veces al día durante tres semanas después del reto.

Fue evidente que la presencia de los anticuerpos Nt se correlacionó con la inmunidad protectora puesto que todos los ratones inmunizados con DNA del virus WN permanecieron saludables después del reto del virus, mientras que todos los ratones de control desarrollaron síntomas de infección CNS 4-6 días después del reto del

virus y murieron en un promedio de 6.9 y 7.4 días después del reto intraperitoneal o del reto del mosquito infectado, respectivamente. En el grupo vacunado, el suero agrupado recopilado tres semanas después del reto con el virus (9 semanas después de la inmunización) tenía titulaciones de anticuerpo Nt de 1:640 o 1:320. Los sueros de los ratones vacunados agrupados reaccionaron solamente con la proteína E en el análisis de secante Western.

Se inmunizaron grupos de diez ratones con 10.0 a 0.1 µg de pCBWN por animal mediante la utilización de electrotransferencia. Todos los grupos que recibieron pCBWN estuvieron completamente protegidos del reto del virus. A las 6 semanas siguientes a la inmunización todos los grupos de ratones electrotransferidos tenía titulaciones Nt inferiores a cuatro veces diferentes a las de los animales que recibieron 100 µg de pCBWN mediante inyección i.m. convencional sin electrotransferencia. Los dos resultados que comprueban la inmunización eficaz sugieren que el protocolo de electrotransferencia mejora la inmunogenicidad y la eficacia protectora de la vacuna DNA del invento cuando se lleva a cabo en la forma descrita en (Mir et al., *Proc. Natl. Acad. Sci. USA.* 96: 4262-4267.(1999)).

Se mostró que las yeguas y los capones de raza mixta de varias edades utilizados en este estudio fueron negativos para el anticuerpo del virus SLE y del virus WN de acuerdo con el ELISA Y EL PRNT. A cuatro caballos se les inyectó intramuscularmente una sola dosis (1.000 µg/1.000 µl en PBS, pH 7.5) de plasmida de pCBWN. Los especimenes de suero se recopilaron día de por medio durante 38 días antes del reto con el virus y la respuesta del anticuerpo específico del virus WN se evaluó mediante MAC- o IgG ELISA y PRNT.

Dos días antes del reto con el virus, 12 caballos (4 vacunados y 8 de control) se relocalizaron en un nivel de bioseguridad (BSL)-3 de una edificación que los albergaba en la Universidad del Estado de Colorado. Los ocho caballos de control no vacunados fueron sujetos a un estudio que se diseñó para investigar la patogénesis inducidas del virus WN en caballos y el potencial de los caballos para servir como huéspedes amplificadores. Los caballos fueron retados cada uno con la picadura de 14 o 15 mosquitos *Aedes albopictus* que habían sido infectados con el virus NY99-6425 o BC787 12 días antes de hacerle el reto al caballo. A los mosquitos se les permitió alimentarse en los caballos por un período de 10 minutos. Los caballos se examinaron para buscar señales de la enfermedad dos veces al día. Se registró la temperatura corporal y se recopilaron especimenes de suero dos veces

al día desde el día 0 (día de la infección) hasta el día 10, y luego una vez al día hasta el día 14. Se registraron las pulsaciones y la respiración diariamente después del reto. Las muestras de suero recopilado se probaron mediante titulación de la placa para detectar la viremia y mediante el MAC- o el IgG ELISA y el PRNT para una respuesta de los anticuerpos.

No se observó ninguna reacción sistémica o local en ningún caballo vacunado. El suero de los caballos individuales se probó con PRNT. Los caballos vacunados desarrollaron un anticuerpo Nt superior o equivalente a 1:5 entre los días 14 y 31. Las titulaciones del punto final para los caballos vacunados, #5, #6, #7, y #8, en el día 37 (dos días antes del reto con el mosquito) fueron de 1:40, 1:5, 1:20, y 1:20, respectivamente. Los caballos vacunados con la plasmida pCBWN continuaron saludables después del reto con el virus. Ninguno de ellos desarrolló una viremia detectable ni fiebre desde los días 1 a 14. Todos los caballos de control no vacunados se infectaron con el virus WN después de exposición a las picaduras de mosquito infectado. Siete de los ocho caballos no vacunados desarrollaron viremia que apareció durante los primeros 6 días después del reto con el virus. Los caballos virémicos desarrollaron un anticuerpo Nt entre el día 7 y el día 9 después del reto con el virus. El único caballo de la totalidad del

000282

estudio que mostró signos clínicos de la enfermedad fue el caballo #11, que tuvo fiebre y mostró señales neurológicas a partir del día 8 después de la infección. Este caballo progresó a una enfermedad clínica severa dentro de las 24 horas y se sometió a eutanasia en el día 9. Cuatro caballos representativos, #9, 10, 14 y 15, que presentaron viremia para los días 0, 2, 4, o 6 se seleccionaron y utilizaron como ejemplos en este ejemplo. Las titulaciones del virus fluctuaron entre $10^{1.0}$ PFU/ml de suero (en el caballo #10), nivel detectable más bajo de nuestro ensayo, a $10^{2.4}$/ml (en el caballo #9). El caballo #14 no desarrolló una viremia detectable durante el período de la prueba. Sin embargo, este caballo se infectó con el virus, tal como lo comprueba el anticuerpo Nt detectado después del día 12.

La respuesta anamnéstica del anticuerpo Nt no se observó en los caballos vacunados tal como se comprueba por el incremento gradual en la titulación Nt durante el experimento. El anticuerpo Nt preexistente en el caballo vacunado antes del reto con el mosquito pudo suprimir la infección y replicación del virus inicial. Sin la replicación del virus, el antígeno del virus del reto suministrado por los mosquitos infectados puede no contener una masa de antígeno suficiente para estimular la respuesta inmune anamnéstica en el caballo vacunado. Todos los caballos vacunados se sometieron a eutanasia a

los 14 días después del reto con el virus. No se observaron lesiones patológicas e histopatológicas graves indicadoras de la infección viral WN.

Ejemplo 12. Preparación de los plamidas recombinantes que contienen secuencias de codificación para proteínas prM y E del virus de fiebre amarilla (YFV) o del virus de encefalitis St. Louis (SLEV).- Una estrategia similar a la de construir el plasmida recombinante pCDJE2-7 se utilizó para preparar plasmidas recombinantes YFV y SLEV. El RNA genómico se extrajo de 150 µL de la cepa TRI-788379 del YFV o semillas de virus 78V-6507 de la sepa SLEV utilizando un Kit RNA Viral Q1Aamp™ (Qiagen, Santa Clarita, CA.). El RNA viral se utilizó como una plantilla para amplificación de las regiones que codifican el gen prM y E del YFV o del SLEV. Se utilizaron imprimadores YFDV389 (secuencia nucleótida, SEQ ID NO:4; secuencia de aminoácido, SEQ ID NO:5), cYFDV2452 (SEQ ID NO:6), SLEDV410 (secuencia nucleótida, SEQ ID NO:7; secuencia aminoácida, SEQ ID NO: 8) y cSLEDV2449 (SEQ ID NO:9) para generar los ácidos nucléicos recombinantes correspondientes, tal como arriba se describe, para la preparación de las plasmidas recombinantes de JEV y de WNV. El cDNA amplificado RT-PCR digerido con las enzimas KpnI y NotI se insertó en un sitio de KpnI-NotI de un vector de plasmida de expresión eucariótica, pCDNA3 (Invitrogen). Se secuenciaron ambas

cepas del cDNA y se verificaron para la identidad con respecto a las secuencias de la cepa TRI-788379 YFV o de la cepa 78V-6507 SLEV. Los plasmidas recombinantes pCDYF2 y pCDSLE4-3, que contenían las secuencias nucleótidas de las regiones codificadoras prM y E para el YFV o SLEV, respectivamente, su purificaron utilizando un Kit de Maxi Plasmida EndoFree™ (Qiagen), y se utilizó para transformación *in vitro* de la inmunización de ratones.

Los antígenos específicos YFV o SLEV se expresaron en células COS-1 transformadas mediante pCDYF2 o pCDSLE4-3, respectivamente. El nivel de proteínas expresadas fue similar al control de la célula COS-1 infectada con YFV- o SLEV. Tal como en el modelo JEV se obtuvieron líneas de células COS-1 transformadas mediante genes que llevaban vectores para los antígenos virales que esencialmente expresan proteínas antigénicas YFV o SLEV. La diagramación epítope mediante IFA utilizando un panel de Mabs específicos-E de YFV o SLEV indicaron que la proteína E auténtica estaba expresada por las células COS-1 transformadas de pCDYF2- o pCDSLE4-3. Un estudio preliminar indicó que el 100% de los ratones ICR hembras de tres semanas, se seroconvirtieron después de inoculación intramuscular con una sola dosis 100 µg/100 µL de plasmida pCDSLE4-3 en agua desionizada.

Ejemplo 13. Preparación de plasmidas recombinantes que contienen secuencias de codificación para el virus de

encefalitis St. Louis con antígenos prM y E de la secuencia de señal JEV. El RNA genómico se extrajo de 150 µL de medio cultivo de célula Vero infectado con cepa MSI-7 del virus de encefalitis St. Louis utilizando el conjunto Kia QIAamp™ de RNA Viral (Qiagen, Santa Clarita, CA). El RNA extraído se levigó y suspendió en 80 µl de agua libre de nucleasa y se utilizó como una plantilla para la amplificación de las secuencias de codificación del gen prM y E del virus de encefalitis St. Louis. Las secuencias del imprimador se obtuvieron de un trabajo de Trent et al. (*Virology* 156: 293-304 (1987)). Un fragmento de cDNA que contenía la región nueclótida genómica se amplificó mediante reacción en cadena inversa de transcriptasa-polimerasa (RT-PCR). Se diseñó el sitio de restricción AfeI al terminal de 5' para el cDNA utilizando amplimer SLE463 (SEQ ID NO:30). En un codon de terminación de traslación en trama, seguido por un sitio de restricción NotI se introdujo en el terminal 3' del cDNA utilizando amplimer cSLE2447 (SEQ ID NO:31), el producto de RT-PCR se purificó mediante el Kit de purificación PCR cuya QIAquick™ (Qiagen).

El amplicon de doble ramificación, producido por la utilización de los dos amplímeros antedichos (SEQ ID NO:30 y SEQ ID NO:31), fue digerido con enzimas AfeI y NotI para generar un fragmento 2004 de DNA (463 a 2466nt), y se insertó en los sitios AfeI y NotI de un

vector pCBJESS-M para formar pCBSLE (secuencia nucletótida, SEQ ID NO:21; secuencia de aminoácido, SEQ ID NO:22). El pCBJESS-M se derivó de la plasmida pCBamp que contenía el promotor del gen precoz de citomegalovirus y el elemento de control translacional y un elemento de secuencia de señal JE modificada con ingeniería (SEQ ID NO:27). El elemento de la secuencia de señal JE comprende la secuencia de señal JE modificada en la posición -4 (Cis a Gli) y -2 (Gli a Ser) en la plasmida original pCBJESS.

La secuencia de DNA automatizada utilizando un Secuenciador 377 prisma ABI (Applied Biosystems/Perkin Elmer, Foster City, CA) se utilizó para confirmar que la plasmida recombinante tenía una secuencia prM y E correcta, tal como lo definió Trent et al. (*Virology* 156: 293-304 (1987)).

Ejemplo 14. Preparación de plasmidas recombinantes que contiene secuencias de codificación de las proteínas prM y E para el virus de fiebre amarilla (YFV) con secuencia de señal JEV.- El RNA genómico se extrajo de 150 µL de medio cultivo de célula Vero infectado con cepa 17D-213 de virus de fiebre amarilla utilizando el Kit RNA Viral QIAamp™ (Qiagen, Santa Clarita, CA). El RNA extraído se levigó y suspendió en 80 µl de agua libre de nucleasa y se utilizó como plantilla para la amplificación de las secuencias de codificación del gen prM y E del virus de

fiebre amarilla. Las secuencias del imprimador se obtuvieron a partir del trabajo de dos Santos y otros (*Virus Research* 35: 35-41 (1995)). Un fragmento de cDNA que contenía la región nucleótida genómica se amplificó mediante la reacción en cadena transcriptasa-polimerasa inversa (RT-PCR). Se diseñó con ingeniería el sitio de restricción AfeI en el terminal 5' del cDNA utilizando amplificador YF482 (SEQ ID NO:28). Un codon de terminación de traslación en trama, seguido por un sitio de restricción NotI se introdujo en el terminal 3' del cDNA utilizando amplímero cYF2433 (SEQ ID NO:29). El producto RT-PCR se purificó con el Kit de Purificación PCR QIAquick™ PCR (Qiagen).

El amplicón de doble ramificación, producido mediante la utilización de los dos amplímeros antedichos (SEQ ID NO:28 y SEQ ID NO:29), se digirió con enzimas AfeI y NotI para generar un fragmento 1971 de DNA (482 a 2452nt), y se insertó en los sitios AfeI y NotI del vector pCBJESS-M para formar pCBYF (secuencia nucleótida, SEQ ID NO:23; secuencia de aminoácido, SEQ ID NO:24). El pCBJESS-M se obtuvo de la plasmida pCBamp que contenía el promotor del gen precoz de citomegalovirus y el elemento de control translacional y un elemento de secuencia de señal JE diseñado con ingeniería (SEQ ID NO:27). El elemento de secuencia de señal JE comprende la secuencia

de señal JE modificada en las posiciones -4 (Cis a Gli) y -2 (Gli a Ser) de la plasmida JESS en el pCBJESS.

La secuencia de DNA automatizada utilizando un Secuenciador 377 prisma ABI (Applied Biosystems/Perkin Elmer, Foster City, CA) se utilizó para confirmar que la plasmida recombinante tenía una secuencia prM y E correcta, tal como lo definió dos Santos y otros (*Virus Research* 35: 35-41 (1995)).

Ejemplo 15. Preparación de plasmidas recombinantes que contiene secuencias de codificación para los antígenos prM y E del virus Powassan con secuencia de señal JEV.- El RNA genómico se extrajo de 150 µL de medio de cultivo de célula Vero infectado con cepa LB del virus Powassan utilizando el Kit RNA Viral QIAamp™ (Qiagen, Santa Clarita, CA). El RNA extraído se levigó y suspendió en 80 µl de agua libre de nucleasa y se utilizó como plantilla para la amplificación de las secuencias de codificación del gen prM y E del virus Powassan. Las secuencias de imprimación se obtuvieron a partir del trabajo de Mandl et al. (*Virology* 194: 173-184 (1993)). Un fragmento de cDNA que contenía la región nucleótida genómica se amplificó mediante reacción en cadena transcriptasa-polimerasa inversa (RT-PCR). El sitio de restricción AfeI se diseñó con ingeniería en el terminal 5' del cDNA utilizando amplímero POW454 (SEQ ID NO:25). Un codon de terminación de traslación en trama,

seguido por un sitio de restricción NotI se introdujo en el terminal 3' del cDNA utilizando el amplímero cPOW2417 (SEQ ID NO:26). El producto RT-PCR se purificó con un Kit de Purificación PCR QIAquick™ (Qiagen).

El amplicón de doble ramificación, producido por la utilización de los dos amplímeros antedichos (SEQ ID NO:25 y SEQ ID NO:26), se digirió con enzimas AfeI y NotI para generar un fragmento 1983 bp de DNA (454 a 2436nt), y se insertó en los sitios AfeI y NotI de un vector pCBJESS-M para formar pCBPOW (secuencia nucleótida, SEQ ID NO:19; secuencia de aminoácido, SEQ ID NO:20). El pCBJESS-M se obtuvo de la plasmida pCBamp que contenía el promotor del gen precoz del citomegalovirus y el elemento de control translacional y un elemento de secuencia de señal JE diseñado con ingeniería (SEQ ID NO:27). El elemento de secuencia de señal JE comprende la secuencia de señal JE modificada en la posición -4 (Cis a Gli) y -2 (Gli a Ser) en la plasmida pCBJESS.

La secuencia DNA automatizada utilizando un Secuenciador 377 prisma ABI (Applied Biosystems/Perkin Elmer, Foster City, CA) se utilizó para confirmar que la plasmida recombinante tenía una secuencia prM y E correcta, tal como lo definió Mandl et al. (*Virology* 194:173-184, (1993)).

Ejemplo 16. Preparación de plasmidas que contiene secuencias de codificación para proteínas estructurales de dengue serotipo 2.- Los procedimientos como los llevados a cabo para otros flaviririus (ver ejemplos 1, 9 y 12-15) deben seguirse para preparar vectores que incluyen TUs de ácido nucléico para antígenos de dengue serotipo 2. De acuerdo con los ejemplos, los amplímeros utilizados para la construcción de los vectores se pueden elegir para diseñar con ingeniería la secuencia de la señal de virus de dengue normal o se pueden elegir para diseñar con ingeniería una secuencia de señal a partir de otro flavivirus, tal como la secuencia de señal del virus de encefalitis japonesa modificada.

Se debe construir una plasmida que contiene la región del gen de dengue serotipo 2 desde el prM hasta E. Los genes prM y E del dengue serotipo 2 (Deubel et al., *Virology* 155:365-377 (1986); Gruenberg et al., *J. Gen. Virol.* 69: 1301-1398 (1988); Hahn et al., *Virology* 162:167-180 (1988)) se deben ligar en una plasmida tal como pCDNA3, y luego escindir y clonar en vectores tales como pCBamp, pCEP4, pREP4, o pRc/RSV (suministrados por Invitrogen, Carlsbad, CA) para permitir la expresión. Si fuere necesario, se puede amplificar una secuencia específica del virus dengue serotipo 2 codificada en una secuencia cDNA utilizando un procedimiento tal como la reacción de polimerasa en cadena (PCR).

Alternativamente, si el RNA viral es la fuente de la región del gen, una secuencia DNA puede ser amplificada mediante un procedimiento RT-PCR. Un fragmento DNA incluyendo un codon de iniciación en el extremo 5', y un codon de terminación en el extremo 3' debe clonarse en un vector de expresión en un sitio apropiado de restricción específica de nucleasa, de tal forma que el promotor precoz inmediato (IE) del citomegalovirus (CMV), un codon de iniciación, y un terminador, estén operablemente ligados a la secuencia del virus dengue serotipo 2.

Ejemplo 17. Vacunación de ratones utilizando una vacuna DNA de dengue serotipo 2.- La vacuna TU nucléica del dengue serotipo 2 que codifica la región del gen a partir de prM a E preparada en el Ejemplo 16 se debe suspender en un portador farmacéutico idóneo, tal como agua para inyección o salina fisiológica tamponada e inyectar intramuscularmente a grupos de ratones destetados. Los grupos de control recibieron una preparación de plasmida comparable que no tenía los genes específicos del dengue serotipo 2. La generación de los anticuerpos específicos del dengue serotipo 2 y/o de las células citotóxicas del sistema inmune específico al dengue serotipo 2 se debe evaluar a intervalos fijos de allí en adelante, por ejemplo a intervalos semanales. Alrededor de dos a cuatro meses después de la administración de la vacuna TU de ácido nucléico, los

ratones se retan con virus del dengue serotipo 2. Los niveles de viremia se deben evaluar de allí en adelante a intervalos apropiados, tal como cada dos días. Se debe evaluar la protección pasiva por el anticuerpo materno, tal como se indica en el Ejemplo 8.

Ejemplo 18. Diseño y construcción de Péptidos de Señal Mejorada.- Los péptidos de señal pueden determinar la translocación y orientación de la proteína insertada, y por consiguiente, la topología de las proteínas prM y E. La característica más común de los péptidos de señal de eucariotos consiste de un tramo de 8 a 12 ácidos aminohidrofóbicos llamados la región-h (von Heijne, "Signal sequences. The limits of variation" *J. Mol. Biol.* 184: 99-105 (1985)). La región entre el Met iniciador y la región-h, que se conoce como la región-n, por lo general tiene de uno a cinco aminoácidos, y normalmente porta aminoácidos cargados positivamente. Entre la región-h y el sitio de separación se encuentra la región-c, que consiste de tres a siete residuos de aminoácidos polares pero en su mayoría no cargados. Durante la síntesis de la poliproteína viral, la modulación del sitio de separación de signalasa desde una conformación críptica hasta la divisible en la unión de las proteínas C y prM depende de la remoción previa de la proteína C por el complejo de proteasa viral, NS2B/NS3 (Lobigs, "Flavivirus premembrane protein cleavage and spike

heterodimer secretion require the function of the viral protinase NS3", *Proc. Natl. Acad. Sci.* U S A, 90: 6218-6222 (1993)). Por consiguiente, es crítico considerar la eficacia de la secuencia de la señal viral cuando las proteínas prM y E deben expresarse solas mediante una plasmida de expresión.

Las diferencias péptidas de señal en diversas construcciones plasmidas pueden responder, por lo menos en parte, por las diferencias en la translocación de proteínas, la presentación del sitio de separación y la topología correcta, y por consiguiente de la formación VLP y secreción de prM y E. La modulación u optimización de estos atributos se puede mejorar mediante la selección o utilización de secuencias de señal con propiedades que impartan las características deseadas. Esto se puede lograr mediante la utilización de programas de computador de aprendizaje mecánico utilizando, por ejemplo, un modelo Markov oculto (HMM) capacitado en eucariotos (ver Henrik Nielsen et al., "Prediction of signal peptides and signal anchors by a hidden Markov model", En Actuaciones para la Sexta Conferencia Internacional sobre Sistemas Inteligentes para Biología Molecular (ISMB 6), Imprenta AAAI,, Menlo Park, California, pp- 122-130 (1998); Nielsen et al., "Machine Learning approaches to the prediction of signal peptides and other protein sorting signals", *Protein Engineering* 12: 3-9 (1999); Nielsen et

al., "A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites", *Int. J. Neural Sys.* 8: 581-599 (1997); "From sequense to sorting; Prediction of signal peptides", Henrik Nielsen, Ph.D. Tesis de Henrik Nielsen, Ph.D., defendida en el Departamento de Bioquímica dela Universidad de Estocolmo, Suecia (25 de mayo 25 1999); cada uno de los cuales se incorpora a este documento por referencia,. particularmente para la enseñanza relacionada con la optimización de las secuencias de señal utilizando algoritmos habilitados por computador).

Un ejemplo del tipo de programa utilizado se encuentra en la dirección http://www.cbs.dtu.dk/services/SignalP-2.0/, al 3 de abril de 2002. El HMM descrito en las referencias que se han mencionado y se han incorporado se aplicó para calcular la probabilidad péptida de señal de las secuencias péptidas de señal prM en estructuras de plasmidas diferentes (Tabla 7). La señal P-HMM busca correctamente predichos los sitios de división de la peptidasa de señal en todas las estructuras. Sin embargo, se observaron considerables diferencias en la probabilidad de separación (que fluctúan entre 0.164 y 1.000) y en la probabilidad péptida de señal (que fluctúan entre 0.165 y 1.00) (Tabla 7). Esto no es sorprendente, puesto que se sabe que el sitio de división y la probabilidad péptida de señal también están

influenciadas por los aminoácidos cargados positivamente en la región-n, la longitud del aminoácido hidrofóbico en la región-h y la composición del aminoácido en la región-c en las estructuras (Chang et al., "Flavivirus DNA vaccines: current statuts and potencial", *Annals of NY Acad. Sci.* 951: 272-285 (2001); Sakaguchi et at., "Functions of Signal and Signal-Anchor Sequences are Determined by the Balance Between the Hidrophobic Segment and the N-Terminal Charge", *Proc. Natl. Acad. Sci.* USA 89: 16-19 (1992)).

Tres estructuras plasmidas del virus JE, cada una derivada de diferente cepa del virus JE mostró potenciales de vacuna diferentes (Lin et al., "DNA immunization with Japanese encephalitis virus nonstructural protein NS1 elicits protective immunity in mice", *J. Virol.* 72: 191-200 (1998); Konishi et al. "Introductin of protective immunity against Japanese encephalitis in mice by immnunizaation with a plasmid encoding Japanese encephalitis virus premembrane and envelope genes". *J. Virol.* 72: 4925-4930 (1998); Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice", *J. Virol.* 74: 4244-4255 (2000); Las secuencias del péptido de señal en estas estructuras son diferentes en longitud de la región-n que pueden contener o no aminoácidos cargados (Tabla 7). La

región-n que contiene aminoácidos cargados positivamente forma un bucle corto en el lado citoplásmico que causa que la región-h (transmembrana helicoidal) se inserte en orientación de cola, exponiendo el sitio de división de la signalasa. En nuestro estudio, los VLPs secretados que contienen proteínas prM/M y E se pudieron purificar a partir de un medio de cultivo de las células COS-1 transformadas transitoriamente pCDJE1-14, o JE4B, o la línea de la célula transformada pCDJES-7. Los VLPs de gradiente purificada y los viriones tienen propiedades bioquímicas e inmunológicas idénticas. El procesamiento de la eficacia desde la proteína prM hasta la proteína M madura, el sello distintivo de la morfogénesis del flavivirus, es también similar entre las partículas de virion y VLPs. Por consiguiente, las proteínas prM y E expresadas por pCDJES-7 y pCDJE1-14 se pueden expresar como proteínas de transmembrana tipo I en la orientación similar a aquella del virion prM y E (Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice", *J. Virol.* 74: 4244-4252 (2000); en contraste, la proteína prM de pcDNA3JEME se puede expresar como una proteína de membrana tipo II con su región-h de transmembrana insertada en orientación de la cabeza debido a la ausencia de aminoácidos cargados positivamente en su región-n Konishi et al. "Induction of protective immunity against Japanese encephalistis in

mice by immunization with a plasmid encoding Japanese encephalitis virus premembrane and envelope genes". *J. Virol.* 72: 4925-4930 (1998)); La síntesis de proteína eficiente en combinación con la proteína expresada que tiene la topología correcta, particularmente del prM y E expresados, puede mejorar la formulación y secreción VLP y por consiguiente promover la inmunogenicidad de la vacuna DNA (Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice", *J. Virol.* 74: 4244-4252 (2000)).

La utilización de cómputos que se basan en computador, tal como arriba se describió, se ha aplicado para optimizar el diseño de la plasmida de expresión. En particular, la facultad predictiva del programa SignalP-HMM se ha aplicado para diseñar la plasmida de expresión del virus WN (Tabla 2) (Davis et al. "West Nile virus recombinant DNA vaccine protects mouse and horse from virus challenge and expresses in vitro a noninfectious recombinant antigen that can be used in enzyme-linked immunosorbent assays", *J. Virol.* 75: 4040-4047 (20001)). La plasmida pCBWN consiste de una versión corta del péptido de señal de virus JE seguido por la secuencia del gen prM-E del virus WN. El potencial de la vacuna de esta estructura se demostró ampliamente, puesto que una sola inyección i.m. de pCBWN DNA no solamente indujo una

inmunidad protectora sino que también impidió la infección del virus WN en ratones y caballos.

Tal como se discutió anteriormente, y como se demostró en los Ejemplos 13-15, la secuencia de señal que codifica el virus del mismo virus como las regiones que codifican el antígeno no es necesariamente el péptido de señal óptima disponible. Además, la secuencia de señal no modificada no es necesariamente óptima. Por ejemplo, el péptido de señal codificado en la plasmida pCBJE1-14 se puede mejorar, medida por la probabilidad de secuencia de señal acortando la región-n, cambiando la región-c, o mediante la combinación de ambas modificaciones (Figura 6). A manera de ilustración, se ha utilizado una versión acortada del péptido de señal del virus JE para las expresiones de los genes prM y E del virus WN, tal como se describe en este documento y en los documentos incorporados a este invento por referencia a las enseñanzas (Davis et al. "West Nile virus recombinant DNA vaccine protects mouse and horse from virus challenge and expresses in vitro a noninfectious recombinant antigen that can be used in enzyme-linked immunosorbent assays", *J. Virol.* 75: 4040-4047 (2001)). Los estudios de titulación de dosis mediante inoculación simple i.m. indicaron que el pCBWN fue por lo menos 2-4 veces más inmunogénico que el pCBJE1-14 en ratones.

Ejemplo 19. Vacunas Multivalentes.- También se pueden producir las vacunas multivalentes y/o de combinación diseñadas para inmunizar contra múltiples flavivirus. Al preparar una vacuna multivalente, los componentes de la vacuna monovalente se preparan incluyendo los elementos relacionados con los patógenos de interés, tales como YF, diferentes serotipos de virus DEN, JE, WN, SLE y TBE (RSSE y CEE) y cualquier otra combinación de flavivirus. El diseño y la producción de estructuras DNA, tal como se describe en los otros ejemplos y en las especificaciones, se llevan a cabo en la forma descrita. Se pueden hacer combinaciones de vacunas apropiadas para proporcionar vacunas de combinación o multivalentes que protejan contra patógenos múltiples. Los datos preliminares de nuestro grupo han demostrado que la inyección i.m. de vacunas combinadas pCBJE1-14 y pCBWN DNA indujeron anticuerpos Nt específicos de virus WN y de virus JE en ratones (Tabla 8). Cada componente monovalente, inclusive si está construido utilizando reguladores translacionales y transcripcionales idénticos preferiblemente se debe probar en un sistema de modelo análogo para garantizar su potencial de vacuna. Un cóctel de vacuna de vacunación se puede entonces formular. Estos cócteles de vacuna se pueden diseñar específicamente para regiones geográficas particulares. Por ejemplo, un cóctel de vacuna para el Asia tropical y subtropical debe incluir cuatro

serotipos de vacuna de virus DEN; WN y JE. Los cócteles de vacunas similarmente útiles para el África y América Latina deberían incluir cuatro serotipos de vacunas DEN, WN y virus YF y cuatro serotipos de vacunas de virus DEN, Roció y YF, respectivamente.

Ejemplo 20. Preparación y prueba de vacunas recombinantes del virus dengue tipo 2.

a. Resumen del Ejemplo.- Se construyeron una serie de plasmidas que codifican las proteínas de premembrana (prM) y las proteínas envolventes (E) del virus dengue tipo 2 (DEN-2). Estos plasmidas incluyeron una estructura del prM-E del DEN-2 auténtico (pCBD2-14-6)(SEQ ID NO:42) que codifica la proteína descrita por SEQ ID NO:43, una estructura quimérica E del virus de encefalitis japonesa (JE) 90% DEN-2 E-10% (pCB9D2-1J-4-3)(SEQ ID NO:44) que codifica la proteína descrita por SEQ ID NO:45 y una estructura quimérica E JE 80% DEN-2E-20% (pCB8D2-2J-9-1)(SEQ ID NO:46) que codifica la proteína descrita por (SEQ ID NO:47). La reactividad del anticuerpo monoclonal (Mab) indicó que todas las tres plasmidas expresaron epítopes de proteína E que reaccionaron con un panel de anticuerpos de dominio 1, 2 y 3. Sin embargo, solamente la estructura pCB8D2-2J-2-9-1 (SEQ ID NO:46) secretó altos niveles de prM, M (prM madura) y E en el medio de las células plasmidotransformadas COS-1. La porción mayor de la proteína prM y E expresada por las células

transformadas COS-1 con plasmida pCBD2-14-6 (SEQ ID NO:42) y por las células COS-1 transformadas con plasmida pCB9D2-4-3 (SEQ ID NO:44) permanecieron con las membranas ligadas. El reemplazo del 20% de la secuencia que codifica la proteína E del DEN-2 E con la secuencia que codifica la correspondiente secuencia de proteína JE E no tuvo ningún efecto en la reactividad MAbs .

En la prueba, grupos de ratones recibieron dos inmunizaciones intramusculares de plasmidas seleccionadas en las semanas 0 y 3, y la respuesta inmune fue evaluada determinando el anticuerpo ELISA y la neutralización especifica. Las prM y E secretadas que expresan plasmida, que puede formar particulas subvirales (SVPs) fue superior a las de otras estructuras en estimular una respuesta de anticuerpo. El 90% de las titulaciones de neutralización que fluctuaron entre 1:40 y >1:1.000 se observaron en 7 de los 9 especimenes séricos de ratones inmunizados pCB8D2-2J-2-9-1.

b. Importancia de las vacunas y del virus DEN-2. La fiebre dengue (DEN) es una infección aguda que ocurre en áreas tropicales y subtropicales. Esta es una de las enfermedades flavivirales más importantes de los humanos. Tal como se anotó anteriormente, existen cuatro serotipos claros DEN (DEN-1, DEN-2, DEN-3 y DEN-4) del virus dengue. La infección con cualquiera de estos es usualmente bien sea asintomática o solamente causa una

.000302

enfermedad febril autolimitada conocida como fiebre dengue (DF). Sin embargo, en un porcentaje pequeño de los casos, la infección del virus dengue origina una enfermedad mucho más seria, la fiebre hemorrágica de dengue que puede ser mortal o el síndrome de choque dengue (DHF/DSS). Por consiguiente, mientras existen aproximadamente 100 millones de casos del DF relativamente suave en todo el mundo anualmente, que presentan preocupación limitada, también se calculan 500.000 casos de DHF/DSS que requieren hospitalización anualmente. Para protegerse contra esta enfermedad, se requiere una vacuna segura y eficaz DEN contra todos los cuatro serotipos para ser administrada a niños y adultos no inmunes en las regiones endémicas y epidémicas de DEN.

Las vacunas seguras deben minimizar el riesgo potencial de infección seria por virus virulentos. Estos virus virulentos pueden surgir por reversión de genes o recombinación de algunos tipos de vacunas derivadas de virus de vacunas atenuados. Estas ocurrencias surgieron en la campaña de erradicación de polivirus (Guillot et al., "Natural Genetic Exchanges between Vaccine and Wild Poliovirus Strains in Humans", *J. Virol.* 74: 8434-8443 (2000): Liu et al. "Molecular Evolutin of a Type 1 Wild-Vaccine Poliovirus Recombinant during Widespread Circulation in China", *J. Virol.* 74: 11153-11151 (2000)). Además, la secuencia genómica de una cepa americana de

virus de fiebre amarilla, TRINID 79A, demuestra que existe una amplia similitud entre esta cepa y el virus de vacuna de fiebre amarilla atenuado, FNV (Chang et al. "Nucleotide sequense variation of the envelope protein gene identifies two distinct genotypes of yellow fever virus", *J. Virol. 69*: 5773-5780 (1995); Pisano et al. "Complete nucleotide sequense and phylogeny of an American strain of yellow fever virus. TRINID79A", *Arch. Virol.* 144: 1837-1843 (1999)). Aún cuando no es concluyente por si misma, la similaridad sugiere fuertemente que el TRINID 79A se deriva del virus de la vacuna FNV.

La utilización de vacunas que se basan en DNA constituye un enfoque novedoso y prometedor de inmunización para el desarrollo de las vacunas de flavivirus (tal como aquí se describe, en Chang et al., "Flavivirus DNA vaccine: current status and potential". *Ann NY Acad. Sci.* 951: 272-285 (2001) y en las referencias aquí citadas). En este ejemplo, se produjeron una serie de vacunas de DEN-2 y la respuesta inmune en ratones después de inmunización i.m. de estructuras DEN-2 se correlacionó con la eficiencia de la secreción de prM/M y E. Una estructura que condujo a la secreción de cantidades significativas de antígenos prM/M y E demostró ser capaz de estimular altas titulaciones de anticuerpos neutralizantes en la plasmida de ratones vacunados.

c. Materiales y métodos.

i. Cultivo de células y cepas de virus. Las células COS-1 (ATCC, Manassas, VA; 1650-CRL) se cultivaron a 37°C con CO_2 al 5% en un medio esencial mínimo Eagle modificado de Dulbecco (DMEM, GIBCO, Grand Island, NY) suplementado con suero bovino fetal inactivado calentado al 10% (FBS, Hyclone Laboratories, Inc. Logan, UT), 1mM de piruvato sódico, 1mM de aminoácidos no esenciales, 30 ml/litro de $NaHCO_3$ al 7.5%, 100 unidades/ml de penicilina, y 100 µg/ml de estreptomicina. Se cultivaron células Vero y C6/36 bajo las mismas condiciones utilizadas para las células COS-1. Se utilizó virus DEN-2, cepa-16681 para la clonación cDNA y IgG ELISA y la prueba de neutralización de reducción de placa (PRNT). El virus se propagó en el cultivo de la célula C6/36. El virus utilizado para los estudios inmunológicos o bioquímicos se purificó mediante precipitación con polietilenglicol al 7% (PEG-8000; Fisher Scientific, Fair Lawn, NJ) seguido de ultracentrifugación en gradientes de tartrato de potasio al45%- glicerol al 30% (Obijeksy et al. "Segmented genome and nucleocapsid of La Crosse virus", *J. Virol. 20:* 664-675 (1976)).

ii. Construcción plasmida. El RNA genómico se extrajo de 150 µg de medio de cultivo de célula C6/36 infectado con DEN-2 cepa 16681 utilizando el Kit RNA

000305

Viral QlAamp™ (Qiagen, Santa Clarita, CA.). El RNA extraído se resuspendió en 80 µl de agua tratada con dietilpirocarbonato (DEPC, Sigma, ST. Louis, MO) y luego se utilizó como una plantilla en amplificación de transcriptasa-PCR (RT-PCR) inversa de genes prM y E del virus DEN-2. Las secuencias del imprimador (Tabla 9) se diseñaron basándose en secuencias publicadas (Gadkari et al., "Critical evaluation of Kyasanur Forest disease virus neutralizing antibodies found in bats (a preliminary report)", *Indian J. Med. Res.* 64: 64-67 (1976); Kinney et al."Construction of infectious cDNA clones for dengue 2 virus; strain 16681 and its attenuated vaccine derivative, strain PDK-53". *Virology* 230: 300-308 (1997)). El reconocimiento y lugar de división para la restricción dela enzima KasI se incorporó al terminal 5' del amplicon cDNA. Un codon de terminación en trama, seguido por un sitio de restricción NotI se introdujo en el terminal 3' del amplicon cDNA. El amplicon cDNA del virus DEN-2 se digirió con enzimas KasI y NotI, y luego se insertó en los sitios KasI y NotI de un vector pCBJESS para formar la plasmida DEN-2-E al 100%, pCBD2-14-6 (SEQ ID NO:42).

Para construir las plasmidas DEN-2 E al 90% y 80%, el 100% de las plasmida DEN-2 pCBD2-14-6 (SEQ ID NO:42), y la plasmida JE, pCBJ1-14-6 (SEQ ID NO:17) se utilizaron como plantillas PCR para amplificar al secuencia DEN-2 y

JE DNA, respectivamente. Los conjuntos de imprimadores utilizados en las reacciones de amplificación para obtener los fragmentos de gen de DEN-2 y de JE se enumeran en la Tabla 9. Los sitios del imprimador T7 y SP6 se encuentran en la plasmida pCBamp derivada de la plasmida original pCDNA-3 (Invitrogen, Carlsbad, CA) y se pueden utilizar como se desee o requiera. Los fragmentos DNA amplificados PCR para el gen de proteína E del DEN-2 al 90% y del JE al 10% se digirieron con endonucleasa de restricción BxtKI, se ligaron utilizando ligaza T4 DNA, se digirieron con enzima KasI y NotI y se insertaron en los sitios KasI y NotI del vector pCBJESS para obtener la plasmida, pCB9D2-1J-4-3 (SEQ ID NO:44). Los fragmentos DNA amplificados PCR para el gen E del DEN-2 al 80% y del JE al 20% se digirieron con BsmBI, se ligaron con ligaza T4 DNA, se digirieron con enzima KasI y NotI y se insertaron en los sitios KasI y NotI del vector pCBJESS para obtener la plasmida, pCB8D2-2J-2-9-1 (SEQ ID NO:46). En la Figura 7 se muestran representaciones esquemáticas de las tres construcciones de plasmida. La E del DEN-2 al 90% y del JE al 10% y las regiones de unión de la proteína E del DEN-2 al 80% y del JE al 20%, respectivamente, se muestran en la Tabla 9.

La secuencia DNA automatizada se realizó en un Secuenciador 377 Prisma ABI (Applied Biosystems/Perkin Elmer, Foster City, CA) de acuerdo con los procedimientos

recomendados por el fabricante. Las plasmidas recombinantes con las secuencias correctas prM y E se identificaron utilizando análisis de secuencia.

iii. Expresiones transitorias de antígeno recombinante DEN-2 en células COS-1 por electroporación. Las células COS-1 se electropolaron separadamente con la plasmida de cada DEN-2 o con control de plasmida de expresión de proteína fluorescente verde (pEGFP, Clonetech, San Francisco CA) utilizando el protocolo descrito en otras partes de los ejemplos y en Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice", *J. Virol.* 74: 4244-4252 (2000)); Las células electroporadas se sembraron en matraces de cultivo de 75 cm^2 y se mantuvieron a 37°C y con CO_2 al 5%. Seis horas después de la electropolación se reemplazó el medio de crecimiento con un medio de mantenimiento que contenía suero bovino fetal al 2%. El medio del cultivo del tejido y las células se cosecharon separadamente 48 horas después de la electroporación para caracterización del antígeno.

iv. Diagramación epítope utilizando anticuerpos monoclonales específicos E del DEN-2. 48 horas después de la electroporación, se tripsinizaron las células adherentes, se resuspendieron en PBS con contenido de suero de cabra al 5%, se mancharon en una lámina de machar

de 12 cavidades y se secaron con aire. Las células adheridas a la placa de manchas fueron fijadas con acetona por 10 minutos a -20°C y se les permitió secarse al aire. Se utilizaron anticuerpos monoclonales específicos de proteína E (Mab) para detectar la expresión de proteínas mediante el ensayo de anticuerpos de inmunofluorescencia indirecta (IFA), tal como se describió previamente (Tabla 10; Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice", *J. Virol.* 74: 4244-4252 (2000));

v. Caracterización del antígeno del virus recombinante DEN-2. El medio de cultivo del tejido se cosechó 48 horas después de la electroporación. La captura del antígeno (Ag-Captura) ELISA se utilizó para detectar el antígeno del virus DEN-2 secretado en el medio de cultivo de células COS-1 transformadas transitoriamente. Se utilizaron Mab 4G2 y peroxidasa conjugada Mab 6B6C-1 de rábano radicular para capturar los antígenos del virus DEN y para detectar el antígeno capturado, respectivamente (Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice", *J. Virol.* 74: 4244-4252 (2000); Hunt et al., "A recombinant particulate antigen of Japanese encephalitis virus produced in stably-

transformed cells is an effective moninfectious antigen and subunit immunogen", *J. Virol. Methods*, 97: 133-149 (2001)).

Cuarenta y ocho horas después de la electroporación, las células transformadas para cada plasmida se tripsinizaron y resuspendieron en PBS como alícuotas que contenían células 5 x 10^6. Estas muestras de células se procesaron para extracción de la proteína de la membrana utilizando el kit reactivo de extracción de proteína de membrana mamífera Mem-PER (Pierce, Rockford, IL) siguiendo el protocolo sugerido por el fabricante. Las proteínas tanto hidrofóbicas como hidrofílicas están aisladas. Este procedimiento se desarrolló para el enriquecimiento de la proteína de la membrana integral encontrada en la fase hidrofóbica. Las fracciones tanto hidrofóbicas como hidrofílicas se analizaron mediante la captura-AG ELISA para el antígeno recombinante del DEN-2.

El antígeno recombinante en el medio se concentró mediante precipitación con polietilenglicol al 10% (PEG)-8000. El precipitante fue resuspendido en el tampón TNE (50 mM Tris, 100 mM NaCI, 10mM EDTA; pH 7.5) hasta 1/100° del volumen original, se aclaró mediante centrifugación y se almacenó a 4°C. El antígeno recombinante concentrado mediante precipitación PEG y resuspendido en el tampón TNE se extrajo con etanol al 4.0% hasta remover el PEG residual (Hunt et al., "A recombinant particulate antigen

of Japanese encephalitis virus produced in stably-transformed cells is an effective moninfectious antigen and subunit immunogen", *J. Virol. Methods,* 97: 133-149 (2001)). Los antígenos extraídos con etanol, las proteínas de la membrana hidrofóbica de las células transformadas y los viriones DEN-2 del gradiente purificado se analizaron en un NuPAGE, gel de gradiente Bis-Tris 4-12% en un Aparato de Electroforesis Excel Plus™ (Invitrogen Corp., Carlsbad, CA), seguido por electrosecante (electroblotting) en membranas de nitrocelulosa utilizando una unidad Blot Excel Plus (Invitrongen Corp.). Se detectó la proteína específica del virus DEN-2 con secante Western utilizando MAbs 1A6A-8 (específicos de E) del virus DEN-2 y 1A2A-1 (específico del cápsido), así como también antisuero de conejo específico para el PrM del DEN-2 y suero de ratón específico para un péptido compuesto de aminoácido 1-34 de la proteína DEN-2 M y se utilizó fluido ascítico de ratón normal como control negativo (Murray et al., "Processing of the dengue virus type 2 proteins prM and C-prM", *J. Gen Virol.* 74 (Pt 2): 175-182 (1993); Roehrig et al., "Monoclonal antibody mapping of the envolope glycoprotein of the dengue 2 virus, Jamaica", *Virology* 246: 317-328 (1998)).

vi. Vacunación de Ratón. En el estudio se utilizaron grupos de diez ratones hembras de 3 semanas de nacidos

reproducios con ICR. A los ratones se les inyectó i.m. con pCBD2-14-6, pCB9D2-1J-4-3, pCB8D2-2J-2-9-1 o pEGFP en la semana 0 y la semana 3 a una dosis de 100 µg en un volumen de 100 µg por ratón. El DNA plasmida se purificó a partir de las células azules XL-1 con Kits EndoFree Plasmid Giga™ (Qiagen) y se resuspendió en PBS, pH 7.5, a una concentración de 1.0 µg/µl. Los ratones que recibieron 100 µg de pEGFP se utilizaron como controles vacunados con plasmida. Los ratones se sangraron cada 3 semanas después de la inyección y la respuesta del anticuerpo específico del virus DEN-2 se evaluó mediante la utilización de ELISA indirecto y PRNT.

vii. Pruebas serológicas. Se probaron especimenes séricos antes y después de la vacunación para la capacidad aglutinante del anticuerpo al virion purificado DEN-2 mediante ELISA, para neutralizar (Nt) el anticuerpo mediante PRNT, y para anticuerpos que reconocen proteínas de virus purificado DEN-2 mediante secante western (Western blotting). Se realizó el PRNT con células Vero, tal como se describió previamente (Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice", *J. Virol.* 74: 4244-4255 (2000)) utilizando virus DEN-2 (cepa-16881) y JE (cepa-Nakayama). Los puntos extremos se determinaron a un nivel de reducción de placa del 90% (Hunt et al., "A recombinant

particulate antigen of Japanese encephalitis virus produced in stably-transformed cells is an effective moninfectious antigen and subunit immunogen", *J. Virol. Methods*, 97: 133-149 (2001)).

d. Resultados.

i. Expresión transitoria del antígeno recombinante del virus DEN-2. La expresión de los genes prM y E del virus DEN-2 o de un gen E quimérico de una combinación de secuencias de virus DEN-2 y JE (80% DEN-20% JE o 90% DEN-10% JE) se logró mediante transformaciones separadas de cada una de las tres plasmidas recombinantes DEN-2 DNA en células COS-1. El diseño básico de plasmida se basó en resultados de estudios anteriores con plasmidas recombinantes del virus JE y del virus WN en los que se expresaron las células transformadas en plasmida y secretaron proteínas virales auténticas en el fluido de cultivo de la célula (Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice", *J. Virol.* 74: 4244-4252 (2000); (Davis et al. "West Nile virus recombinant DNA vaccine protects mouse and horse from virus challenge and expresses in vitro a noninfectious recombinant antigen that can be used in enzyme-linked immunosorbent assays", *J. Virol.* 75: 4040-4047 (2001)). La expresión transciente de las proteínas recombinantes DEN-2 inicialmente se

evaluó mediante captura-Ag ELISA de los licores flotantes del cultivo de la célula y mediante IFA de células COS-1 transformadas, fijadas con acetona (Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice", *J. Virol.* 74: 4244-4252 (2000)). Se determinó que el punto óptimo de la expresión del antígeno fue de 48 horas después de la electroporación.

ii. Diagramación epítope de la proteína E expresada mediante células COS-1 transformadas transitoriamente.- La proteína DEN-2 expresada por cada uno de los plasmidas recombinantes se evaluó mediante IFA utilizando un panel de MAbs murina con reactividad conocida al virus DEN-2 (Tabla 10; Henchal et al., "Epitopic analysis of antigenic determinants on the surface of dengue-2 virions using monoclonal antibodies", *Am. J. Trop. Med. Hyg.* 34:162-169 (1895); Roehrig et al., "Monoclonal antibody mapping of the envolope glycoprotein of the dengue 2 virus, Jamaica", *Virology* 246: 317-328 (1998)). El panel MAb incluyó anticuerpos reactivos con cada uno delos tres dominios antigénicos de la proteína E del flavivirus así como proteínas prM y C. (Mandl et al., "Antigenic structure of the flavivirus envelope protein E at the molecular level, using tick-borne encephalitis virus as a model", *J. Virol.* 63: 564-571 (1989); Rey et al., "The envelope glycoprotein from tick-borne encephalitis virus

at 2 A resolution", *Nature* 375: 291-298 (1995)). El MAbs específico para los Dominios 2 y 3 antigénicos del flavivirus mostró reactividad cualitativa casi idéntica con el virus DEN-2 y cada una de las tres proteínas expresadas con plasmida. Uno de los MAbs específicos del Dominio 1, 1B4C-2, también mostró un patrón de reactividad similar con todas las proteínas expresadas. Sin embargo, dos de los MAbs específicos del Dominio 1, el 2B3A-1 y el 9A4D-1, fueron mucho menos reactivos con la proteína E expresada mediante los plasmidas pCBD2-14-6 y pCB9D2-1J-4-3, tal como se muestra por la titulación del punto final (valores en paréntesis, Tabla 10). La comparación de las titulaciones del punto final revelaron la expresión aparentemente pobre de los epítopes C3 y C4 en estructuras que contienen 100% DEN-2 E y 90% DEN-2 E-10%JE E. El MAbs 2H2, específico para prM tuvo la misma reactividad con el antígeno expresado por todas las tres plasmidas. El MBb Anti-C 1A2A-1 reaccionó bien con el virus DEN-2 y tuvo reactividad de bajo nivel, no específica con las proteínas virales expresadas en plasmida que incluyeron prM y E pero no C.

iii. Comparación de la proteína secretada y de la proteína con membrana ligada producida por cada uno de los tres plasmidas recombinantes DEN-2.- Se cosecharon cantidades similares de fluido de cultivo de célula a partir de las células COS-1 48 horas después de la

transformación para cada plasmida recombinante DEN-2. El antígeno recombinante secretado encontrado en el fluido del cultivo se concentró 100 veces mediante precipitación PEG, seguida por extracción de etanol para remover el PEG que interfirió con los análisis posteriores, mediante electroforesis de gel de poliacrilamida. La cantidad relativa de antígeno secretado expresado por cada plasmida se determinó mediante análisis de captura-Ag ELISA tanto del precipitado PEG como de las preparaciones del fluido del cultivo de la célula extraída con etanol (Tabla 11). El antígeno secretado se detectó solamente a partir de las células transfectadas con pCB8D2-2J-9-1 (SEQ ID NO:34), que contenían genes E del DEN-2 al 80% y del JE al 20%. Los plasmidas recombinantes que contenían bien sea genes E del DEN-2 al 100% o del DEN-2 al 90% o del JE al 10% no produjeron antígeno detectable por ELISA en el fluido del cultivo, a pesar de los esfuerzos para concentrar la proteína expresada.

También se utilizó análisis de secante western (Western Blot) para evaluar la producción de antígeno secretado por cada uno de los plasmidas recombinantes de DEN-2. Para propósitos de comparación, se corrieron volúmenes equivalentes de precipitado de PEG, licor flotante del cultivo de la célula extraída con etanol, en gels de gradiente NuPAGE, electrosecados a nitrocelulosa, y se analizaron utilizando MAbs o antisueros policlonales

capaces de reaccionar con todas las proteínas estructurales del DEN-2 (Figura 8A). El análisis de secante western mostró mayor sensibilidad en la detección del antígeno recombinante que en la captura-Ag ELISA puesto que las proteínas específicas al DEN-2 se detectaron en el fluido del cultivo de dos de las plasmidas, pCB8D2-2J-9-1 y pCB9D2-1J-4-3 (SEQ ID NO:46 y 44, respectivamente). La plasmida pCB8D2-2J-9-1 (SEQ ID NO:46) expresó la cantidad mayor de antígeno secretado que se mostró estaba compuesta de proteínas E, prM, y M. Se produjo antígeno relativamente menos secretado por pCB9D2-1J-4-3 (SEQ ID NO:44) y se encontraron niveles apenas detectables para la preparación pCBD2-14-6 (SEQ ID NO:42), que parecieron contener relativamente menos proteína E expresada, especialmente si se tomaron en consideración la reactividad no específica del MAb E-específico, 1A6A-8 en el control pEGFP (Figura 8A, líneas a, b para 14-6 y GFP).

Como E, prM y M son proteínas asociadas con membrana a lo largo de su síntesis intracelular, cualquier evaluación de la expresión de estas proteínas mediante las tres plasmidas recombinantes DEN-2 debe incluir una evaluación de las preparaciones de la membrana de la célula a partir de células de plasmida transformada. El kit de Reactivo de Extracción de Proteína de Membrana Mamífera Mem-PER (Pierce) se utilizó para aislar las

proteínas de membrana integral de los números equivalentes de células transformadas por cada una de las plasmidas recombinantes. Las proteínas hidrofóbicas se separaron de las proteínas hidrofílicas mediante partición de fases. Los análisis preliminares realizados con captura-AG de ELISA indicaron que la fracción de proteína hidrofílica no fue reactiva; sin embargo, las fracciones de proteína hidrofóbica de las células transformadas COS-1 con cada uno de los plasmidas recombinantes DEN-2 tuvo titulaciones similares en las pruebas ELISA (Tabla 11). Estos resultados indicaron que el antígeno recombinante codificado con todos los tres plasmidas se expresó después de la transformación, pero que los antígenos recombinantes expresados no fueron todos secretados al mismos nivel.

La confirmación de los resultados de la captura-Ag para las fracciones de proteína hidrofóbica estuvieron acompañados por secante western (Figura 8B). Se diluyeron volúmenes equivalentes de fracciones de proteína hidrofobica de cada una de las células plasmidotransformadas de acuerdo con las recomendaciones del fabricante para la electroforesis de gel de poliacrilamida-SDS con el fin de reducir la distorsión de banda y de línea. El inmunosecante con MAb E-, prM-, C- y M- específicos o con antisueros policlonales demostraron que todos los tres plasmidas recombinantes del DEN-2

indujeron la producción de cantidades similares de antígeno recombinante compuesto de E y de prM. No se detectó ninguna proteína M, bien sea porque no se procesó a partir de prM o porque los niveles fueron demasiado bajos para ser detectados. A pesar de los esfuerzos para reducir la distorsión de banda, los altos niveles de detergente en las muestras de proteína hidrofóbica aparentemente causados por E y por prM para correr en una forma ligeramente aberrante (migración más lenta) se compararon con muestras sin esas altas concentraciones de detergentes (comparar la migración E y prM en la Figura 8A y 8B).

iv. Comparación de la respuesta inmune en ratones vacunados con tres plasmidas de DNA recombinantes diferentes del DEN-2.- Los ratones ICR de tres semanas de nacidos se inmunizaron mediante inyección i.m. con 100 µg de pCB8D2-2J-9-1 (SEQ ID NO:46), pCB9D2-1J-4-3 (SEQ ID NO:44, pCBD2-146 (SEQ ID NO:42), o PEGFP en las semanas 0 y 3. Los ratones se sangraron a las 3, 6 y 9 semanas después de la inmunización primaria. Se probaron los sueros individuales y agrupados mediante ELISA indirecto, utilizando diluciones para el examen de 1:100 y 1:400 a la 3ª y 6ª semanas siguientes a la vacunación y titulaciones del punto final 9 semanas después de la vacunación. También se probaron sueros de 9 semanas mediante PRNT con virus tanto DEN-2 como JE. Los

resultados ELISA mostraron que después de una inmunización (sueros de 3 semanas) todos los ratones a quienes se les había administrado pCB8D2-2J-9-1 habían seroconvertido, mientras que solamente el 50% de los ratones vacunados con pCB9D2-1J-4-3 y 20% de los vacunados con pCBD2-14-6 reaccionaron con el virus DEN-2 (Tabla 12). A las 9 semanas siguientes a la vacunación, todos los ratones vacunados bien sea con pCB8D2-2J-9-1 o pCBD2-1J-4-3 demostraron reactividad ELISA anti-DEN-2; sin embargo, las titulaciones medias geométricas difirieron significativamente (titulaciones de 1:20.000 versus 1:708, respectivamente). Solamente 40% de los ratones inmunizados con pCBD2-14-6 tuvieron titulaciones ELISA anti-DEN-2 superiores a 1:100. Un secante western de sueros agrupados de 9 semanas provenientes de ratones inmunizados con pCB8D2-2J-2-9-1 en virus purificado de DEN-2 mostró que la respuesta inmunodominante fue a la glicoproteína E. También se detectó una ligera reactividad al prM y al M.

Más significativamente, en términos de evaluar el potencial de la vacuna de las tres plasmidas DEN-2, la inducción del anticuerpo virus-neutralizante en 7 de 9 ratones inmunizados con pCB8D2-2J-9-1 (SEQ ID NO.46) se observó basándose en un punto final de reducción de placa de 90% (Tabla 10). Sin embargo, si se utiliza un punto final de neutralización de 50%, entonces todos los 9 de

los 9 sueros tienen titulaciones PRNT de ^{3} 1:40. El noventa por ciento de las titulaciones de neutralización fluctuaron entre 1:40 y >1:1.000 para los 7 sueros con actividad neutralizante. Ninguno de los ratones inmunizados con pCB9D2-1J-4-3 produjó anticuerpo neutralizante, y solamente 1 de los 10 sueros de ratones vacunados con pCBD2-14-6, neutralizó el virus, pero tuvo una titulación de solo 1:8.

Como dos de los plasmidas recombinantes, principalmente el pCB9D2-1J-4-3 (SEQ ID NO:44) y el pCB8D2-2J-9-1 (SEQ ID NO:46) contenían secuencias del gen E del virus JE, todos los sueros fueron también evaluados para la presencia de actividad neutralizante del virus JE. Sin embargo, no se detectó esta actividad en el 90% del punto final de neutralización para ratones en ninguno de los grupos de inmunización. No sorprendentemente, los ratones inmunizados con la plasmida de control pEGFP no exhibieron reactividad para los virus DEN-2 o JE.

e. Discusión.- Los mismos pasos utilizados anteriormente para las vacunas JE y WN se utilizaron inicialmente para construir un plasmida recombinante DEN-2, pCB9D2-14-16 (SEQ ID NO:42), consistente de la región del gen prM y E del DEN-2 auténtico. La diagramación antigénica de las proteínas DEN-2 expresadas por las células COS-1 transformadas por esta plasmida, utilizando un panel de MAb mediante IFA indicó que la proteína prM y

E tenía una intensidad fluorescente compatible y una reactividad MAb similar a la de las células infectadas con virus (Tabla 10). Sin embargo, estas células COS-1 transformadas por la plasmida que codifica la región prM y E del DEN-2 auténtico no secretaron antígeno DEN-2 detectable en el fluido del cultivo (medido para la captura de antígeno ELISA). Además, la vacunación que utilizó la plasmida codificando una región prM y E del DEN-2 auténtico no estimuló el anticuerpo neutralizante del virus anti-DEN-2 en los ratones inmunizados i.m. (Tabla 13). Es interesante anotar que la transformación de las células por pCBD2-14-6 originó una mancha fluorescente globular-punteada que sugirió que el terminal-C de la proteína E del DEN-2 puede contribuir a la señal de retención de la membrana de esa proteína. Este patrón de mancha IFA no se observó en las células de construcción formada JE o WN (Chang et al., "A single intramuscular injection of recombinant plasmid DNA induces protective immunity and prevents Japanese encephalitis in mice", *J. Virol.* 74: 4244-4252 (2000); (Davis et al. "West Nile virus recombinant DNA vaccine protects mouse and horse from virus challenge and expresses in vitro a noninfectious recombinant antigen that can be used in enzyme-linked immunosorbent assays", *J. Virol.* 75: 4040-4047 (2001)). Por consiguiente, a la luz de la observación efectuada de conformidad con las enseñanzas de esta solicitud de patente, dos plasmidas

adicionales, el pCB9D2-1J-4-3 (SEQ ID NO:44) y el pCB8D2-2J-9-1 (SEQ ID NO:46), en los se hizo manipulación apropiada de la secuencia DNA para el 10% o el 20% del E del terminal-C del DEN-2 para que fueran reemplazadas con la región correspondiente de proteína E del virus JE, respectivamente. En la Tabla 13 se muestra la eficacia relativa de las estructuras diferentes en la estimulación del anticuerpo ELISA del anti-DEN-2 detectable en ratones vacunados.

Estos resultados con compatibles son el modelo que las interacciones entre prM y E pueden influenciar los procesos del conjunto y secreción de partículas. Se puede encontrar soporte para este modelo en un estudio del virus de encefalitis transmitido por garrapatas que putativamente sugiere que la interacción entre el prM y el ectodominio del E están involucrados en el transporte intracelular del prM-E mediado por prM, y por consiguiente, en la secreción de una partícula similar al virus (Allison et al., "Mapping of functional elements in the stem-anchor region of tick-borne encephalitis virus envelope protein E", *J. Virol.* 73: 5606-5612 (1999)).

En el presente ejemplo, reemplazar una porción del terminal-C de la proteína E del DEN-2 con proteína E del JE, correspondiente a TBE HI^{pred} por TM2, originó la proteína prM del DEN-2 y la secreción de la proteína E quimérica. Sin embargo, en contraste, la reposición del

TM1 y el TM2 en el TBE solo causó una mejora menor en la secreción de antígeno. La porción principal de la proteína prM y E expresada bien sea por pCBD2-14-6 transformado por COS-1 o por la plasmida pCB9D2-4-3 permaneció ligada a la membrana (Tabla 13). Estos resultados indicaron que una secuencia de retención de membrana no identificada está localizada en la región del tallo del terminal-C de la proteína E del DEN-2. La sustitución de esta región del tallo del terminal-C con secuencia del virus JE remueve o hace ineficaz esta secuencia de retención.

Otros han afirmado que la proteína prM es esencial para mantener conformación y secreción adecuada de la proteína E durante la maduración prM-E (Aberle et al., "A DNA immunization model study with constructs expressing the tick-borne encephalitis virus envelope protein E in different physical forms". *J. Immnunol.* 163: 6756-6761 (1999), Allison et al., "Synthesis and secretion of recombinant tick-borne encephalitis virus protein E in soluble and particulate forms", *J. Virol.* 69:5816-5820 (1995)). Además, también se ha demostrado que el ectodominio de la proteína E interactúa con prM. Se ha estimado que esta interacción involucra una secuencia de aminoácido dentro de los residuos del aminoácido 200-327 del E en el virus de encefalitis Murray Valley (Guirakhoo et al., "The Murray Valley encephalitis virus prM protein

confers acid resistance to virus particle and alters the expression of epitopes within the R2 domain of E glycoprotein", *Virology* 191:921-931(1992)).

Muy posiblemente se mantienen interacciones adecuadas E y prM e integridad retenida de la estructura de la proteína E en la proteína expresada por todas las tres estructuras de DEN-2, por lo menos en cuanto se requieren para inmunoreactividad. Además, la reposición del 20% de E en el terminal-C en el pCB8D2-2J-2-9-1 originó una proteína que mantuvo 395 aminoácidos de la E del DEN-2 auténtico. Se espera que cualquiera de estas modificaciones tenga una influencia mínima en las interacciones E y prM-E y su influencia en la naturaleza antigénica de la proteína E quimérica. Como reemplazar la región del terminal-C de la E del DEN-2 con secuencia ancla-tallo JE no tuvo efecto en la reactividad MAbs (Tabla 10), la retención de la secuencia del DEN-2 a ser reemplazada solo puede ser opcional para lograr la respuesta inmunológica específica del DEN-2.

Previamente se mostró que la estructura de un plasmida que codifica una partícula subviral secretada de la proteína E y prM del virus de encefalitis transmitido por garrapatas fue superior en términos del alcance y funcionalidad de la respuesta del anticuerpo y en términos de respuesta al reto del virus, con respecto a otras estructuras que codificaron el dímero E soluble

truncado terminalmente C que fue secretado, la longitud total E que no fue secretada, o una E truncada que no fue suficientemente secretada (Aberle et al., "A DNA immunization model study with constructs expressing the tick-borne encephalitis virus envelope protein E in different physical forms". *J. Immnunol.* 163:6756-6761 (1999). Sin embargo, aquí hemos demostrado que la potencia de la vacuna de DEN-2 DNA está correlacionada con la secreción de prM/M y E (Tabla 13). La morfología y naturaleza física del prM y E secretados no se demostraron en este estudio. Sin embargo, el prM y el E, secretados por la estructura pCB8D2-2J-2-9-1 muy probablemente forman una partícula viral parecida. Se considera que la presentación de los antígenos protectores múltiples en la superficie de la partícula mejora la potencia de la vacuna de esta construcción.

Los intentos anteriores para los desarrollos de la vacuna DNA del virus DEN-2 han originado diversos grados de éxitos (Kochel et al., "Innoculation of plasmids expressing the dengue-2 envelope gene elicit neutralizing antibodies in mice", *Vaccine* 15:547-552 (1997); Konishi et al. "A DNA vaccine expressing dengue type 2 virus premembrane and envelope genes induces neutralizing antibodies and memory B cells in mice", *Vaccine* 18: 1133-1139 (2000)). Para mejorar el nivel de la eficacia, se ha utilizado la coinmunización del motivo

inmunoestimulatorio CpG que contiene plasmida pUC19, murina que expresa plasmida GM-CSF en el régimen de la vacuna, o reemplazar el terminal-C de aminoácido 43 de E con secuencia de retención de membrana asociada con lisosoma mejoró la respuesta del anticuerpo para la vacuna DEN-2 (Porter et al. "Protective efficacy of a dengue 2 DNA vaccine in mice and the effect of CpG immuno-stimulatory motifs on antibody responses", *Arch. Virol.* 143:997-1003 (1998), Raviprakash et al. "Synergistic Neutralizing Antibody Response to a Dengue Virus Type 2 DNA Vaccine by Incorporation of Lysosome-Associated Membrane Protein Sequences and Use of Plasmid Expressing GM-CSF", *Virology* 290:74-82 (2001)). Los motivos CpG no metilados directamente activan los macrófagos, las células supresoras naturales y los linfocitos para secretar citocinas y quimiocinas, y soportan el desarrollo de respuestas inmunes mediadas por citocinas Th1 (Manders et al. "Immunology of DNA vaccines: CpG motifs and antigen presentation", *Inflamm. Res.* 49: 199-205(2000)). Sin embargo, la inclusión del CpG podría sesgar el perfil citocino del huésped y por ello contribuir tanto al desarrollo de trastornos autoinmunes Th1 mediados por órganos específicos como interferir con la homeostasis inmune (Smith et al. "The regulation of DNA vaccines" *Curr. Opin. Biotech.* 12:299-303(2000)). También existe evidencia en ratones de que niveles excesivos de citocina, aún cuando aumentan la

respuesta de determinadas células ayudantes-T, pueden disminuir o cerrar la respuesta de otros actores en la respuesta inmune, conduciendo a inmunosupresión generalizada o inflamación crónica (Robertson et al. "Assuring the quality, safety, and efficacy of DNA vaccines" *Mol. Biotechnol.* 17:143-149 (2001)). Por consiguiente, la seguridad y la eficacia de la inmunización DNA de flavivirus se podría beneficiar con la manipulación de una plasmida de expresión para mejorar la transcripción y la translación y objetivar las proteínas prM y E para secreción que promueva el procesamiento y ensamblaje correctos de poliproteína (Chang et al., "Flavivirus DNA vaccines: current statuts and potencial", *Annals of NY Acad. Sci.* 951: 272-285 (2001)). Las mejoras futuras se pueden centrar en mejorar la toma de DNA por parte de las células que presentan el antigeno o por parte de las células musculares (Rodríguez et al.,"Enhacing DNA immunization", *Virology* 268:233-238 (2000)).

Tabla 1. Expresión Transitoria de las proteínas prM y E del JE mediante diversas plasmidas recombinantes en dos líneas de células transferidas.

Vector					Plasmida Recombinante	Intensidad/porcentaje de las células antigenopositivas IFA *	
	Promotor	Intron	Poli (A)	ORI		COS-1	COS-7
pCDNA3	CMV	No	BGH	SV40	pCDJE2-7	3+/40	3+/35
PCBamp	CMV	No	BGH	No	pCBJE1-14	3+/45	nd
pC1Bamp	CMV	Si	BGH	No	pC1BJES14	3+/39	nd
pCEP4	CMV	No	SV40	OriP	pCEJE	2+/4	2+/3
pREP4	RSV	No	SV40	OriP	pREJE	1+/3	1+/2
pRe/RSV	RSV	No	BGH	SV40	pRCJE	1+/3	1+/3
pCDNA3	CMV	No	BGH	SV40	pCDNA3/CAT	-	-

*Varias líneas de células fueron transformadas con pCDNA3/CAT (control negativo), pCDJE2-7, pCBJE1-14, pC1BJES14, pCEJEm pREJE, o pRCJE. Las células se tripsimizaron 48 horas más tarde y se probaron mediante un ensayo de anticuerpos inmunofluorescente y directo (IFA) con HIAF específico del virus JE. Los datos se presentan como la intensidad (escala de 1 + a 4+) y el porcentaje de las células positivas IFA. Las células transformadas pCDNA3/CAT se utilizaron como control negativo

000328

Tabla 2. Caracterización de las proteínas expresadas por un clon transformado establemente pCDJE2-7 (JE-4B) de células COS-1 con anticuerpos reactivos al virus JE.

Mab o antisuero	Actividad Biológica del Mab		Intensidad inmunofluorescente de las células	
	Especificidad	Función Biológica	Infectadas JEV	4B
Mab:				
MC3	Específica JEV		2+	2+
2F2	Específica JEV	HI, N	4+	4+
112	Específica JEV		4+	4+
503	Específica JEV	N	4+	3+
109	Subgrupo	HI	2+	1+
N.04	Subgrupo	HI, N	3+	4+
201	Subgrupo		1+	1+
203	Subgrupo		4+	3+
204	Subgrupo		2+	2+
301	Subgrupo	HI	2+	2+
504	Flavivirus		4+	4+
6B6C-1	Flavivirus		2+	2+
3B4C-4	VEE		-	-
HIAF:				
Anti-JEV			4+	3+
Anti-WEE			-	-
PBS			-	-

Tabla 3. Persistencia de la respuesta inmune en ratones inmunizados con vacuna pCDJE2-7 o JE-VEX.

	Titulación ELISA (log_{10})						Titulación $PRNT_{90\%}$		
	3 sem.	6 sem.	9 sem.	23 sem.	40 sem.	60 sem.*	3 sem.	6 sem.	9 sem.
1x pCDJE2-7	2.6-3.2	3.8-5.0	3.8-4.4	>3.2	>3.2	2.4, 2.4, 3.8, 4.4	<20	20	40-160
2x pCDJE2-7	2.6-3.8	4.4	3.8-4.4	>3.2	>3.2	2.6, 3.8, 3.8	<20	20-40	40-160
2x JE-VAX	2.6-3.8	4.4-5.0	3.8-5.6	>3.2	>3.2	<2,<2,<2, 4.4	<20	20-40	20-160
2x pCDNA3/CAT	<2	<2	<2	ND	ND	<2	<20	<20	<20

Se inocularon ratones con 1 o 2 dosis de 100 µg/de DNA plasmida, o 1/5 de la dosis humana de la vacuna VAX-JE. Se recopilaron sueros para prueba antes de la segunda inmunización.

* Titulaciones séricas individuales.

Tabla 4. La rata seropositiva porcentual dependiente de la edad en ratones después de vacunación con diversas vacunas JEV.

	3-días de nacidos		3-semanas de nacidos	
	3 semanas PV	7 semanas PV	3 semanas PV	7 semanas PV
JE-VAX	0	0	100	100
PCDNA3/CAT	0	0	0	0
PCDJE2-7	40	60	90	90
PC1BJES14	10	60	80	100
PCBJE1-14	80	100	100	100

Tabla 5. Protección contra el reto JEV en ratones de 8 semanas de nacidos después de la vacunación a los 3 días de nacidos con diversas vacunas JEV.

Vacuna	Seroconversión JEV anterior al reto	Días de tasa de supervivencia posteriores al reto (%)				
		6	7	8	9	21
JE-VAX	0	100	100	60	40	40
PCDNA3/CAT	0	100	80	30	30	30
PCDJE2-7	60	100	100	100	100	100
pC1BJES14	60	100	100	100	100	100
PCBJE1-14	100	100	100	100	100	100

Tabla 6. Evaluación de la capacidad del anticuerpo materno del ratón hembra vacunado con ácido nucléico-JEV para proteger a sus crías de la encefalitis JEV fatal.

Madre vacunada		Crías retadas JEV		
Vacuna	$PRNT_{90\%}$	Edad reto (días)	No. de supervivencia[1]	ELISA[2]
1 x pCDJE2-7	40	4	0/11	
2 x pCDJE2-7	80	4	12/12	12/12
2 x JE-VAX	20	3	0/16	
2 x pCDNA-3/CAT	<10	5	0/14	
1 x pCDJE2-7	20	15	5/11	5/5
2 x pCDJE2-7	40	14	8/12	7/8
2 x JE-VAX	80	13	5/5	5/5
2 x pCDNA-3/CAT	<10	14	0/14	

Los ratones fueron inoculados intramuscularmente con 1 o 2, 100 µg dosis de plasmida DNA, o subcutáneamente con dos dosis humanas 1/5 de vacuna JE-VAX. Se recopilaron sueros 9 semanas después de la vacunación para pruebas PRNT anteriores del apareamiento con machos no inmunes.

[1]: No hubo sobreviventes/total para cada titulación.

[2]: Número de animales JEV ELISA con anticuerpos positivos (titulación $\geq$ 1:400)/No. de los sobrevivientes; los sueros fueron recopilaos para prueba 12 semanas después del reto.

Tabla 7. Característica de los péptidos de señal y su vacuna de potenciales entre las estructuras de vacuna de flavivirus DNA

Plasmida	Secuencia de péptido de señal que precede la proteína prM	Probabilidad de Péptido de Señal[a] SP	AP	Corte Sitio C	protocolo /protección de inmunización
pSLE1	?LDTIMKRPSKKRGGTRSLLGLAALIGLASS/LQLLSTYQG (SEQ ID NO:32)	0.702	0.292	0.352	im x 2/Partial
pJME	MWLASLAVVIACAGA/MKLSNFQGK (SEQ ID NO:33)	0.998	0.000	0.778	im x 2/Partial
pCJEME	MNEGSIMWLASLAVVIACAGA/MKLSNFQGK (SEQ ID NO:34)	0.985	0.012	0.785	im x 2/100%
pCBJE1-14	MGRKQNKRGGNEGSIMWLASLAVVIACAGA/MKLSNFQGK (SEQ ID NO:35)	0.791	0.199	0.623	im x 1/100%
pcDNA3prM-E	MSKKRGGSETSVLMVIFMLIGFAAA/LKLSNFQGK (SEQ ID NO:36)	0.721	0.277	0.622	im x 4 / Partial gg x 2-4/100%
pCBWN	MGKRSAGSIMWLASLAVVIACAGA/VTLSNFQGK (SEQ ID NO:37)	0.976	0.024	0.526	im x 1/100%
p1012DZME	MNVLRGFKKEIGRMLNILNRRRRTAGMIIMLIPTVMA/FKLTTRNGE (SEQ ID NO:38)	0.165	0.778	0.164	id x 2/None
SV-PE	MVGLQKRGKRRSATDWMSWLLVITLLGMTLA/ATVRKERGD (SEQ ID NO:39)	0.943	0.056	0.899	im or gg x 2/100%
pWRG7077-RSSE	MGWLLVVVLLGVTLA/ATVRKERGD (SEQ ID NO:40)	1.000	0.000	0.912	gg x 2/100%
pWRG7077-CEE	MSWLLVITLLGMTLA/ATVRKERGD (SEQ ID NO:41)	0.999	0.000	0.821	gg x 2/100%

[a] Se aplicó el programa SignalP HMM para calcular las probabilidades del péptido de señal (SP), el péptido de anclaje (AP), y el sitio de división de la signalasa (sitio C).

Se utilizaron códigos de aminoácidos simples, y se destacaron los aminoácidos cargados con letras en negrilla y subrayadas. El sitio de división de la signalasa que separa el SP y el prM se indican mediante "/". Las vacunas DNA se inocularon mediante método intramuscular (im), intradérmico (id) o con pistola de genes (gg).

WO 02/081754

PCT/US02/10764

Tabla 8: Las respuestas de anticuerpo neutralizante (Nt) en ratones inmunizados con diferentes dosis de vacunas DNA combinadas del virus JE y WN.

	pCBWN+pCBJEI-14				pCB Control
Dosis por plasmida (μg)	100+100	40+40	20+20	10+10	100
Porcentaje de ratones con Nt:					
Virus WN/virus JE	100/100	100/70	70/0	60/0	0/0
Rango de titulación $PRNT_{90}$					
Virus WN	1:320 - 1:80	1:80 - 1:20	1:80 - <1:10	1:20 - <1:10	<1:10
Virus JE	1:40 - 1:10	1:10 - <1:10	<1:10	<1:10	<1:10

Se inyectaron i.m. grupos de diez ratones hembras de tres semanas de nacidos reproducidos ICR, con una sola dosis de plasmida combinada DNAs según lo indicado. Se ensayaron especimenes séricos recopilados 12 semanas después de la inmunización mediante prueba de neutralización de reducción de placa (PRNT). Se calcularon las titulaciones de punto final frente a los virus JE y WN basándose en la reducción de placa al 90% utilizando virus JE (cepa SA-14) y virus West Nile (cepa NY-6480), respectivamente.

WO 02/081754

PCT/US02/10764

Tabla 9. Oligonucleótidos utilizados para construir plasmidas de expresión prM-# del virus DEN-2 y la región de unión del E indicado DEN-2 quimérico y JE

100% DEN-2 prM-E:	
D2KasI-438[a]	5' TGTGCA***GGCGCC***TTCCATTTAACCACACGTAACG (SEQ ID NO:48)
CD2NotI-2402	5' TCGA***GCGGCCGC***TCAACTAATTAGGCCTGCACCATGACTC (SEQ ID NO:49)
90% DEN-2 E & 10% JE E:	
T7	5' CTTATCGAAATTAATACGACTCACTATAGG (SEQ ID NO:50)
CD2BstXI-2244	5' ATAGATTGCT***CCA***AACACT***TGG***TGG (SEQ ID NO:51)
JE-2281	5' ACTCCATAGGAAAAGCCGTTCACC (SEQ ID NO:52)
CSP6	5' GCGAGCTCTAGCATTTAGGTGACACTATAG (SEQ ID NO:53)

DEN-2 → ← JE

Unión 90 - 10 : Leu His Gln Val Fe Gli Gli Ala Fe Arg Tr (SEQ ID NO:55)
CTC CAC CAA GTG TTT GGT GGT GCC TTC AGA ACA (SEQ ID NO:54)

80% DEN-2 E & 20% JE E:	
T7	5' CTTATCGAAATTAATACGACTCACTATAGG (SEQ ID NO:56)
CD2BsmBI-2097	5' GAATT***CGTCTC***ACTTCCTTTCTTAAACCAGTTGAGCTTC (SEQ ID NO:57)
JEBsmBI-2175	5' GGAATT***CGTCTC***GGAAGCACGCTGGGCAAGG (SEQ ID NO:58)
CSP6	5' GCGAGCTCTAGCATTTAGGTGACACTATAG 3' (SEQ ID NO:59)

DEN-2 → ← JE

Unión 80-20 : Asn Trp Lis Lis Gli Ser Tr Leu Gli Lis Ala (SEQ ID NO:61)
AAC TGG TTT AAG AAA GGA AGC ACG CTG GGC GCC (SEQ ID NO:60)

[a] Los sitios de la enzima de restricción codificados en oliglonucleótidos se destacaron con negrilla, cursiva y subrayados.

WO 02/081754

Tabla 10 Caracterizacion de los epitopes de glicoproteina E del DEN-2 expresados por plasmidas recombinantes DEN-2 segun lo determinado por el ensayo de anticuerpos fluorescente indirecto (IFA)

Anticuerpo			Controles[a]		Estructura del Plasmida[a]	
MAb (Epitope)[b]	Dominio Antigénico C	PRNT[c]	Células Infectadas Con DEN-2	Células normales	pCBD2-14-6	pCB9D2-1J-4-3	pCB8D2-2J-2-9-1
4G2 (A1)	2	+/-	4+		4+	4+	4+
4E5 (A2)	2	Si	3+		3-4+	3-4+	2-3+
1B7 (A5)	2	Si	3-4+		4+	4+	2-3+
1B4C-2(C1)	1	No	3-4+ (8000)		2-3+ (4000)	2-3+ (4000)	2-3+ (8000)
2B3A-1 (C3)	1	No	3-4+ (≥3200)		3+ (100)	2+ (100)	2-3+ (≥3200)
9A4D-1 (C4)	1	No	3-4+		2-3+ (400)	1 3+ (400)	3+ (≥12800)
3H5 (B2)	3	Si	4+		4+	4+	4+
10A4D-2 (B3)	3	Si	2-3+		3-4+	3-4+	2-3+
1A1D-2 (B4)	3	Si	4+		3-4+	4+	3-4+
9D12-6		Si	2-4+		2-3+	2-3+	3-4+
2H2	prM	No	4+		4+	3-4+	3-4+
1A2A-1	Cápsido	No	2-3+		1+	2+	1-2+

PCT/US02/10764 000337

[a] Los substratos IFA fueron células COS-1 fijadas con acetona bien sea infectadas con DEN-2 16681 controles sin infectar o transformados con un plasmida recombinante DEN-2

[b] Se utilizaron anticuerpos monoclonales a una dilución óptima predeterminada basada en reactividad con el virus DEN-2 16681 Para algunos MAbs titulaciones de punto final que se muestran en paréntesis se reportan y para otros solo se reportan valores cualitativos basándose en una escala de 1+ a 4+ considerándose positivo 3-4+ 2+ equivocado y 1+ negativo

[c] Dominios antigénicos basados en la glicoproteína-E del virus TBE (Mandl et al Antigenic structure of the flavivirus envelope protein E at the molecular level using tick-borne encephalitis virus as a model" *J Virol* 63 564-571 (1989) Rey et al The envelope glycoprotein from tick-borne encephalitis virus at 2 A resolution" *Nature* 375 291-298 (1995))

[d] Actividad de neutralización de placas a una dilución de 1 100 de fluido ascítico utilizando un punto final de reducción de placa de 90% excepto para el 4G2 y el 9D12-6 para los cuales se reportó un punto final de neutralización de 50% (Henchal et al Epitopic analysis of antigenic determinants on the surface of dengue-2 virions using monoclonal antibodies" *Am J Trop Med Hyg* 34 162-169 (1895) Roehrig et al Monoclonal antibody mapping of the envolope glycoprotein of the dengue 2 virus Jamaica" *Virology* 246 317-328 (1998))

000338

Tabla 11: Detección de proteína recombinante secretada y con membrana del DEN-2, mediante captura de antígen ELISA.

Plasmida	Tipo de Muestra	Titulación ELISA del punto final
PCBD2-146	Fluido de cultivo precipitado PEG[a]	<1:10
PCBD2-146	Precipitado PEG, fluido de cultivo extraído con etanol[b]	<1:20
PCBD2-146	Preparación de la proteína de membrana hidrofóbica[c]	1:160
PCB9D2-1J-4-3	Fluido de cultivo precipitado PEG[a]	<1:10
PCB9D2-1J-4-3	Precipitado PEG, fluido de cultivo extraído con etanol[b]	<1:20
PCB9D2-1J-4-3	Preparación de la proteína de membrana hidrofóbica[c]	1:80
PCB8D2-2J-2-9-1	Fluido de cultivo precipitado PEG[a]	1:640
PCB8D2-2J-2-9-1	Precipitado PEG, fluido de cultivo extraído con etanol[b]	1:80
PCB8D2-2J-2-9-1	Preparación de la proteína de membrana hidrofóbica[c]	1:80
PEGFP	Fluido de cultivo precipitado PEG[a]	<1:10
PEGFP	Precipitado PEG, fluido de cultivo extraído con etanol[b]	<1:10
PEGFP	Preparación de la proteína de membrana hidrofóbica[c]	<1:10

[a]El licor flotante del cultivo de las células plasmidotransformadas se precipitó con polietilenglicol al 10% (PEG) y se resuspendió en 1/100° del volumen original.
[b]El licor flotante del cultivo precipitado PEG se extrajo con etanol al 4% para remover el PEG y el gránulo se resuspendió en 1/5 parte del volumen extraído.
[c]Se prepararon fracciones de la membrana hidrofóbica tal como se describe en los materiales y métodos.

Tabla 12. Inmunogenicidad de tres plasmidas recombinantes de DEN-2 en ratones ICR

		Elisa en virus DEN-2					PRNT en DEN-2 virus[a]	PRNT en JE virus[a]
		Examen 3 semanas p.v.[c]		Examen 6 semanas p.v.[c]				
Plasmid DNA[b]	Mouse #	1:100	1:400	1:100	1:400	Titulación punto final 9 semanas, p.v.	Titulación punto final 9 semanas, p.v.	Titulación punto final 9 semanas, p.v.
[illegible]	Grupo Prom. 1,2,4-10	ND[d]	ND	+	+	[illegible]	ND	ND
	1	+	+	+	+	[illegible]	>[illegible]	<[illegible]
	3	+	+	+	+	[illegible]	>[illegible]	<[illegible]
	4	+	+	+	+	[illegible]	[illegible]	<[illegible]
	5	+	+	+	+	[illegible]	<[illegible]	<[illegible]
	6	+	+	+	+	[illegible]	[illegible]	<[illegible]
	7	+	+	+	+	[illegible]	[illegible]	<[illegible]
	8	+	+	+	+	[illegible]	[illegible]	<[illegible]
	9	+	+	+	+	[illegible]	<[illegible]	<[illegible]
	10	+	+	+	+	[illegible]	>[illegible]	<[illegible]
[illegible]	Grupo Prom. 1-10	ND	ND	+	+	[illegible]	ND	<[illegible]
	1	+	+	+	+	[illegible]	<[illegible]	ND
	3	+	+	+	+	[illegible]	<[illegible]	ND
	3	+	+	+	+	[illegible]	<[illegible]	[illegible]
	4	+	+	+	+	[illegible]	<[illegible]	ND
	5	+	+	+	+	[illegible]	<[illegible]	ND
	6	+	+	+	+	[illegible]	<[illegible]	[illegible]
	7	+	+/-	+	+	[illegible]	<[illegible]	ND
	8	+	+	+	+	[illegible]	<[illegible]	ND
	9	+	+	+	+	[illegible]	<[illegible]	<[illegible]
	10	+	+	+	+	[illegible]	<[illegible]	<[illegible]
[illegible]	Grupo Prom. 1-10	ND	ND	+	+	[illegible]	<[illegible]	<[illegible]
	1	+	+	+	+	[illegible]	<[illegible]	ND
	2,3,6-9	+	+	+	+	[illegible]	ND	ND
	4	+	+	+	+	[illegible]	<[illegible]	<[illegible]
	5	+	+	+	+	[illegible]	[illegible]	<[illegible]
	10	+	+	+	+	[illegible]	ND	ND
[illegible]	Prom. 1-10 Grupo	+	ND	+	ND	[illegible]	<[illegible]	<[illegible]

[a] PRNT prueba de neutralización de reducción de placa, punto final de neutralización 90%

[b] Los ratones se inmunizaron intramuscularmente con 100 µg de DNA plasmida en las semanas 0 y 3.

[c] Los exámenes ELISA utilizaron sueros diluidos a 1:100 y 1:400

[d] ND, no se hizo

[e] Grupo, 1, 2, 4, 5, 7, 8

[F] Grupo, 2,3,6-10

[g] Grupo, 1-3, 6-10

Tabla 13. Resumen de las características de tres plasmidas recombinantes del DEN-2

Plasmida	IFA[a]		Titulación ELISA captura-AG		Titulación ELISA en DEN-2[b]		DEN-2 PRNT[c]
	+/-	Globular / Difuso	Antigeno Secretado	prep. proteína Membrana Hidrofóbica	No. suero ≥ 1:100	titulación sueros agrupados	No. suero ≥1:10
PCB8D2-2J-2-9-1	+	Difuso	1:640	1:80	9/9	1:64000	7/9[d]
PCB9D2-1J-4-3	+	Globular	<1:10	1:80	10/10	1:1000	0/10
PCB8D2-14-6	+	Globular	<1:10	1:80	3/10	1:200	0/10

[a] Características de las manchas del ensayo de anticuerpo fluorescente directo (IFA), + o -, y patrón difuso o globular.

[b] Titulación ELISA ant-DEN-2 de sueros de ratones inmunizados con los plasmidas recombinantes. Se recopilaron los sueros 9 semanas después de la vacunación (semanas 0 y 3). El número de ratones con titulación de ≥ 1:100/número total de ratones se muestra, incluyendo la titulación ELISA del punto final de la muestra del suero agrupado.

[c] Número de ratones con titulaciones de neutralización de reducción de placa (PRNT, reducción al 90%) ≥ 1:10/número total de ratones. Los sueros se recopilaron 9 semanas después de la vacunación.

[d] De los 7 ratones con anticuerpo neutralizante, 3 ratones tuvieron titulaciones PRNT de ≥1:1000, 3 tuvieron titulaciones de ≥1:100<1:1.000 y uno tuvo una titulación de 1:40.

VACUNAS DE ÁCIDO NUCLÉICO PARA PREVENCIÓN DE LA INFECCIÓN DE FLAVIVIRUS

Lo que se reivindica es:

1. Un ácido nucléico aislado que comprende una unidad transcripcional que codifica una secuencia de señal de una proteína estructural de un primer flavivirus y un antígeno de flavivirus inmunogénico de un segundo flavivirus, en donde la unidad transcripcional dirige las síntesis del antígeno.

2. El ácido nucléico de la reivindicación 1, en donde la secuencia de señal es una secuencia de señal de virus de encefalitis japonesa.

3. El ácido nucléico de la reivindicación 1, en donde el antígeno del flavivirus inmunogénico es de un flavivirus seleccionado entre el grupo que consiste de virus de fiebre amarilla, virus dengue serotipo 1, virus dengue serotipo 2, virus dengue serotipo 3, virus dengue serotipo 4, virus de encefalitis japonesa, virus Powassan y virus West Nile.

4. El ácido nucléico de la reivindicación 1, en donde la unidad transcripcional codifica una secuencia de señal del virus de encefalitis japonesa y una proteína M y una proteína E del virus West Nile.

5. El ácido nucléico de la reivindicación 1, en donde la unidad transcripcional codifica una secuencia de señal del virus de encefalitis japonesa y una proteína M y una proteína E del virus de fiebre amarilla.

6. El ácido nucléico de la reivindicación 1, en donde la unidad transcripcional codifica una secuencia de señal del virus de encefalitis japonesa y una proteína M y una proteína E del virus de encefalitis St. Louis.

7. El ácido nucléico de la reivindicación 1, en donde la unidad transcripcional codifica una secuencia de señal del virus de encefalitis japonesa y una proteína M y una proteína E del virus Powassan.

8. El ácido nucléico de la reivindicación 1, en donde el antígeno se selecciona entre el grupo que consiste de una proteína M de un flavivirus, una proteína E de un flavivirus, tanto una proteína M como una Proteína E de un flavivirus, una parte de una proteína M de un flavivirus, una parte de una proteína E de un flavivirus y tanto una porción de una proteína M de un flavivirus como una porción de una proteína E de un flavivirus o una combinación de las mismas.

9. El ácido nucléico de la reivindicación 8, en donde el antígeno es tanto la proteína M como la proteína E de un flavivirus.

10. El ácido nucléico de la reivindicación 1, en donde el ácido nucléico es DNA.

11. El ácido nucléico de la reivindicación 10, que comprende una secuencia nucleótida seleccionada entre el grupo que consiste de SEQ ID NO:15, SEQ ID NO:19, SEQ ID NO:21 y SEQ ID NO:23.

12. El ácido nucléico de la reivindicación 1, en donde la unidad transcripcional comprende una secuencia de control dispuesta apropiadamente de forma tal que controla en forma operable la síntesis del antígeno.

13. El ácido nucléico de la reivindicación 12, en donde la secuencia de control es el promotor precoz inmediato del citomegalovirus.

14. El ácido nucléico de la reivindicación 1, que comprende una secuencia de consenso Kozak localizada en el sitio de inicio translacional para un polipétido que comprende el antígeno codificado por la unidad transcripcional TU.

15. El ácido nucléico de la reivindicación 1, en donde la unidad transcripcional comprende un terminador poli-A.

16. Una célula que comprende el ácido nucléico de la reivindicación 1.

17. Una composición que comprende el ácido nucléico de la reivindicación 1 y un portador aceptable desde el punto de vista farmacéutico.

18. Un método de inmunizar a un sujeto contra infección por flavivirus, que comprende administrarle una cantidad eficaz de la composición de la reivindicación 17.

19. El método de la reivindicación 18, en donde el antígeno del flavivirus es de un flavivirus seleccionado entre el grupo que consiste del virus de fiebre amarilla, virus dengue serotipo 1, virus dengue serotipo 2, virus dengue serotipo 3, virus dengue serotipo 4, virus de encefalitis japonesa, virus Powassan y virus West Nile.

20. El método de la reivindicación 18, en donde el antígeno se selecciona entre el grupo que consiste de una proteína M de un flavivirus, una proteína E de un flavivirus, tanto una proteína M como una proteína E de un flavivirus, una porción de una proteína M de un flavivirus, una porción de una proteína E de un flavivirus y tanto una porción de una proteína M como una porción de una proteína E de un flavivirus, o cualquier combinación de las mismas.

21. El método de la reivindicación 20, en donde el antígeno es tanto la proteína M como la proteína E de un flavivirus, y donde una célula dentro del cuerpo del sujeto, después de incorporar el ácido nucléico dentro de

ella, secreta partículas subvirales que comprenden la proteína M y la proteína E.

22. El método de la reivindicación 18, en donde la unidad transcripcional codifica una secuencia de señal del virus de encefalitis japonesa y una proteína M y una proteína E del virus West Nile.

23. El método de la reivindicación 18, en donde la unidad transcripcional codifica una secuencia de señal del virus de encefalitis japonesa y una proteína M y una proteína E del virus de fiebre amarilla.

24. El método de la reivindicación 18, en donde la unidad transcripcional codifica una secuencia de señal del virus de encefalitis japonesa y una proteína M y una proteína E del virus de encefalitis St. Louis.

25. El método de la reivindicación 18, en donde la unidad transcripcional codifica una secuencia de señal del virus de encefalitis japonesa y una proteína M y una proteína E del virus Powassan.

26. El método de la reivindicación 18, que comprende administrar la composición al sujeto en una sola dosis.

27. El método de la reivindicación 18, en donde la composición se administra vía parenteral.

28. El ácido nucléico de la reivindicación 1, en donde el antígeno es un antígeno de virus de encefalitis St. Louis.

29. El ácido nucléico de la reivindicación 18, en donde el antígeno es un antígeno del virus de encefalitis St. Louis.

30. El ácido nucléico de la reivindicación 1, en donde el antígeno es un antígeno de virus de encefalitis japonesa.

31. El método de la reivindicación 18, en donde el antígeno es un antígeno del virus de encefalitis japonesa.

32. El ácido nucléico de la reivindicación 1, en donde el antígeno es un antígeno de virus de fiebre amarilla.

33. El método de la reivindicación 18, en donde el antígeno es un antígeno de virus de fiebre amarilla.

34. El ácido nucléico de la reivindicación 1, en donde el antígeno es un antígeno de virus dengue.

35. El método de la reivindicación 18, en donde el antígeno es un antígeno del virus dengue.

36. El ácido nucléico de la reivindicación 1, en donde el antígeno es un antígeno de virus West Nile.

37. El método de la reivindicación 18, en donde el antígeno es un antígeno de virus West Nile.

38. El antígeno producido a partir del ácido nucléico de la reivindicación 1.

39. Un método de detectar un anticuerpo de flavivirus en una muestra, que comprende:

(a) poner en contacto la muestra con el antígeno de la reivindicación 38 bajo condiciones por medio de las cuales se pueda formar un antígeno/anticuerpo complejo; y

(b) detectar la formación del antígeno/anticuerpo complejo, detectando así un anticuerpo de flavivirus en la muestra.

40. Un anticuerpo producido en respuesta a la inmunización mediante el antígeno de la reivindicación 38.

41. Un método de detectar un antígeno de flavivirus en una muestra, que comprende:

(a) poner en contacto la muestra con el anticuerpo de la reivindicación 40 bajo condiciones por medio de las cuales se pueda formar un antígeno/anticuerpo complejo; y

(b) detectar la formación del antígeno/anticuerpo complejo, detectando por ello un antígeno de flavivirus en una muestra.

42. Un método de diagnosticar una infección de flavivirus en un sujeto, que comprende:

(a) poner en contacto una muestra del sujeto con el antígeno de la reivindicación 38 bajo condiciones por medio de las cuales se pueda formar un antígeno/anticuerpo; y

(b) detectar la formación del antígeno/anticuerpo complejo, diagnosticando así una infección de flavivirus en un sujeto.

43. Un método de diagnosticar una infección de flavivirus en un sujeto, que comprende:

(a) poner en contacto una muestra del sujeto con el anticuerpo de la reivindicación 40 bajo condiciones por medio de las cuales se pueda formar un antígeno/anticuerpo complejo; y

(b) detectar la formación del antígeno/anticuerpo complejo, diagnosticando así una infección de flavivirus en un sujeto.

Es traducción fiel y completa efectuada por María de Pilar Iglesias el día 20 de octubre de 2003, de un documento escrito en Inglés, al cual para su identificación se le estampa el sello de María del Pilar Iglesias, Traductora e Intérprete Oficial por Resolución

No. 819 de abril de 1971 expedida por el Ministerio de Justicia de la República de Colombia.

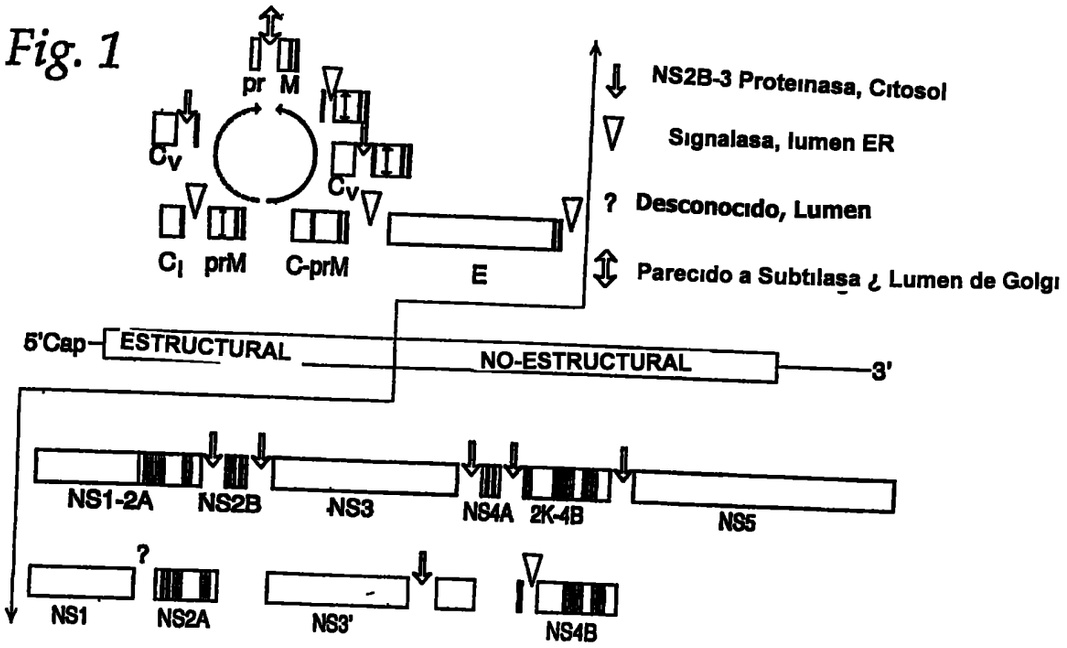
María del Pilar Iglesias.
María del Pilar Iglesias
C.C. 41.339.183 de Bogotá

María del Pilar Iglesias
C. C. 41.339.183 de Bogotá
Intérprete Oficial según Resolución No. 819 de
15 de Abril de 1971 del Ministerio de Justicia

352

PCT/US92/10764

Fig. 1



WO 93/06214

PCT/US92/10764

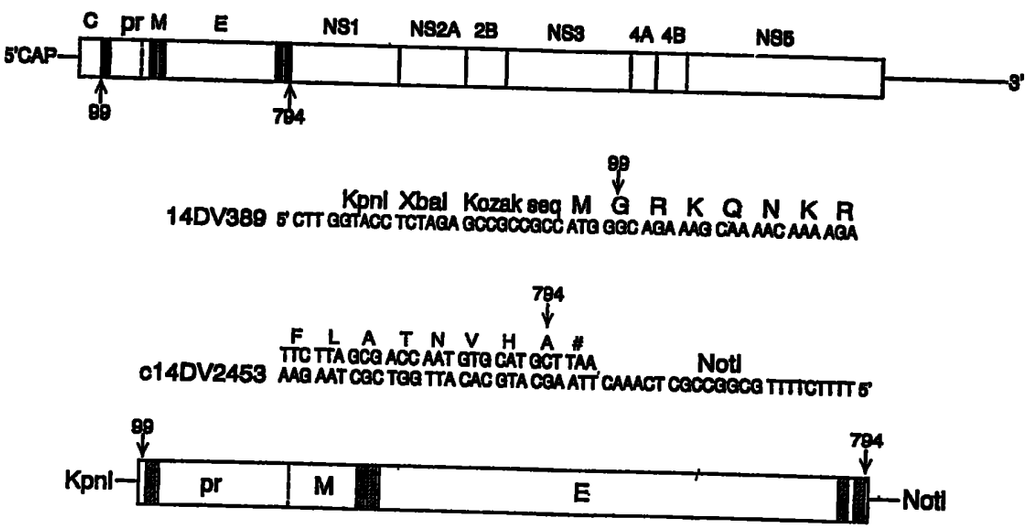


Fig. 2

WO 02/081754

PCT/US02/10764

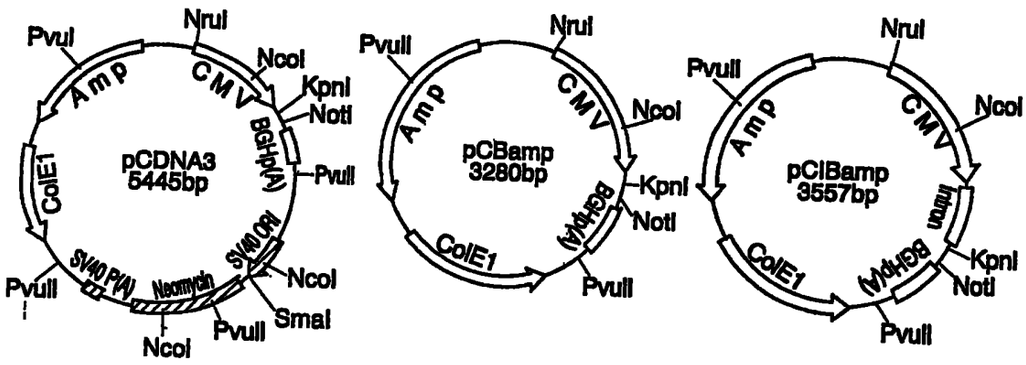


Fig. 3

WO 02/081754

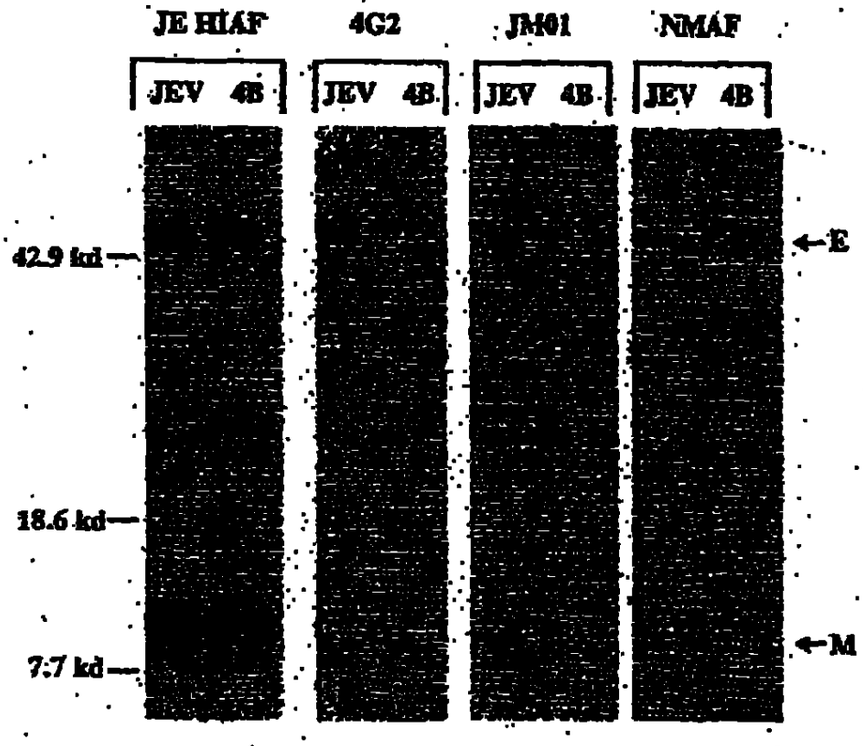

JE HIAF
4G2
JM01
NMAF
JEV 4B
JEV 4B
JEV 4B
JEV 4B
42.9 kd
18.6 kd
7.7 kd
E
M

Fig. 4

000356
PCT/US02/10754
5/8
WO 02/081754

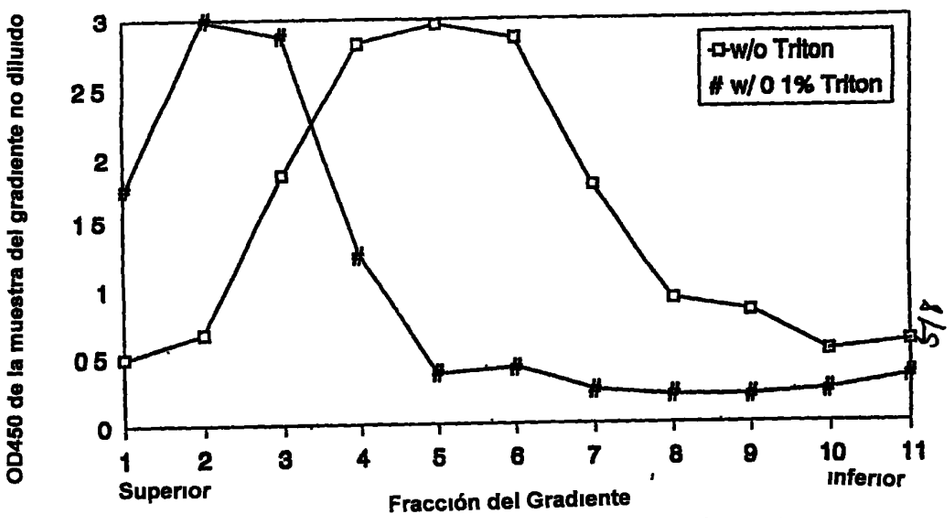


Fig. 5

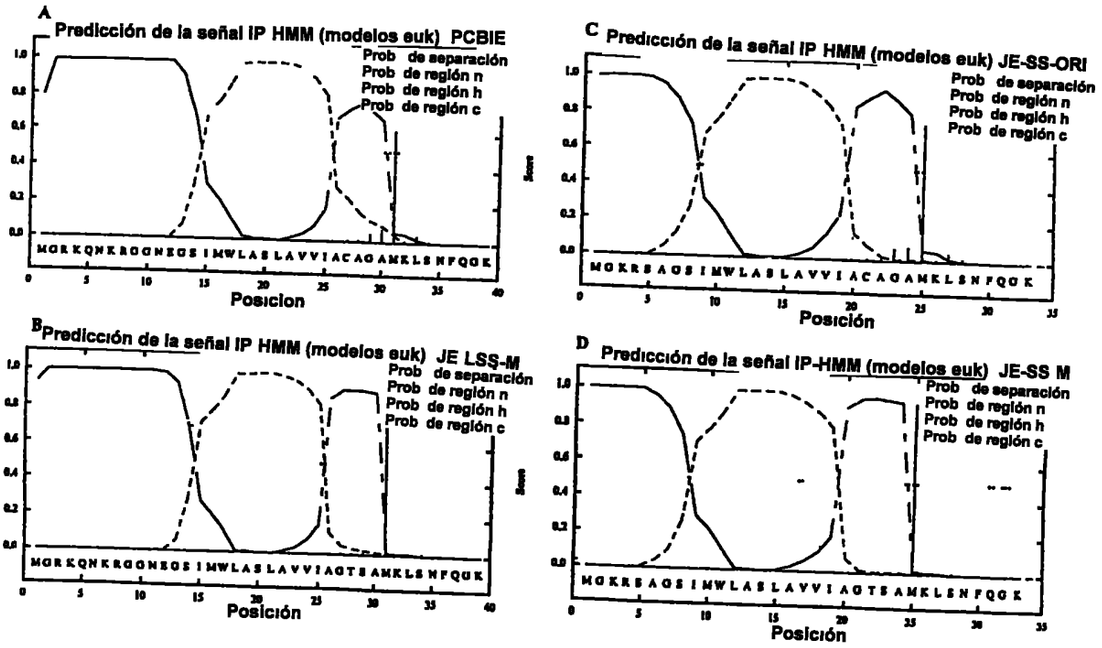
PCT/US02/10764

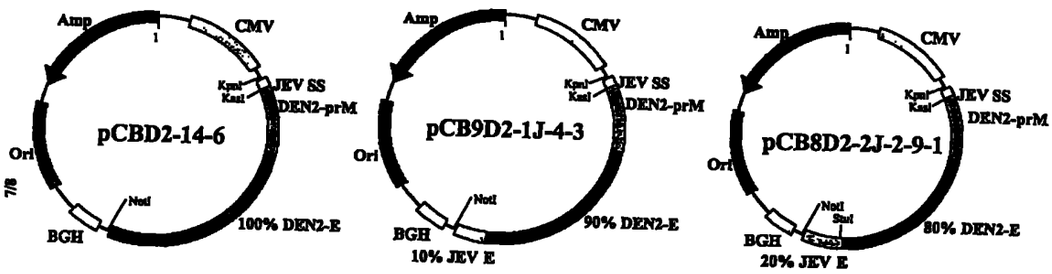
WO 02/081754

6/8

357

FIG 6





WO 02/081754

FIG 7

PCT/US02/19764

WO 02/081754

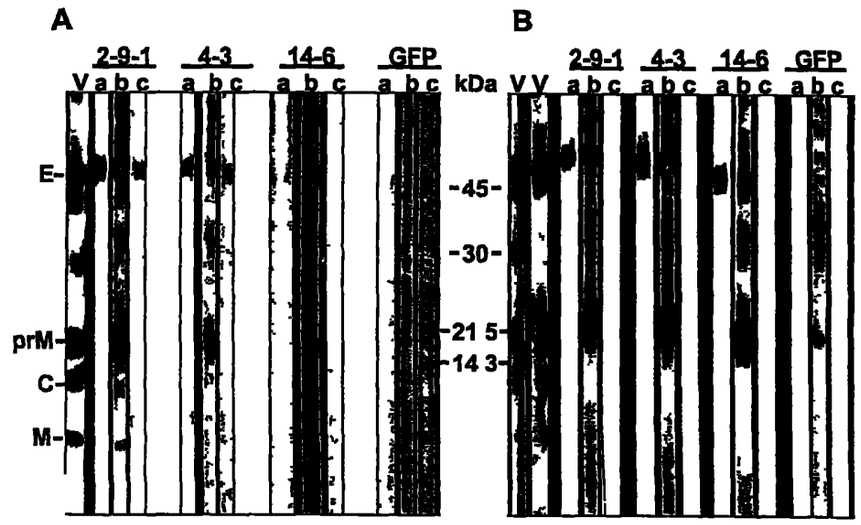


FIG 8

María del Pilar Iglesias
C. C. 41.339 183 de Bogotá
Intérprete Oficial según Resolución No 819 de
a de Abril de 1971 del Ministerio de Justicia

LISTADO DE SECUENCIAS

<110> El Gobierno de Estados Unidos de América, representado por la secretaría, Departamento de Salud y Servicios Humanos, c/o Centros para el control y prevención de enfermedades.

Chang, Gwong-Jen J

<120> Vacunas de ácido nucleico para prevención de la infección del flavivirus.

<130> 14114.0332P2

<150> 09/826,115
<151> 2001-04-04

<150> 09/701,536
<151> 2000-11-29

<150> PCT/US99/12298
<151> 1999-06-03

<150> 60/087,908
<151> 1998-06-04

<160> 61

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (25)...(48)

<221> misc_feature
<222> 1-48
<223> Amplimer 14DV389

<400> 1
cttggtacct ctagagccgc cgcc atg ggc aga aag caa aac aaa aga 48
Met Gly Arg Lys Gln Asn Lys Arg
1 5

<210> 2
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 2
Met Gly Arg Lys Gln Asn Lys Arg
1 5

<210> 3
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-50
<223> Amplimer c14DV2453

<400> 3
ttttcttttg cggccgctca aacttaagca tgcacattgg tcgctaagaa 50

<210> 4
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (25)...(48)

<221> misc_feature
<222> (1)...(48)
<223> Amplimer YFDV389

<400> 4
cttggtacct ctagagccgc cgcc atg cgt tcc cat gat gtt ctg act 48
Met Arg Ser His Asp Val Leu Thr
1 5

<210> 5
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 5
Met Arg Ser His Asp Val Leu Thr
1 5

<210> 6
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-41
<223> Amplimer cYFDV2452

<400> 6
ttttcttttg cggccgctca cgccccaact cctagagaaa c 41

<210> 7
<211> 51
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (25)...(54)

<221> misc_feature
<222> 1-54
<223> Amplimer SLEDV410

<400> 7
cttggtacct ctagagccgc cgcc atg tct aaa aaa aga gga ggg acc aga 51
Met Ser Lys Lys Arg Gly Gly Thr Arg
1 5

<210> 8
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 8
Met Ser Lys Lys Arg Gly Gly Thr Arg
1 5

<210> 9
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-38
<223> Amplimer cSLEDV2449

<400> 9
ttttcttttg cggccgctta ggcttgcacg ctggttgc 38

<210> 10
<211> 7500
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (916)...(3009)

<221> misc_feature
<222> 1-7500
<223> pCDJE 2-7

<400> 10
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc 900
gagctcgccg ccgcc atg ggc aga aag caa aac aaa aga gga gga aat gaa 951
Met Gly Arg Lys Gln Asn Lys Arg Gly Gly Asn Glu
1 5 10

ggc tca atc atg tgg ctc gcg agc ttg gca gtt gtc ata gct tgt gcg 999
Gly Ser Ile Met Trp Leu Ala Ser Leu Ala Val Val Ile Ala Cys Ala
15 20 25

gga gcc atg aag ttg tcg aat ttc cag ggg aag ctt ttg atg acc atc 1047
Gly Ala Met Lys Leu Ser Asn Phe Gln Gly Lys Leu Leu Met Thr Ile
30 35 40

aac aac acg gac att gca gac gtt atc gtg att ccc acc tca aaa gga 1095
Asn Asn Thr Asp Ile Ala Asp Val Ile Val Ile Pro Thr Ser Lys Gly
45 50 55 60

gag aac aga tgc tgg gtc cgg gca atc gac gtc ggc tac atg tgt gag 1143
Glu Asn Arg Cys Trp Val Arg Ala Ile Asp Val Gly Tyr Met Cys Glu
65 70 75

gac act atc acg tac gaa tgt cct aag ctt acc atg ggc aat gat cca 1191
Asp Thr Ile Thr Tyr Glu Cys Pro Lys Leu Thr Met Gly Asn Asp Pro
80 85 90

gag gat gtg gat tgc tgg tgt gac aac caa gaa gtc tac gtc caa tat 1239
Glu Asp Val Asp Cys Trp Cys Asp Asn Gln Glu Val Tyr Val Gln Tyr
95 100 105

```
gga cgg tgc acg cgg acc agg cat tcc aag cga agc agg aga tcc gtg     1287
Gly Arg Cys Thr Arg Thr Arg His Ser Lys Arg Ser Arg Arg Ser Val
    110                 115                 120

tcg gtc caa aca cat ggg gag agt tca cta gtg aat aaa aaa gag gct     1335
Ser Val Gln Thr His Gly Glu Ser Ser Leu Val Asn Lys Lys Glu Ala
125                 130                 135                 140

tgg ctg gat tca acg aaa gcc aca cga tat ctc atg aaa act gag aac     1383
Trp Leu Asp Ser Thr Lys Ala Thr Arg Tyr Leu Met Lys Thr Glu Asn
                145                 150                 155

tgg atc ata agg aat cct ggc tat gct ttc ctg gcg gcg gta ctt ggc     1431
Trp Ile Ile Arg Asn Pro Gly Tyr Ala Phe Leu Ala Ala Val Leu Gly
            160                 165                 170

tgg atg ctt ggc agt aac aac ggt caa cgc gtg gta ttt acc atc ctc     1479
Trp Met Leu Gly Ser Asn Asn Gly Gln Arg Val Val Phe Thr Ile Leu
        175                 180                 185

ctg ctg ttg gtc gct ccg gct tac agt ttt aat tgt ctg gga atg ggc     1527
Leu Leu Leu Val Ala Pro Ala Tyr Ser Phe Asn Cys Leu Gly Met Gly
    190                 195                 200

aat cgt gac ttc ata gaa gga gcc agt gga gcc act tgg gtg gac ttg     1575
Asn Arg Asp Phe Ile Glu Gly Ala Ser Gly Ala Thr Trp Val Asp Leu
205                 210                 215                 220

gtg ctg gaa gga gat agc tgc ttg aca atc atg gca aac gac aaa cca     1623
Val Leu Glu Gly Asp Ser Cys Leu Thr Ile Met Ala Asn Asp Lys Pro
                225                 230                 235

aca ttg gac gtc cgc atg att aac atc gaa gct agc caa ctt gct gag     1671
Thr Leu Asp Val Arg Met Ile Asn Ile Glu Ala Ser Gln Leu Ala Glu
            240                 245                 250

gtc aga agt tac tgc tat cat gct tca gtc act gac atc tcg acg gtg     1719
Val Arg Ser Tyr Cys Tyr His Ala Ser Val Thr Asp Ile Ser Thr Val
        255                 260                 265

gct cgg tgc ccc acg act gga gaa gcc cac aac gag aag cga gct gat     1767
Ala Arg Cys Pro Thr Thr Gly Glu Ala His Asn Glu Lys Arg Ala Asp
    270                 275                 280

agt agc tat gtg tgc aaa caa ggc ttc act gac cgt ggg tgg ggc aac     1815
Ser Ser Tyr Val Cys Lys Gln Gly Phe Thr Asp Arg Gly Trp Gly Asn
285                 290                 295                 300

gga tgt gga ctt ttc ggg aag gga agc att gac aca tgt gca aaa ttc     1863
Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Asp Thr Cys Ala Lys Phe
                305                 310                 315

tcc tgc acc agt aaa gcg att ggg aga aca atc cag cca gaa aac atc     1911
Ser Cys Thr Ser Lys Ala Ile Gly Arg Thr Ile Gln Pro Glu Asn Ile
            320                 325                 330

aaa tac gaa gtt ggc att ttt gtg cat gga acc acc act tcg gaa aac     1959
Lys Tyr Glu Val Gly Ile Phe Val His Gly Thr Thr Thr Ser Glu Asn
        335                 340                 345
```

```
cat ggg aat tat tca gcg caa gtt ggg gcg tcc cag gcg gca aag ttt     2007
His Gly Asn Tyr Ser Ala Gln Val Gly Ala Ser Gln Ala Ala Lys Phe
    350                 355                 360

aca gta aca ccc aat gct cct tcg ata acc ctc aaa ctt ggt gac tac     2055
Thr Val Thr Pro Asn Ala Pro Ser Ile Thr Leu Lys Leu Gly Asp Tyr
365                 370                 375                 380

gga gaa gtc aca ctg gac tgt gag cca agg agt gga ctg aac act gaa     2103
Gly Glu Val Thr Leu Asp Cys Glu Pro Arg Ser Gly Leu Asn Thr Glu
                385                 390                 395

gcg ttt tac gtc atg acc gtg ggg tca aag tca ttt ctg gtc cat agg     2151
Ala Phe Tyr Val Met Thr Val Gly Ser Lys Ser Phe Leu Val His Arg
            400                 405                 410

gag tgg ttt cat gac ctc gct ctc ccc tgg acg tcc cct tcg agc aca     2199
Glu Trp Phe His Asp Leu Ala Leu Pro Trp Thr Ser Pro Ser Ser Thr
        415                 420                 425

gcg tgg aga aac aga gaa ctc ctc atg gaa ttt gaa gag gcg cac gcc     2247
Ala Trp Arg Asn Arg Glu Leu Leu Met Glu Phe Glu Glu Ala His Ala
    430                 435                 440

aca aaa cag tcc gtt gtt gct ctt ggg tca cag gaa gga ggc ctc cat     2295
Thr Lys Gln Ser Val Val Ala Leu Gly Ser Gln Glu Gly Gly Leu His
445                 450                 455                 460

cag gcg ttg gca gga gcc atc gtg gtg gag tac tca agc tca gtg aag     2343
Gln Ala Leu Ala Gly Ala Ile Val Val Glu Tyr Ser Ser Ser Val Lys
                465                 470                 475

tta aca tca ggc cac ctg aaa tgt agg ctg aaa atg gac aaa ctg gct     2391
Leu Thr Ser Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Ala
            480                 485                 490

ctg aaa ggc aca acc tat ggc atg tgt aca gaa aaa ttc tcg ttc gcg     2439
Leu Lys Gly Thr Thr Tyr Gly Met Cys Thr Glu Lys Phe Ser Phe Ala
        495                 500                 505

aaa aat ccg gcg gac act ggt cac gga aca gtt gtc att gaa ctc tcc     2487
Lys Asn Pro Ala Asp Thr Gly His Gly Thr Val Val Ile Glu Leu Ser
    510                 515                 520

tac tct ggg agt gat ggc ccc tgc aaa att ccg att gct tcc gtt gcg     2535
Tyr Ser Gly Ser Asp Gly Pro Cys Lys Ile Pro Ile Ala Ser Val Ala
525                 530                 535                 540

agc ctc aat gac atg acc ccc gtt ggg cgg ctg gtg aca gtg aac ccc     2583
Ser Leu Asn Asp Met Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro
                545                 550                 555

ttc gtc gcg act tcc agt gcc agc tca aag gtg ctg gtc gag atg gaa     2631
Phe Val Ala Thr Ser Ser Ala Ser Ser Lys Val Leu Val Glu Met Glu
            560                 565                 570

ccc ccc ttc gga gac tcc tac atc gta gtt gga agg gga gac aag cag     2679
Pro Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Asp Lys Gln
        575                 580                 585
```

```
atc aac cac cat tgg cac aaa gct gga agc acg ctg ggc aag gcc ttt     2727
Ile Asn His His Trp His Lys Ala Gly Ser Thr Leu Gly Lys Ala Phe
    590                 595                 600

tca aca act ttg aag gga gct caa aga ctg gca gcg ttg ggc gac aca     2775
Ser Thr Thr Leu Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr
605                 610                 615                 620

gcc tgg gac ttt ggc tct att gga ggg gtc ttc aac tcc ata gga aaa     2823
Ala Trp Asp Phe Gly Ser Ile Gly Gly Val Phe Asn Ser Ile Gly Lys
                625                 630                 635

gcc gtt cac caa gtg ttt ggt ggt gcc ttc aga aca ctc ttt ggg gga     2871
Ala Val His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly
            640                 645                 650

atg tct tgg atc aca caa ggg cta atg ggt gcc cta ctg ctc tgg atg     2919
Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met
        655                 660                 665

ggc gtc aac gca cga gac cga tca att gct ttg gcc ttc tta gcc aca     2967
Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr
    670                 675                 680

ggg ggt gtg ctc gtg ttc tta gcg acc aat gtg cat gct taa             3009
Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala  *
685                 690                 695

ttagtttgag cggccgctcg agcatgcatc tagagggccc tattctatag tgtcacctaa   3069
atgctagagc tcgctgatca gcctcgactg tgccttctag ttgccagcca tctgttgttt   3129
gcccctcccc cgtgccttcc ttgaccctgg aaggtgccac tcccactgtc ctttcctaat   3189
aaaatgagga aattgcatcg cattgtctga gtaggtgtca ttctattctg gggggtgggg   3249
tggggcagga cagcaagggg gaggattggg aagacaatag caggcatgct ggggatgcgg   3309
tgggctctat ggcttctgag gcggaaagaa ccagctgggg ctctaggggg tatccccacg   3369
cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta   3429
cacttgccag cgccctagcg cccgctcctt tcgctttctt cccttccttt ctcgccacgt   3489
tcgccggctt tccccgtcaa gctctaaatc ggggcatccc tttagggttc cgatttagtg   3549
ctttacggca cctcgacccc aaaaaacttg attagggtga tggttcacgt agtgggccat   3609
cgccctgata gacggttttt cgccctttga cgttggagtc cacgttcttt aatagtggac   3669
tcttgttcca aactggaaca acactcaacc ctatctcggt ctattctttt gatttataag   3729
ggattttggg gatttcggcc tattggttaa aaaatgagct gatttaacaa aaatttaacg   3789
cgaattaatt ctgtggaatg tgtgtcagtt agggtgtgga aagtccccag gctccccagg   3849
caggcagaag tatgcaaagc atgcatctca attagtcagc aaccaggtgt ggaaagtccc   3909
caggctcccc agcaggcaga agtatgcaaa gcatgcatct caattagtca gcaaccatag   3969
tcccgcccct aactccgccc atcccgcccc taactccgcc cagttccgcc cattctccgc   4029
cccatggctg actaattttt tttatttatg cagaggccga ggccgcctct gcctctgagc   4089
tattccagaa gtagtgagga ggcttttttg gaggcctagg cttttgcaaa aagctcccgg   4149
gagcttgtat atccattttc ggatctgatc aagagacagg atgaggatcg tttcgcatga   4209
ttgaacaaga tggattgcac gcaggttctc cggccgcttg ggtggagagg ctattcggct   4269
atgactgggc acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc   4329
aggggcgccc ggttcttttt gtcaagaccg acctgtccgg tgccctgaat gaactgcagg   4389
acgaggcagc gcggctatcg tggctggcca cgacgggcgt tccttgcgca gctgtgctcg   4449
acgttgtcac tgaagcggga agggactggc tgctattggg cgaagtgccg gggcaggatc   4509
tcctgtcatc tcaccttgct cctgccgaga aagtatccat catggctgat gcaatgcggc   4569
ggctgcatac gcttgatccg gctacctgcc cattcgacca ccaagcgaaa catcgcatcg   4629
agcgagcacg tactcgggatg gaagccggtc ttgtcgatca ggatgatctg gacgaagagc   4689
atcaggggct cgcgccagcc gaactgttcg ccaggctcaa ggcgcgcatg cccgacggcg   4749
aggatctcgt cgtgacccat ggcgatgcct gcttgccgaa tatcatggtg gaaaatggcc   4809
gcttttctgg attcatcgac tgtggccggc tgggtgtggc ggaccgctat caggacatag   4869
cgttggctac ccgtgatatt gctgaagagc ttggcggcga atgggctgac cgcttcctcg   4929
tgctttacgg tatcgccgct cccgattcgc agcgcatcgc cttctatcgc cttcttgacg   4989
```

agttcttctg agcgggactc tggggttcga aatgaccgac caagcgacgc ccaacctgcc 5049
atcacgagat ttcgattcca ccgccgcctt ctatgaaagg ttgggcttcg gaatcgtttt 5109
ccgggacgcc ggctggatga tcctccagcg cggggatctc atgctggagt tcttcgccca 5169
ccccaacttg tttattgcag cttataatgg ttacaaataa agcaatagca tcacaaattt 5229
cacaaataaa gcatttttt cactgcattc tagttgtggt ttgtccaaac tcatcaatgt 5289
atcttatcat gtctgtatac cgtcgacctc tagctagagc ttggcgtaat catggtcata 5349
gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag 5409
cataaagtgt aaagcctggg gtgcctaatg agtgagctaa ctcacattaa ttgcgttgcg 5469
ctcactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca 5529
acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc tcactgactc 5589
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 5649
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 5709
ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gcccccctga 5769
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 5829
ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 5889
taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc aatgctcacg 5949
ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 6009
ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt 6069
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 6129
tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggac 6189
agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 6249
ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 6309
tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 6369
tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt 6429
cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta 6489
aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 6549
atttcgttca tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg 6609
cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga 6669
tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt 6729
atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt 6789
taatagtttg cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt 6849
tggtatggct tcattcagct ccggttccca acgatcaagg cgagttacat gatcccccat 6909
gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc gttgtcagaa gtaagttggc 6969
cgcagtgtta tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc 7029
cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat 7089
gcggcgaccg agttgctctt gcccggcgtc aatacgggat aataccgcgc cacatagcag 7149
aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt 7209
accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc 7269
ttttactttc accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa 7329
gggaataagg gcgacacgga aatgttgaat actcatactc ttcctttttc aatattattg 7389
aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa 7449
taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg t 7500

<210> 11
<211> 697
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<223> pCDJE 2-7

<400> 11
Met Gly Arg Lys Gln Asn Lys Arg Gly Gly Asn Glu Gly Ser Ile Met
1 5 10 15
Trp Leu Ala Ser Leu Ala Val Val Ile Ala Cys Ala Gly Ala Met Lys
20 25 30
Leu Ser Asn Phe Gln Gly Lys Leu Leu Met Thr Ile Asn Asn Thr Asp
35 40 45

```
Ile Ala Asp Val Ile Val Ile Pro Thr Ser Lys Gly Glu Asn Arg Cys
     50                  55                  60
Trp Val Arg Ala Ile Asp Val Gly Tyr Met Cys Glu Asp Thr Ile Thr
65                  70                  75                  80
Tyr Glu Cys Pro Lys Leu Thr Met Gly Asn Asp Pro Glu Asp Val Asp
                85                  90                  95
Cys Trp Cys Asp Asn Gln Glu Val Tyr Val Gln Tyr Gly Arg Cys Thr
            100                 105                 110
Arg Thr Arg His Ser Lys Arg Ser Arg Arg Ser Val Ser Val Gln Thr
        115                 120                 125
His Gly Glu Ser Ser Leu Val Asn Lys Lys Glu Ala Trp Leu Asp Ser
    130                 135                 140
Thr Lys Ala Thr Arg Tyr Leu Met Lys Thr Glu Asn Trp Ile Ile Arg
145                 150                 155                 160
Asn Pro Gly Tyr Ala Phe Leu Ala Ala Val Leu Gly Trp Met Leu Gly
                165                 170                 175
Ser Asn Asn Gly Gln Arg Val Val Phe Thr Ile Leu Leu Leu Leu Val
            180                 185                 190
Ala Pro Ala Tyr Ser Phe Asn Cys Leu Gly Met Gly Asn Arg Asp Phe
        195                 200                 205
Ile Glu Gly Ala Ser Gly Ala Thr Trp Val Asp Leu Val Leu Glu Gly
    210                 215                 220
Asp Ser Cys Leu Thr Ile Met Ala Asn Asp Lys Pro Thr Leu Asp Val
225                 230                 235                 240
Arg Met Ile Asn Ile Glu Ala Ser Gln Leu Ala Glu Val Arg Ser Tyr
                245                 250                 255
Cys Tyr His Ala Ser Val Thr Asp Ile Ser Thr Val Ala Arg Cys Pro
            260                 265                 270
Thr Thr Gly Glu Ala His Asn Glu Lys Arg Ala Asp Ser Ser Tyr Val
        275                 280                 285
Cys Lys Gln Gly Phe Thr Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu
    290                 295                 300
Phe Gly Lys Gly Ser Ile Asp Thr Cys Ala Lys Phe Ser Cys Thr Ser
305                 310                 315                 320
Lys Ala Ile Gly Arg Thr Ile Gln Pro Glu Asn Ile Lys Tyr Glu Val
                325                 330                 335
Gly Ile Phe Val His Gly Thr Thr Thr Ser Glu Asn His Gly Asn Tyr
            340                 345                 350
Ser Ala Gln Val Gly Ala Ser Gln Ala Ala Lys Phe Thr Val Thr Pro
        355                 360                 365
Asn Ala Pro Ser Ile Thr Leu Lys Leu Gly Asp Tyr Gly Glu Val Thr
    370                 375                 380
Leu Asp Cys Glu Pro Arg Ser Gly Leu Asn Thr Glu Ala Phe Tyr Val
385                 390                 395                 400
Met Thr Val Gly Ser Lys Ser Phe Leu Val His Arg Glu Trp Phe His
                405                 410                 415
Asp Leu Ala Leu Pro Trp Thr Ser Pro Ser Ser Thr Ala Trp Arg Asn
            420                 425                 430
Arg Glu Leu Leu Met Glu Phe Glu Glu Ala His Ala Thr Lys Gln Ser
        435                 440                 445
Val Val Ala Leu Gly Ser Gln Glu Gly Gly Leu His Gln Ala Leu Ala
    450                 455                 460
Gly Ala Ile Val Val Glu Tyr Ser Ser Ser Val Lys Leu Thr Ser Gly
465                 470                 475                 480
His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Ala Leu Lys Gly Thr
                485                 490                 495
Thr Tyr Gly Met Cys Thr Glu Lys Phe Ser Phe Ala Lys Asn Pro Ala
            500                 505                 510
Asp Thr Gly His Gly Thr Val Val Ile Glu Leu Ser Tyr Ser Gly Ser
        515                 520                 525
```

```
Asp Gly Pro Cys Lys Ile Pro Ile Ala Ser Val Ala Ser Leu Asn Asp
    530                 535                 540
Met Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val Ala Thr
545                 550                 555                 560
Ser Ser Ala Ser Ser Lys Val Leu Val Glu Met Glu Pro Pro Phe Gly
                565                 570                 575
Asp Ser Tyr Ile Val Val Gly Arg Gly Asp Lys Gln Ile Asn His His
            580                 585                 590
Trp His Lys Ala Gly Ser Thr Leu Gly Lys Ala Phe Ser Thr Thr Leu
        595                 600                 605
Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr Ala Trp Asp Phe
    610                 615                 620
Gly Ser Ile Gly Gly Val Phe Asn Ser Ile Gly Lys Ala Val His Gln
625                 630                 635                 640
Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly Met Ser Trp Ile
                645                 650                 655
Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met Gly Val Asn Ala
            660                 665                 670
Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr Gly Gly Val Leu
        675                 680                 685
Val Phe Leu Ala Thr Asn Val His Ala
    690                 695
```

<210> 12
<211> 46
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-46
<223> WN 466

<400> 12
cttggtaccc gtctcggcgc cgtgaccctc tcgaacttcc agggca 46

<210> 13
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-43
<223> CWN2444

<400> 13
agaggcactt gcacgtgcgg actttccgcc gcgaaaaaga aaa 43

<210> 14
<211> 24
<212> PRT
<213> Artificial Sequence

<220>

<223> Description of artificial sequence; note =
synthetic construct

<223> JE Signal

<400> 14
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
1 5 10 15
Val Val Ile Ala Cys Ala Gly Ala
20

<210> 15
<211> 5308
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (911)...(2987)

<221> misc_feature
<222> (1)...(5308)
<223> pCBWN

<400> 15
gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggtttttgg cagtacatca atgggcgtgg atagcggttt 660
gactcacggg gatttccaag tctccaccca attgacgtca atgggagttt gttttggcac 720
caaaatcaac gggactttcc aaaatgtcgt aacaactccg ccccattgac gcaaatgggc 780
ggtaggcgtg tacggtgggg ggtctatata agcagagctc tctggctaac tagagaaccc 840
actgcttact ggcttatcga aattaatacg actcactata gggagaccca agcttggtac 900
cgccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg 949
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala
1 5 10

agc ttg gca gtt gtc ata gct tgt gca ggc gcc gtg acc ctc tcg aac 997
Ser Leu Ala Val Val Ile Ala Cys Ala Gly Ala Val Thr Leu Ser Asn
15 20 25

ttc cag ggc aag gtg atg atg acg gta aat gct act gac gtc aca gat 1045
Phe Gln Gly Lys Val Met Met Thr Val Asn Ala Thr Asp Val Thr Asp
30 35 40 45

gtc atc acg att cca aca gct gct gga aag aac cta tgc att gtc aga 1093
Val Ile Thr Ile Pro Thr Ala Ala Gly Lys Asn Leu Cys Ile Val Arg
50 55 60

```
gca atg gat gtg gga tac atg tgc gat gat act atc act tat gaa tgc     1141
Ala Met Asp Val Gly Tyr Met Cys Asp Asp Thr Ile Thr Tyr Glu Cys
                65                  70                  75

cca gtg ctg tcg gct ggt aat gat cca gaa gac atc gac tgt tgg tgc     1189
Pro Val Leu Ser Ala Gly Asn Asp Pro Glu Asp Ile Asp Cys Trp Cys
            80                  85                  90

aca aag tca gca gtc tac gtc agg tat gga aga tgc acc aag aca cgc     1237
Thr Lys Ser Ala Val Tyr Val Arg Tyr Gly Arg Cys Thr Lys Thr Arg
        95                  100                 105

cac tca aga cgc agt cgg agg tca ctg aca gtg cag aca cac gga gaa     1285
His Ser Arg Arg Ser Arg Arg Ser Leu Thr Val Gln Thr His Gly Glu
    110                 115                 120                 125

agc act cta gcg aac aag aag ggg gct tgg atg gac agc acc aag gcc     1333
Ser Thr Leu Ala Asn Lys Lys Gly Ala Trp Met Asp Ser Thr Lys Ala
                130                 135                 140

aca agg tat ttg gta aaa aca gaa tca tgg atc ttg agg aac cct gga     1381
Thr Arg Tyr Leu Val Lys Thr Glu Ser Trp Ile Leu Arg Asn Pro Gly
            145                 150                 155

tat gcc ctg gtg gca gcc gtc att ggt tgg atg ctt ggg agc aac acc     1429
Tyr Ala Leu Val Ala Ala Val Ile Gly Trp Met Leu Gly Ser Asn Thr
        160                 165                 170

atg cag aga gtt gtg ttt gtc gtg cta ttg ctt ttg gtg gcc cca gct     1477
Met Gln Arg Val Val Phe Val Val Leu Leu Leu Leu Val Ala Pro Ala
    175                 180                 185

tac agc ttc aac tgc ctt gga atg agc aac aga gac ttc ttg gaa gga     1525
Tyr Ser Phe Asn Cys Leu Gly Met Ser Asn Arg Asp Phe Leu Glu Gly
190                 195                 200                 205

gtg tct gga gca aca tgg gtg gat ttg gtt ctc gaa ggc gac agc tgc     1573
Val Ser Gly Ala Thr Trp Val Asp Leu Val Leu Glu Gly Asp Ser Cys
                210                 215                 220

gtg act atc atg tct aag gac aag cct acc atc gat gtg aag atg atg     1621
Val Thr Ile Met Ser Lys Asp Lys Pro Thr Ile Asp Val Lys Met Met
            225                 230                 235

aat atg gag gcg gcc aac ctg gca gag gtc cgc agt tat tgc tat ttg     1669
Asn Met Glu Ala Ala Asn Leu Ala Glu Val Arg Ser Tyr Cys Tyr Leu
        240                 245                 250

gct acc gtc agc gat ctc tcc acc aaa gct gcg tgc ccg acc atg gga     1717
Ala Thr Val Ser Asp Leu Ser Thr Lys Ala Ala Cys Pro Thr Met Gly
    255                 260                 265

gaa gct cac aat gac aaa cgt gct gac cca gct ttt gtg tgc aga caa     1765
Glu Ala His Asn Asp Lys Arg Ala Asp Pro Ala Phe Val Cys Arg Gln
270                 275                 280                 285

gga gtg gtg gac agg ggc tgg ggc aac ggc tgc gga cta ttt ggc aaa     1813
Gly Val Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys
                290                 295                 300
```

```
gga agc att gac aca tgc gcc aaa ttt gcc tgc tct acc aag gca ata      1861
Gly Ser Ile Asp Thr Cys Ala Lys Phe Ala Cys Ser Thr Lys Ala Ile
            305                 310                 315

gga aga acc atc ttg aaa gag aat atc aag tac gaa gtg gcc att ttt      1909
Gly Arg Thr Ile Leu Lys Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe
        320                 325                 330

gtc cat gga cca act act gtg gag tcg cac gga aac tac tcc aca cag      1957
Val His Gly Pro Thr Thr Val Glu Ser His Gly Asn Tyr Ser Thr Gln
    335                 340                 345

gtt gga gcc act cag gca ggg aga ttc agc atc act cct gcg gcg cct      2005
Val Gly Ala Thr Gln Ala Gly Arg Phe Ser Ile Thr Pro Ala Ala Pro
350                 355                 360                 365

tca tac aca cta aag ctt gga gaa tat gga gag gtg aca gtg gac tgt      2053
Ser Tyr Thr Leu Lys Leu Gly Glu Tyr Gly Glu Val Thr Val Asp Cys
                370                 375                 380

gaa cca cgg tca ggg att gac acc aat gca tac tac gtg atg act gtt      2101
Glu Pro Arg Ser Gly Ile Asp Thr Asn Ala Tyr Tyr Val Met Thr Val
            385                 390                 395

gga aca aag acg ttc ttg gtc cat cgt gag tgg ttc atg gac ctc aac      2149
Gly Thr Lys Thr Phe Leu Val His Arg Glu Trp Phe Met Asp Leu Asn
        400                 405                 410

ctc cct tgg agc agt gct gga agt act gtg tgg agg aac aga gag acg      2197
Leu Pro Trp Ser Ser Ala Gly Ser Thr Val Trp Arg Asn Arg Glu Thr
    415                 420                 425

tta atg gag ttt gag gaa cca cac gcc acg aag cag tct gtg ata gca      2245
Leu Met Glu Phe Glu Glu Pro His Ala Thr Lys Gln Ser Val Ile Ala
430                 435                 440                 445

ttg ggc tca caa gag gga gct ctg cat caa gct ttg gct gga gcc att      2293
Leu Gly Ser Gln Glu Gly Ala Leu His Gln Ala Leu Ala Gly Ala Ile
                450                 455                 460

cct gtg gaa ttt tca agc aac act gtc aag ttg acg tcg ggt cat ttg      2341
Pro Val Glu Phe Ser Ser Asn Thr Val Lys Leu Thr Ser Gly His Leu
            465                 470                 475

aag tgt aga gtg aag atg gaa aaa ttg cag ttg aag gga aca acc tat      2389
Lys Cys Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr
        480                 485                 490

ggc gtc tgt tca aag gct ttc aag ttt ctt ggg act ccc gcg gac aca      2437
Gly Val Cys Ser Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr
    495                 500                 505

ggt cac ggc act gtg gtg ttg gaa ttg cag tac act ggc acg gat gga      2485
Gly His Gly Thr Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly
510                 515                 520                 525

cct tgc aaa gtt cct atc tcg tca gtg gct tca ttg aac gac cta acg      2533
Pro Cys Lys Val Pro Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr
                530                 535                 540
```

```
cca gtg ggc aga ttg gtc act gtc aac cct ttt gtt tca gtg gcc acg      2581
Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val Ser Val Ala Thr
            545                 550                 555

gcc aac gct aag gtc ctg att gaa ttg gaa cca ccc ttt gga gac tca      2629
Ala Asn Ala Lys Val Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser
        560                 565                 570

tac ata gtg gtg ggc aga gga gaa caa cag atc aat cac cat tgg cac      2677
Tyr Ile Val Val Gly Arg Gly Glu Gln Gln Ile Asn His His Trp His
    575                 580                 585

aag tct gga agc agc att ggc aaa gcc ttt aca acc acc ctc aaa gga      2725
Lys Ser Gly Ser Ser Ile Gly Lys Ala Phe Thr Thr Thr Leu Lys Gly
590                 595                 600                 605

gcg cag aga cta gcc gct cta gga gac aca gct tgg gac ttt gga tca      2773
Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser
                610                 615                 620

gtt gga ggg gtg ttc acc tca gtt ggg aag gct gtc cat caa gtg ttc      2821
Val Gly Gly Val Phe Thr Ser Val Gly Lys Ala Val His Gln Val Phe
            625                 630                 635

gga gga gca ttc cgc tca ctg ttc gga ggc atg tcc tgg ata acg caa      2869
Gly Gly Ala Phe Arg Ser Leu Phe Gly Gly Met Ser Trp Ile Thr Gln
        640                 645                 650

gga ttg ctg ggg gct ctc ctg ttg tgg atg ggc atc aat gct cgt gat      2917
Gly Leu Leu Gly Ala Leu Leu Leu Trp Met Gly Ile Asn Ala Arg Asp
    655                 660                 665

agg tcc ata gct ctc acg ttt ctc gca gtt gga gga gtt ctg ctc ttc      2965
Arg Ser Ile Ala Leu Thr Phe Leu Ala Val Gly Gly Val Leu Leu Phe
670                 675                 680                 685

ctc tcc gtg aac gtg cac gcc t gaaggcggcc gctcgagcat gcatctagag      3017
Leu Ser Val Asn Val His Ala
                690

ggccctattc tatagtgtca cctaaatgct agagctcgct gatcagcctc gactgtgcct   3077
tctagttgcc agccatctgt tgtttgcccc tccccccgtgc cttccttgac cctggaaggt   3137
gccactccca ctgtcctttc ctaataaaat gaggaaattg catcgcattg tctgagtagg   3197
tgtcattcta ttctgggggg tggggtgggg caggacagca agggggagga ttgggaagac   3257
aatagcaggc atgctgggga tgcggtgggc tctatggctt ctgaggcgga aagaaccagc   3317
tgcattaatg aatcggccaa cgcgcggggg gaggcggttt gcgtattggg cgctcttccg   3377
cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc   3437
actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt   3497
gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc   3557
ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa   3617
acccgacagg actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc   3677
ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg   3737
cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc   3797
tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc   3857
gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca   3917
ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact   3977
acggctacac tagaagaaca gtatttggta tctgcgctct gctgaagcca gttaccttcg   4037
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt   4097
ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct   4157
tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga   4217
gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa   4277
```

tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac 4337
ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga 4397
taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc 4457
cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca 4517
gaagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta 4577
gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg 4637
tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc 4697
gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt cctccgatcg 4757
ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt 4817
ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac tcaaccaagt 4877
cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca atacgggata 4937
ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc 4997
gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac 5057
ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa 5117
ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct 5177
tcctttttca atattattga agcatttatc agggttattg tctcatgagc ggatacatat 5237
ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc 5297
cacctgacgt c 5308

<210> 16
<211> 692
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<223> pCEMN

<400> 16
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
1 5 10 15
Val Val Ile Ala Cys Ala Gly Ala Val Thr Leu Ser Asn Phe Gln Gly
20 25 30
Lys Val Met Met Thr Val Asn Ala Thr Asp Val Thr Asp Val Ile Thr
35 40 45
Ile Pro Thr Ala Ala Gly Lys Asn Leu Cys Ile Val Arg Ala Met Asp
50 55 60
Val Gly Tyr Met Cys Asp Asp Thr Ile Thr Tyr Glu Cys Pro Val Leu
65 70 75 80
Ser Ala Gly Asn Asp Pro Glu Asp Ile Asp Cys Trp Cys Thr Lys Ser
85 90 95
Ala Val Tyr Val Arg Tyr Gly Arg Cys Thr Lys Thr Arg His Ser Arg
100 105 110
Arg Ser Arg Arg Ser Leu Thr Val Gln Thr His Gly Glu Ser Thr Leu
115 120 125
Ala Asn Lys Lys Gly Ala Trp Met Asp Ser Thr Lys Ala Thr Arg Tyr
130 135 140
Leu Val Lys Thr Glu Ser Trp Ile Leu Arg Asn Pro Gly Tyr Ala Leu
145 150 155 160
Val Ala Ala Val Ile Gly Trp Met Leu Gly Ser Asn Thr Met Gln Arg
165 170 175
Val Val Phe Val Val Leu Leu Leu Leu Val Ala Pro Ala Tyr Ser Phe
180 185 190
Asn Cys Leu Gly Met Ser Asn Arg Asp Phe Leu Glu Gly Val Ser Gly
195 200 205
Ala Thr Trp Val Asp Leu Val Leu Glu Gly Asp Ser Cys Val Thr Ile
210 215 220
Met Ser Lys Asp Lys Pro Thr Ile Asp Val Lys Met Met Asn Met Glu
225 230 235 240

```
Ala Ala Asn Leu Ala Glu Val Arg Ser Tyr Cys Tyr Leu Ala Thr Val
                245                 250                 255
Ser Asp Leu Ser Thr Lys Ala Ala Cys Pro Thr Met Gly Glu Ala His
            260                 265                 270
Asn Asp Lys Arg Ala Asp Pro Ala Phe Val Cys Arg Gln Gly Val Val
        275                 280                 285
Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile
    290                 295                 300
Asp Thr Cys Ala Lys Phe Ala Cys Ser Thr Lys Ala Ile Gly Arg Thr
305                 310                 315                 320
Ile Leu Lys Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe Val His Gly
                325                 330                 335
Pro Thr Thr Val Glu Ser His Gly Asn Tyr Ser Thr Gln Val Gly Ala
            340                 345                 350
Thr Gln Ala Gly Arg Phe Ser Ile Thr Pro Ala Ala Pro Ser Tyr Thr
        355                 360                 365
Leu Lys Leu Gly Glu Tyr Gly Glu Val Thr Val Asp Cys Glu Pro Arg
    370                 375                 380
Ser Gly Ile Asp Thr Asn Ala Tyr Tyr Val Met Thr Val Gly Thr Lys
385                 390                 395                 400
Thr Phe Leu Val His Arg Glu Trp Phe Met Asp Leu Asn Leu Pro Trp
                405                 410                 415
Ser Ser Ala Gly Ser Thr Val Trp Arg Asn Arg Glu Thr Leu Met Glu
            420                 425                 430
Phe Glu Glu Pro His Ala Thr Lys Gln Ser Val Ile Ala Leu Gly Ser
        435                 440                 445
Gln Glu Gly Ala Leu His Gln Ala Leu Ala Gly Ala Ile Pro Val Glu
    450                 455                 460
Phe Ser Ser Asn Thr Val Lys Leu Thr Ser Gly His Leu Lys Cys Arg
465                 470                 475                 480
Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val Cys
                485                 490                 495
Ser Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr Gly His Gly
            500                 505                 510
Thr Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly Pro Cys Lys
        515                 520                 525
Val Pro Ile Ser Ser Val Ala Ser Leu Asn Asp Leu Thr Pro Val Gly
    530                 535                 540
Arg Leu Val Thr Val Asn Pro Phe Val Ser Val Ala Thr Ala Asn Ala
545                 550                 555                 560
Lys Val Leu Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val
                565                 570                 575
Val Gly Arg Gly Glu Gln Gln Ile Asn His His Trp His Lys Ser Gly
            580                 585                 590
Ser Ser Ile Gly Lys Ala Phe Thr Thr Thr Leu Lys Gly Ala Gln Arg
        595                 600                 605
Leu Ala Ala Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Val Gly Gly
    610                 615                 620
Val Phe Thr Ser Val Gly Lys Ala Val His Gln Val Phe Gly Gly Ala
625                 630                 635                 640
Phe Arg Ser Leu Phe Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Leu
                645                 650                 655
Gly Ala Leu Leu Leu Trp Met Gly Ile Asn Ala Arg Asp Arg Ser Ile
            660                 665                 670
Ala Leu Thr Phe Leu Ala Val Gly Gly Val Leu Leu Phe Leu Ser Val
        675                 680                 685
Asn Val His Ala
    690
```

<210> 17
<211> 5334

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (916)...(3007)

<221> misc_feature
<222> (1)...(5334)
<223> pCBJE 1-14

<400> 17
gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc 900
tctagagccg ccgcc atg ggc aga aag caa aac aaa aga gga gga aat gaa 951
Met Gly Arg Lys Gln Asn Lys Arg Gly Gly Asn Glu
1 5 10

ggc tca atc atg tgg ctc gcg agc ttg gca gtt gtc ata gct tgt gcg 999
Gly Ser Ile Met Trp Leu Ala Ser Leu Ala Val Val Ile Ala Cys Ala
15 20 25

gga gcc atg aag ttg tcg aat ttc cag ggg aag ctt ttg atg acc atc 1047
Gly Ala Met Lys Leu Ser Asn Phe Gln Gly Lys Leu Leu Met Thr Ile
30 35 40

aac aac acg gac att gca gac gtt atc gtg att ccc acc tca aaa gga 1095
Asn Asn Thr Asp Ile Ala Asp Val Ile Val Ile Pro Thr Ser Lys Gly
45 50 55 60

gag aac aga tgc tgg gtc cgg gca atc gac gtc ggc tac atg tgt gag 1143
Glu Asn Arg Cys Trp Val Arg Ala Ile Asp Val Gly Tyr Met Cys Glu
65 70 75

gac act atc acg tac gaa tgt cct aag ctt acc atg ggc aat gat cca 1191
Asp Thr Ile Thr Tyr Glu Cys Pro Lys Leu Thr Met Gly Asn Asp Pro
80 85 90

gag gat gtg gat tgc tgg tgt gac aac caa gaa gtc tac gtc caa tat 1239
Glu Asp Val Asp Cys Trp Cys Asp Asn Gln Glu Val Tyr Val Gln Tyr
95 100 105

```
gga cgg tgc acg cgg acc agg cat tcc aag cga agc agg aga tcc gtg     1287
Gly Arg Cys Thr Arg Thr Arg His Ser Lys Arg Ser Arg Arg Ser Val
    110                 115                 120

tcg gtc caa aca cat ggg gag agt tca cta gtg aat aaa aaa gag gct     1335
Ser Val Gln Thr His Gly Glu Ser Ser Leu Val Asn Lys Lys Glu Ala
125                 130                 135                 140

tgg ctg gat tca acg aaa gcc aca cga tat ctc atg aaa act gag aac     1383
Trp Leu Asp Ser Thr Lys Ala Thr Arg Tyr Leu Met Lys Thr Glu Asn
                145                 150                 155

tgg atc ata agg aat cct ggc tat gct ttc ctg gcg gcg gta ctt ggc     1431
Trp Ile Ile Arg Asn Pro Gly Tyr Ala Phe Leu Ala Ala Val Leu Gly
            160                 165                 170

tgg atg ctt ggc agt aac aac ggt caa cgc gtg gta ttt acc atc ctc     1479
Trp Met Leu Gly Ser Asn Asn Gly Gln Arg Val Val Phe Thr Ile Leu
        175                 180                 185

ctg ctg ttg gtc gct ccg gct tac agt ttt aat tgt ctg gga atg ggc     1527
Leu Leu Leu Val Ala Pro Ala Tyr Ser Phe Asn Cys Leu Gly Met Gly
    190                 195                 200

aat cgt gac ttc ata gaa gga gcc agt gga gcc act tgg gtg gac ttg     1575
Asn Arg Asp Phe Ile Glu Gly Ala Ser Gly Ala Thr Trp Val Asp Leu
205                 210                 215                 220

gtg ctg gaa gga gat agc tgc ttg aca atc atg gca aac gac aaa cca     1623
Val Leu Glu Gly Asp Ser Cys Leu Thr Ile Met Ala Asn Asp Lys Pro
                225                 230                 235

aca ttg gac gtc cgc atg att aac atc gaa gct agc caa ctt gct gag     1671
Thr Leu Asp Val Arg Met Ile Asn Ile Glu Ala Ser Gln Leu Ala Glu
            240                 245                 250

gtc aga agt tac tgc tat cat gct tca gtc act gac atc tcg acg gtg     1719
Val Arg Ser Tyr Cys Tyr His Ala Ser Val Thr Asp Ile Ser Thr Val
        255                 260                 265

gct cgg tgc ccc acg act gga gaa gcc cac aac gag aag cga gct gat     1767
Ala Arg Cys Pro Thr Thr Gly Glu Ala His Asn Glu Lys Arg Ala Asp
    270                 275                 280

agt agc tat gtg tgc aaa caa ggc ttc act gac cgt ggg tgg ggc aac     1815
Ser Ser Tyr Val Cys Lys Gln Gly Phe Thr Asp Arg Gly Trp Gly Asn
285                 290                 295                 300

gga tgt gga ctt ttc ggg aag gga agc att gac aca tgt gca aaa ttc     1863
Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Asp Thr Cys Ala Lys Phe
                305                 310                 315

tcc tgc acc agt aaa gcg att ggg aga aca atc cag cca gaa aac atc     1911
Ser Cys Thr Ser Lys Ala Ile Gly Arg Thr Ile Gln Pro Glu Asn Ile
            320                 325                 330

aaa tac gaa gtt ggc att ttt gtg cat gga acc acc act tcg gaa aac     1959
Lys Tyr Glu Val Gly Ile Phe Val His Gly Thr Thr Thr Ser Glu Asn
        335                 340                 345
```

```
cat ggg aat tat tca gcg caa gtt ggg gcg tcc cag gcg gca aag ttt     2007
His Gly Asn Tyr Ser Ala Gln Val Gly Ala Ser Gln Ala Ala Lys Phe
    350                 355                 360

aca gta aca ccc aat gct cct tcg ata acc ctc aaa ctt ggt gac tac     2055
Thr Val Thr Pro Asn Ala Pro Ser Ile Thr Leu Lys Leu Gly Asp Tyr
365                 370                 375                 380

gga gaa gtc aca ctg gac tgt gag cca agg agt gga ctg aac act gaa     2103
Gly Glu Val Thr Leu Asp Cys Glu Pro Arg Ser Gly Leu Asn Thr Glu
                385                 390                 395

gcg ttt tac gtc atg acc gtg ggg tca aag tca ttt ctg gtc cat agg     2151
Ala Phe Tyr Val Met Thr Val Gly Ser Lys Ser Phe Leu Val His Arg
            400                 405                 410

gag tgg ttt cat gac ctc gct ctc ccc tgg acg tcc cct tcg agc aca     2199
Glu Trp Phe His Asp Leu Ala Leu Pro Trp Thr Ser Pro Ser Ser Thr
        415                 420                 425

gcg tgg aga aac aga gaa ctc ctc atg gaa ttt gaa gag gcg cac gcc     2247
Ala Trp Arg Asn Arg Glu Leu Leu Met Glu Phe Glu Glu Ala His Ala
    430                 435                 440

aca aaa cag tcc gtt gtt gct ctt ggg tca cag gaa gga ggc ctc cat     2295
Thr Lys Gln Ser Val Val Ala Leu Gly Ser Gln Glu Gly Gly Leu His
445                 450                 455                 460

cag gcg ttg gca gga gcc atc gtg gtg gag tac tca agc tca gtg aag     2343
Gln Ala Leu Ala Gly Ala Ile Val Val Glu Tyr Ser Ser Ser Val Lys
                465                 470                 475

tta aca tca ggc cac ctg aaa tgt agg ctg aaa atg gac aaa ctg gct     2391
Leu Thr Ser Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Ala
            480                 485                 490

ctg aaa ggc aca acc tat ggc atg tgt aca gaa aaa ttc tcg ttc gcg     2439
Leu Lys Gly Thr Thr Tyr Gly Met Cys Thr Glu Lys Phe Ser Phe Ala
        495                 500                 505

aaa aat ccg gcg gac act ggt cac gga aca gtt gtc att gaa ctc tcc     2487
Lys Asn Pro Ala Asp Thr Gly His Gly Thr Val Val Ile Glu Leu Ser
    510                 515                 520

tac tct ggg agt gat ggc ccc tgc aaa att ccg att gct tcc gtt gcg     2535
Tyr Ser Gly Ser Asp Gly Pro Cys Lys Ile Pro Ile Ala Ser Val Ala
525                 530                 535                 540

agc ctc aat gac atg acc ccc gtt ggg cgg ctg gtg aca gtg aac ccc     2583
Ser Leu Asn Asp Met Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro
                545                 550                 555

ttc gtc gcg act tcc agt gcc agc tca aag gtg ctg gtc gag atg gaa     2631
Phe Val Ala Thr Ser Ser Ala Ser Ser Lys Val Leu Val Glu Met Glu
            560                 565                 570

ccc ccc ttc gga gac tcc tac atc gta gtt gga agg gga gac aag cag     2679
Pro Pro Phe Gly Asp Ser Tyr Ile Val Val Gly Arg Gly Asp Lys Gln
        575                 580                 585
```

```
atc aac cac cat tgg cac aaa gct gga agc acg ctg ggc aag gcc ttt     2727
Ile Asn His His Trp His Lys Ala Gly Ser Thr Leu Gly Lys Ala Phe
    590                 595                 600

tca aca act ttg aag gga gct caa aga ctg gca gcg ttg ggc gac aca     2775
Ser Thr Thr Leu Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr
605                 610                 615                 620

gcc tgg gac ttt ggc tct att gga ggg gtc ttc aac tcc ata gga aaa     2823
Ala Trp Asp Phe Gly Ser Ile Gly Gly Val Phe Asn Ser Ile Gly Lys
                625                 630                 635

gcc gtt cac caa gtg ttt ggt ggt gcc ttc aga aca ctc ttt ggg gga     2871
Ala Val His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly
            640                 645                 650

atg tct tgg atc aca caa ggg cta atg ggt gcc cta ctg ctc tgg atg     2919
Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met
        655                 660                 665

ggc gtc aac gca cga gac cga tca att gct ttg gcc ttc tta gcc aca     2967
Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr
    670                 675                 680

ggg ggt gtg ctc gtg ttc tta gcg acc aat gtg cat gct t aattagtttg    3017
Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala
685                 690                 695

agcggccgct cgagcatgca tctagagggc cctattctat agtgtcacct aaatgctaga   3077
gctcgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt ttgcccctcc   3137
cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tccttttcct ataaaatgag   3197
gaaattgcat cgcattgtct gagtaggtgt cattctattc tgggggggtgg ggtggggcag   3257
gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct   3317
atggcttctg aggcggaaag aaccagctgc attaatgaat cggccaacgc gcggggagag   3377
gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg cgctcggtcg   3437
ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta tccacagaat   3497
caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta   3557
aaaaggccgc gttgctggcg tttttccata ggctccgccc cctgacgag catcacaaaa   3617
atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc   3677
cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt   3737
ccgcctttct cccttcgggg agcgtggcgc tttctcatag ctcacgctgt aggtatctca   3797
gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaacccccc gttcagcccg   3857
accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat   3917
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta   3977
cagagttctt gaagtggtgg cctaactacg gctacactag aagaacagta tttggtatct   4037
gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac   4097
aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa   4157
aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa   4217
actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt   4277
taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact tggtctgaca   4337
gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca   4397
tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta ccatctggcc   4457
ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta tcagcaataa   4517
accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc   4577
agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca   4637
acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat   4697
tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag   4757
cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac   4817
tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt   4877
ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg cgaccgagtt   4937
gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc   4997
```

tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat 5057
ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt actttcacca 5117
gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaagggа ataagggcga 5177
cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg 5237
gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg 5297
ttccgcgcac atttccccga aaagtgccac ctgacgt 5334

<210> 18
<211> 697
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<223> pCBJE 1-14

<400> 18
Met Gly Arg Lys Gln Asn Lys Arg Gly Gly Asn Glu Gly Ser Ile Met
1 5 10 15
Trp Leu Ala Ser Leu Ala Val Val Ile Ala Cys Ala Gly Ala Met Lys
20 25 30
Leu Ser Asn Phe Gln Gly Lys Leu Leu Met Thr Ile Asn Asn Thr Asp
35 40 45
Ile Ala Asp Val Ile Val Ile Pro Thr Ser Lys Gly Glu Asn Arg Cys
50 55 60
Trp Val Arg Ala Ile Asp Val Gly Tyr Met Cys Glu Asp Thr Ile Thr
65 70 75 80
Tyr Glu Cys Pro Lys Leu Thr Met Gly Asn Asp Pro Glu Asp Val Asp
85 90 95
Cys Trp Cys Asp Asn Gln Glu Val Tyr Val Gln Tyr Gly Arg Cys Thr
100 105 110
Arg Thr Arg His Ser Lys Arg Ser Arg Arg Ser Val Ser Val Gln Thr
115 120 125
His Gly Glu Ser Ser Leu Val Asn Lys Lys Glu Ala Trp Leu Asp Ser
130 135 140
Thr Lys Ala Thr Arg Tyr Leu Met Lys Thr Glu Asn Trp Ile Ile Arg
145 150 155 160
Asn Pro Gly Tyr Ala Phe Leu Ala Ala Val Leu Gly Trp Met Leu Gly
165 170 175
Ser Asn Asn Gly Gln Arg Val Val Phe Thr Ile Leu Leu Leu Leu Val
180 185 190
Ala Pro Ala Tyr Ser Phe Asn Cys Leu Gly Met Gly Asn Arg Asp Phe
195 200 205
Ile Glu Gly Ala Ser Gly Ala Thr Trp Val Asp Leu Val Leu Glu Gly
210 215 220
Asp Ser Cys Leu Thr Ile Met Ala Asn Asp Lys Pro Thr Leu Asp Val
225 230 235 240
Arg Met Ile Asn Ile Glu Ala Ser Gln Leu Ala Glu Val Arg Ser Tyr
245 250 255
Cys Tyr His Ala Ser Val Thr Asp Ile Ser Thr Val Ala Arg Cys Pro
260 265 270
Thr Thr Gly Glu Ala His Asn Glu Lys Arg Ala Asp Ser Ser Tyr Val
275 280 285
Cys Lys Gln Gly Phe Thr Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu
290 295 300
Phe Gly Lys Gly Ser Ile Asp Thr Cys Ala Lys Phe Ser Cys Thr Ser
305 310 315 320
Lys Ala Ile Gly Arg Thr Ile Gln Pro Glu Asn Ile Lys Tyr Glu Val
325 330 335

```
Gly Ile Phe Val His Gly Thr Thr Thr Ser Glu Asn His Gly Asn Tyr
            340                 345                 350
Ser Ala Gln Val Gly Ala Ser Gln Ala Ala Lys Phe Thr Val Thr Pro
        355                 360                 365
Asn Ala Pro Ser Ile Thr Leu Lys Leu Gly Asp Tyr Gly Glu Val Thr
    370                 375                 380
Leu Asp Cys Glu Pro Arg Ser Gly Leu Asn Thr Glu Ala Phe Tyr Val
385                 390                 395                 400
Met Thr Val Gly Ser Lys Ser Phe Leu Val His Arg Glu Trp Phe His
                405                 410                 415
Asp Leu Ala Leu Pro Trp Thr Ser Pro Ser Ser Thr Ala Trp Arg Asn
            420                 425                 430
Arg Glu Leu Leu Met Glu Phe Glu Glu Ala His Ala Thr Lys Gln Ser
        435                 440                 445
Val Val Ala Leu Gly Ser Gln Glu Gly Gly Leu His Gln Ala Leu Ala
    450                 455                 460
Gly Ala Ile Val Val Glu Tyr Ser Ser Ser Val Lys Leu Thr Ser Gly
465                 470                 475                 480
His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Ala Leu Lys Gly Thr
                485                 490                 495
Thr Tyr Gly Met Cys Thr Glu Lys Phe Ser Phe Ala Lys Asn Pro Ala
            500                 505                 510
Asp Thr Gly His Gly Thr Val Val Ile Glu Leu Ser Tyr Ser Gly Ser
        515                 520                 525
Asp Gly Pro Cys Lys Ile Pro Ile Ala Ser Val Ala Ser Leu Asn Asp
    530                 535                 540
Met Thr Pro Val Gly Arg Leu Val Thr Val Asn Pro Phe Val Ala Thr
545                 550                 555                 560
Ser Ser Ala Ser Ser Lys Val Leu Val Glu Met Glu Pro Pro Phe Gly
                565                 570                 575
Asp Ser Tyr Ile Val Val Gly Arg Gly Asp Lys Gln Ile Asn His His
            580                 585                 590
Trp His Lys Ala Gly Ser Thr Leu Gly Lys Ala Phe Ser Thr Thr Leu
        595                 600                 605
Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr Ala Trp Asp Phe
    610                 615                 620
Gly Ser Ile Gly Gly Val Phe Asn Ser Ile Gly Lys Ala Val His Gln
625                 630                 635                 640
Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly Met Ser Trp Ile
                645                 650                 655
Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met Gly Val Asn Ala
            660                 665                 670
Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr Gly Gly Val Leu
        675                 680                 685
Val Phe Leu Ala Thr Asn Val His Ala
    690                 695
```

<210> 19
<211> 5283
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (910)...(2965)

<400> 19
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60

382

```
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg    120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc    180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt    240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata    300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc    360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc    420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt    480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt    540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca    600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg    660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc    720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg    780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacccа    840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc    900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc    951
          Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser
           1               5                  10

ttg gca gtt gtc ata gct ggt aca agc gct acc acc atc cac cgg gac      999
Leu Ala Val Val Ile Ala Gly Thr Ser Ala Thr Thr Ile His Arg Asp
 15                  20                  25                  30

agg gaa gga tac atg gtt atg cgg gcc agt gga agg gac gct gca agc     1047
Arg Glu Gly Tyr Met Val Met Arg Ala Ser Gly Arg Asp Ala Ala Ser
                 35                  40                  45

cag gtc agg gta caa aac gga acg tgc gtc atc ctg gca aca gac atg     1095
Gln Val Arg Val Gln Asn Gly Thr Cys Val Ile Leu Ala Thr Asp Met
             50                  55                  60

gga gag tgg tgt gaa gat tca atc acc tac tct tgc gtc acg att gac     1143
Gly Glu Trp Cys Glu Asp Ser Ile Thr Tyr Ser Cys Val Thr Ile Asp
         65                  70                  75

cag gag gaa gaa ccc gtt gac gtg gac tgc ttc tgc cga ggt gtt gat     1191
Gln Glu Glu Glu Pro Val Asp Val Asp Cys Phe Cys Arg Gly Val Asp
     80                  85                  90

agg gtt aag tta gag tat gga cgc tgt gga agg caa gct gga tct agg     1239
Arg Val Lys Leu Glu Tyr Gly Arg Cys Gly Arg Gln Ala Gly Ser Arg
 95                  100                 105                 110

ggg aaa agg tct gtg gtc att cca aca cat gca caa aaa gac atg gtc     1287
Gly Lys Arg Ser Val Val Ile Pro Thr His Ala Gln Lys Asp Met Val
                115                 120                 125

ggg cga ggt cat gca tgg ctt aaa ggt gac aat att cga gat cat gtc     1335
Gly Arg Gly His Ala Trp Leu Lys Gly Asp Asn Ile Arg Asp His Val
            130                 135                 140

acc cga gtc gag ggc tgg atg tgg aag aac aag ctt cta act gcc gcc     1383
Thr Arg Val Glu Gly Trp Met Trp Lys Asn Lys Leu Leu Thr Ala Ala
        145                 150                 155

att gtg gcc ttg gct tgg ctc atg gtt gat agt tgg atg gcc aga gtg     1431
Ile Val Ala Leu Ala Trp Leu Met Val Asp Ser Trp Met Ala Arg Val
    160                 165                 170

act gtc atc ctc ttg gcg ttg agt cta ggg cca gtg tac gcc acg agg     1479
Thr Val Ile Leu Leu Ala Leu Ser Leu Gly Pro Val Tyr Ala Thr Arg
175                 180                 185                 190
```

```
tgc acg cat ctt gag aac aga gat ttt gtg aca gga act caa ggg acc      1527
Cys Thr His Leu Glu Asn Arg Asp Phe Val Thr Gly Thr Gln Gly Thr
                195                 200                 205

acc aga gtg tcc cta gtt ttg gaa ctt gga ggc tgc gtg acc atc aca      1575
Thr Arg Val Ser Leu Val Leu Glu Leu Gly Gly Cys Val Thr Ile Thr
            210                 215                 220

gct gag ggc aag cca tcc att gat gta tgg ctc gaa gac att ttt cag      1623
Ala Glu Gly Lys Pro Ser Ile Asp Val Trp Leu Glu Asp Ile Phe Gln
        225                 230                 235

gaa agc ccg gct gaa acc aga gaa tac tgc ctg cac gcc aaa ttg acc      1671
Glu Ser Pro Ala Glu Thr Arg Glu Tyr Cys Leu His Ala Lys Leu Thr
    240                 245                 250

aac aca aaa gtg gag gct cgc tgt cca acc act gga ccg gcg aca ctt      1719
Asn Thr Lys Val Glu Ala Arg Cys Pro Thr Thr Gly Pro Ala Thr Leu
255                 260                 265                 270

ccg gag gag cat cag gct aat atg gtg tgc aag aga gac caa agc gac      1767
Pro Glu Glu His Gln Ala Asn Met Val Cys Lys Arg Asp Gln Ser Asp
                275                 280                 285

cgt gga tgg gga aac cac tgc ggg ttt ttt ggg aag ggc agt ata gtg      1815
Arg Gly Trp Gly Asn His Cys Gly Phe Phe Gly Lys Gly Ser Ile Val
            290                 295                 300

gct tgt gca aag ttt gaa tgc gag gaa gca aaa aaa gct gtg ggc cac      1863
Ala Cys Ala Lys Phe Glu Cys Glu Glu Ala Lys Lys Ala Val Gly His
        305                 310                 315

gtc tat gac tcc aca aag atc acg tat gtt gtc aag gtt gag ccc cac      1911
Val Tyr Asp Ser Thr Lys Ile Thr Tyr Val Val Lys Val Glu Pro His
    320                 325                 330

aca ggg gat tac ttg gct gca aat gag acc aat tca aac agg aaa tca      1959
Thr Gly Asp Tyr Leu Ala Ala Asn Glu Thr Asn Ser Asn Arg Lys Ser
335                 340                 345                 350

gca cag ttt acg gtg gca tcc gag aaa gtg atc ctg cgg ctc ggc gac      2007
Ala Gln Phe Thr Val Ala Ser Glu Lys Val Ile Leu Arg Leu Gly Asp
                355                 360                 365

tat gga gat gtg tcg ctg acg tgt aaa gtg gca agt ggg att gat gtc      2055
Tyr Gly Asp Val Ser Leu Thr Cys Lys Val Ala Ser Gly Ile Asp Val
            370                 375                 380

gcc caa act gtg gtg atg tca ctc gac agc agc aag gac cac ctg cct      2103
Ala Gln Thr Val Val Met Ser Leu Asp Ser Ser Lys Asp His Leu Pro
        385                 390                 395

tct gca tgg caa gtg cac cgt gac tgg ttt gag gac ttg gcg ctg ccc      2151
Ser Ala Trp Gln Val His Arg Asp Trp Phe Glu Asp Leu Ala Leu Pro
    400                 405                 410

tgg aaa cac aag gac aac caa gat tgg aac agt gtg gag aaa ctt gtg      2199
Trp Lys His Lys Asp Asn Gln Asp Trp Asn Ser Val Glu Lys Leu Val
415                 420                 425                 430
```

```
gaa ttt gga cca cca cat gct gtg aaa atg gat gtt ttc aat ctg ggg      2247
Glu Phe Gly Pro Pro His Ala Val Lys Met Asp Val Phe Asn Leu Gly
                435                 440                 445

gac cag acg gct gtg ctg ctc aaa tca ctg gca gga gtt ccg ctg gcc      2295
Asp Gln Thr Ala Val Leu Leu Lys Ser Leu Ala Gly Val Pro Leu Ala
            450                 455                 460

agt gtg gag ggc cag aaa tac cac ctg aaa agc ggc cat gtt act tgt      2343
Ser Val Glu Gly Gln Lys Tyr His Leu Lys Ser Gly His Val Thr Cys
        465                 470                 475

gat gtg gga ctg gaa aag ctg aaa ctg aaa ggc aca acc tac tcc atg      2391
Asp Val Gly Leu Glu Lys Leu Lys Leu Lys Gly Thr Thr Tyr Ser Met
    480                 485                 490

tgt gac aaa gca aag ttc aaa tgg aag aga gtt cct gtg gac agc ggc      2439
Cys Asp Lys Ala Lys Phe Lys Trp Lys Arg Val Pro Val Asp Ser Gly
495                 500                 505                 510

cat gac aca gta gtc atg gag gta tca tac aca gga agc gac aag cca      2487
His Asp Thr Val Val Met Glu Val Ser Tyr Thr Gly Ser Asp Lys Pro
                515                 520                 525

tgt cgg atc ccg gtg cgg gct gtg gca cat ggt gtc cca gcg gtt aat      2535
Cys Arg Ile Pro Val Arg Ala Val Ala His Gly Val Pro Ala Val Asn
            530                 535                 540

gta gcc atg ctc ata acc ccc aat cca acc att gaa aca aat ggt ggc      2583
Val Ala Met Leu Ile Thr Pro Asn Pro Thr Ile Glu Thr Asn Gly Gly
        545                 550                 555

gga ttc ata gaa atg cag ctg cca cca ggg gat aac atc atc tat gtg      2631
Gly Phe Ile Glu Met Gln Leu Pro Pro Gly Asp Asn Ile Ile Tyr Val
    560                 565                 570

gga gac ctt agc cag cag tgg ttt cag aaa ggc agt acc att ggt aga      2679
Gly Asp Leu Ser Gln Gln Trp Phe Gln Lys Gly Ser Thr Ile Gly Arg
575                 580                 585                 590

atg ttt gaa aaa acc cgc agg gga ttg gaa agg ctc tct gtg gtt gga      2727
Met Phe Glu Lys Thr Arg Arg Gly Leu Glu Arg Leu Ser Val Val Gly
                595                 600                 605

gaa cat gca tgg gac ttt ggc tca gta ggc ggg gta ctg tct tct gtg      2775
Glu His Ala Trp Asp Phe Gly Ser Val Gly Gly Val Leu Ser Ser Val
            610                 615                 620

ggg aag gca atc cac acg gtg ctg ggg gga gct ttc aac acc ctt ttt      2823
Gly Lys Ala Ile His Thr Val Leu Gly Gly Ala Phe Asn Thr Leu Phe
        625                 630                 635

ggg ggg gtt gga ttc atc cct aag atg ctg ctg ggg gtt gct ctg gtc      2871
Gly Gly Val Gly Phe Ile Pro Lys Met Leu Leu Gly Val Ala Leu Val
    640                 645                 650

tgg ttg gga cta aat gcc agg aat cca acg atg tcc atg acg ttt ctt      2919
Trp Leu Gly Leu Asn Ala Arg Asn Pro Thr Met Ser Met Thr Phe Leu
655                 660                 665                 670
```

```
gct gtg ggg gct ttg aca ctg atg atg aca atg gga gtt ggg gca t       2965
Ala Val Gly Ala Leu Thr Leu Met Met Thr Met Gly Val Gly Ala
                675                 680                 685

gagcggccgc tcgagcatgc atctagaggg ccctatteta tagtgtcacc taaatgctag   3025
agctcgctga tcagcctcga ctgtgccttc tagttgccag ccatctgttg tttgcccctc   3085
ccccgtgcct tccttgaccc tggaaggtgc cactcccact gtcctttcct aataaaatga   3145
ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg gggtggggca  3205
ggacagcaag ggggaggatt gggaagacaa tagcaggcat gctggggatg cggtgggctc   3265
tatggcttct gaggcggaaa gaacagctgc attaatgaat cggccaacgc gcggggagag   3325
gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg cgctcggtcg   3385
ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta tccacagaat   3445
caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta   3505
aaaaggccgc gttgctggcg tttttccata ggctccgccc cctgacgag catcacaaaa    3565
atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc   3625
cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt   3685
ccgcctttct cccttcggga agcgtggcgc tttctcaatg ctcacgctgt aggtatctca   3745
gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaacccccc gttcagcccg   3805
accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat   3865
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta   3925
cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta tttggtatct   3985
gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac   4045
aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa   4105
aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa   4165
actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt   4225
taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact tggtctgaca   4285
gttaccaatg cttaatcagt gaggcacctа tctcagcgat ctgtctattt cgttcatcca   4345
tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta ccatctggcc   4405
ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta tcagcaataa   4465
accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc   4525
agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca   4585
acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat   4645
tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag   4705
cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac   4765
tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt   4825
ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg cgaccgagtt   4885
gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc   4945
tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat   5005
ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt actttcacca   5065
gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataagggcga   5125
cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg   5185
gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg   5245
ttccgcgcac atttccccga aaagtgccac ctgacgtc                           5283

<210> 20
<211> 681
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
      synthetic construct

<400> 20
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
 1               5                  10                  15
Val Val Ile Ala Gly Thr Ser Ala Val Thr Leu Val Arg Lys Asn Arg
            20                  25                  30
Trp Leu Leu Leu Asn Val Thr Ser Glu Asp Leu Gly Lys Thr Phe Ser
        35                  40                  45
```

```
Val Gly Thr Gly Asn Cys Thr Thr Asn Ile Leu Glu Ala Lys Tyr Trp
    50                  55                  60
Cys Pro Asp Ser Met Glu Tyr Asn Cys Pro Asn Leu Ser Pro Arg Glu
65                  70                  75                  80
Glu Pro Asp Asp Ile Asp Cys Trp Cys Tyr Gly Val Glu Asn Val Arg
                85                  90                  95
Val Ala Tyr Gly Lys Cys Asp Ser Ala Gly Arg Ser Arg Arg Ser Arg
            100                 105                 110
Arg Ala Ile Asp Leu Pro Thr His Glu Asn His Gly Leu Lys Thr Arg
        115                 120                 125
Gln Glu Lys Trp Met Thr Gly Arg Met Gly Glu Arg Gln Leu Gln Lys
    130                 135                 140
Ile Glu Arg Trp Phe Val Arg Asn Pro Phe Phe Ala Val Thr Ala Leu
145                 150                 155                 160
Thr Ile Ala Tyr Leu Val Gly Ser Asn Met Thr Gln Arg Val Val Ile
                165                 170                 175
Ala Leu Leu Val Leu Ala Val Gly Pro Ala Tyr Ser Ala His Cys Ile
            180                 185                 190
Gly Ile Thr Asp Arg Asp Phe Ile Glu Gly Val His Gly Gly Thr Trp
        195                 200                 205
Val Ser Ala Thr Leu Glu Gln Asp Lys Cys Val Thr Val Met Ala Pro
    210                 215                 220
Asp Lys Pro Ser Leu Asp Ile Ser Leu Glu Thr Val Ala Ile Asp Arg
225                 230                 235                 240
Pro Ala Glu Val Arg Lys Val Cys Tyr Asn Ala Val Leu Thr His Val
                245                 250                 255
Lys Ile Asn Asp Lys Cys Pro Ser Thr Gly Glu Ala His Leu Ala Glu
            260                 265                 270
Glu Asn Glu Gly Asp Asn Ala Cys Lys Arg Thr Tyr Ser Asp Arg Gly
        275                 280                 285
Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Val Ala Cys
    290                 295                 300
Ala Lys Phe Thr Cys Ala Lys Ser Met Ser Leu Phe Glu Val Asp Gln
305                 310                 315                 320
Thr Lys Ile Gln Tyr Val Ile Arg Ala Gln Leu His Val Gly Ala Lys
                325                 330                 335
Gln Glu Asn Trp Thr Thr Asp Ile Lys Thr Leu Lys Phe Asp Ala Leu
            340                 345                 350
Ser Gly Ser Gln Glu Val Glu Phe Ile Gly Tyr Gly Lys Ala Thr Leu
        355                 360                 365
Glu Cys Gln Val Gln Thr Ala Val Asp Phe Gly Asn Ser Tyr Ile Ala
    370                 375                 380
Glu Met Glu Thr Glu Ser Trp Ile Val Asp Arg Gln Trp Ala Gln Asp
385                 390                 395                 400
Leu Thr Leu Pro Trp Gln Ser Gly Ser Gly Gly Val Trp Arg Glu Met
                405                 410                 415
His His Leu Val Glu Phe Glu Pro Pro His Ala Ala Thr Ile Arg Val
            420                 425                 430
Leu Ala Leu Gly Asn Gln Glu Gly Ser Leu Lys Thr Ala Leu Thr Gly
        435                 440                 445
Ala Met Arg Val Thr Lys Asp Thr Asn Asp Asn Asn Leu Tyr Lys Leu
    450                 455                 460
His Gly Gly His Val Ser Cys Arg Val Lys Leu Ser Ala Leu Thr Leu
465                 470                 475                 480
Lys Gly Thr Ser Tyr Lys Ile Cys Thr Asp Lys Met Phe Phe Val Lys
                485                 490                 495
Asn Pro Thr Asp Thr Gly His Gly Thr Val Val Met Gln Val Lys Val
            500                 505                 510
Ser Lys Gly Ala Pro Cys Arg Ile Pro Val Ile Val Ala Asp Asp Leu
        515                 520                 525
```

```
Thr Ala Ala Ile Asn Lys Gly Ile Leu Val Thr Val Asn Pro Ile Ala
    530                 535                 540
Ser Thr Asn Asp Asp Glu Val Leu Ile Glu Val Asn Pro Pro Phe Gly
545                 550                 555                 560
Asp Ser Tyr Ile Ile Val Gly Arg Gly Asp Ser Arg Leu Thr Tyr Gln
                565                 570                 575
Trp His Lys Glu Gly Ser Ser Ile Gly Lys Leu Phe Thr Gln Thr Met
            580                 585                 590
Lys Gly Val Glu Arg Leu Ala Val Met Gly Asp Thr Ala Trp Asp Phe
        595                 600                 605
Ser Ser Ala Gly Gly Phe Phe Thr Ser Val Gly Lys Gly Ile His Thr
    610                 615                 620
Val Phe Gly Ser Ala Phe Gln Gly Leu Phe Gly Gly Leu Asn Trp Ile
625                 630                 635                 640
Thr Lys Val Ile Met Gly Ala Val Leu Ile Trp Val Gly Ile Asn Thr
                645                 650                 655
Arg Asn Met Thr Met Ser Met Ser Met Ile Leu Val Gly Val Ile Met
            660                 665                 670
Met Phe Leu Ser Leu Gly Val Gly Ala
        675                 680

<210> 21
<211> 5304
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
      synthetic construct

<221> CDS
<222> (910)...(2986)

<400> 21
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg     60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg    120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc    180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt    240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata    300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc    360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc    420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt    480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt    540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca    600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg    660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc    720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg    780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacccа    840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc    900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc    951
          Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser
           1               5                  10

ttg gca gtt gtc ata gct ggt aca agc gct ttg cag tta tca acc tat     999
Leu Ala Val Val Ile Ala Gly Thr Ser Ala Leu Gln Leu Ser Thr Tyr
 15                  20                  25                  30

cag ggg aaa gtg tta atg tca atc aac aag act gac gct caa agc gcc    1047
Gln Gly Lys Val Leu Met Ser Ile Asn Lys Thr Asp Ala Gln Ser Ala
                35                  40                  45
```

```
ata aac att cct agt gcc aac gga gca aac act tgc att gtg agg gct     1095
Ile Asn Ile Pro Ser Ala Asn Gly Ala Asn Thr Cys Ile Val Arg Ala
            50                  55                  60

cta gat gtg ggg gtc atg tgc aaa gat gac atc aca tac ctg tgc cca     1143
Leu Asp Val Gly Val Met Cys Lys Asp Asp Ile Thr Tyr Leu Cys Pro
        65                  70                  75

gtg ctt tca gcg gga aat gat ccc gag gac att gac tgt tgg tgt gac     1191
Val Leu Ser Ala Gly Asn Asp Pro Glu Asp Ile Asp Cys Trp Cys Asp
    80                  85                  90

gtc gaa gag gtg tgg gtg cac tac ggc aga tgc acg cgc atg gga cat     1239
Val Glu Glu Val Trp Val His Tyr Gly Arg Cys Thr Arg Met Gly His
 95                  100                 105                 110

tcg agg cgt agc cga cgg tca atc tct gtg cag cat cat gga gat tcc     1287
Ser Arg Arg Ser Arg Arg Ser Ile Ser Val Gln His His Gly Asp Ser
                115                 120                 125

aca ctg gca aca aag aac acg cca tgg ttg gac acc gtg aaa acc acc     1335
Thr Leu Ala Thr Lys Asn Thr Pro Trp Leu Asp Thr Val Lys Thr Thr
            130                 135                 140

aaa tac ttg aca aaa gta gaa aac tgg gtt ttg cgc aat cct gga tat     1383
Lys Tyr Leu Thr Lys Val Glu Asn Trp Val Leu Arg Asn Pro Gly Tyr
        145                 150                 155

gcc cta gtt gcg ctg gcg att gga tgg atg ctc ggt agc aac aac aca     1431
Ala Leu Val Ala Leu Ala Ile Gly Trp Met Leu Gly Ser Asn Asn Thr
    160                 165                 170

cag aga gtg gtt ttt gtg atc atg ctg atg ctg att gct ccg gca tac     1479
Gln Arg Val Val Phe Val Ile Met Leu Met Leu Ile Ala Pro Ala Tyr
175                 180                 185                 190

agc ttc aac tgt ctg gga aca tca aac agg gac ttt gtc gag gga gcc     1527
Ser Phe Asn Cys Leu Gly Thr Ser Asn Arg Asp Phe Val Glu Gly Ala
                195                 200                 205

agt ggg gca aca tgg att gac ttg gta ctt gaa ggg gga agc tgt gtc     1575
Ser Gly Ala Thr Trp Ile Asp Leu Val Leu Glu Gly Gly Ser Cys Val
            210                 215                 220

aca gtg atg gca cca gag aaa cca aca ctg gac ttc aaa gtg atg aag     1623
Thr Val Met Ala Pro Glu Lys Pro Thr Leu Asp Phe Lys Val Met Lys
        225                 230                 235

atg gag gct acc gag tta gcc act gtg cgt gag tat tgt tac gaa gca     1671
Met Glu Ala Thr Glu Leu Ala Thr Val Arg Glu Tyr Cys Tyr Glu Ala
    240                 245                 250

acc ttg gac acg ctg tca aca gtg gca agg tgc ccc aca aca gga gaa     1719
Thr Leu Asp Thr Leu Ser Thr Val Ala Arg Cys Pro Thr Thr Gly Glu
255                 260                 265                 270

gct cac aac acc aaa agg agt gac cca aca ttt gtc tgc aaa aga gat     1767
Ala His Asn Thr Lys Arg Ser Asp Pro Thr Phe Val Cys Lys Arg Asp
                275                 280                 285
```

```
gtt gtg gac cgc gga tgg ggt aac gga tgt ggt ctg ttt gga aaa ggg     1815
Val Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly
            290                 295                 300

agc att gac aca tgc gct aag ttc aca tgc aaa aac aag gca aca ggg     1863
Ser Ile Asp Thr Cys Ala Lys Phe Thr Cys Lys Asn Lys Ala Thr Gly
        305                 310                 315

aag acg atc ttg aga gaa aac atc aag tat gag gtt gca atc ttt gtg     1911
Lys Thr Ile Leu Arg Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe Val
    320                 325                 330

cat ggt tca acg gac tct acg tca cat ggc aat tac tct gag cag att     1959
His Gly Ser Thr Asp Ser Thr Ser His Gly Asn Tyr Ser Glu Gln Ile
335                 340                 345                 350

gga aaa aac caa gcg gct aga ttc acc ata agc ccg caa gca ccg tcc     2007
Gly Lys Asn Gln Ala Ala Arg Phe Thr Ile Ser Pro Gln Ala Pro Ser
                355                 360                 365

ttt acg gcc aac atg ggc gag tat gga aca gtt acc att gat tgt gaa     2055
Phe Thr Ala Asn Met Gly Glu Tyr Gly Thr Val Thr Ile Asp Cys Glu
            370                 375                 380

gca aga tca gga atc aac acg gag gat tat tat gtt ttc act gtc aag     2103
Ala Arg Ser Gly Ile Asn Thr Glu Asp Tyr Tyr Val Phe Thr Val Lys
        385                 390                 395

gag aag tca tgg cta gtg aac agg gac tgg ttt cac gac ttg aac ctt     2151
Glu Lys Ser Trp Leu Val Asn Arg Asp Trp Phe His Asp Leu Asn Leu
    400                 405                 410

cca tgg acg agc cct gcc aca act gat tgg cgc aac aga gaa aca ctg     2199
Pro Trp Thr Ser Pro Ala Thr Thr Asp Trp Arg Asn Arg Glu Thr Leu
415                 420                 425                 430

gtg gaa ttt gag gaa ccg cat gcc acc aag caa act gta gta gcc cta     2247
Val Glu Phe Glu Glu Pro His Ala Thr Lys Gln Thr Val Val Ala Leu
                435                 440                 445

gga tcg caa gaa ggt gcc ctg cac aca gca ttg gct gga gcc att cca     2295
Gly Ser Gln Glu Gly Ala Leu His Thr Ala Leu Ala Gly Ala Ile Pro
            450                 455                 460

gcc act gtt agc agc tca acc cta acc ttg caa tca ggg cat ttg aaa     2343
Ala Thr Val Ser Ser Ser Thr Leu Thr Leu Gln Ser Gly His Leu Lys
        465                 470                 475

tgc aga gct aag ctt gac aag gtc aaa atc aag gga acg aca tat ggc     2391
Cys Arg Ala Lys Leu Asp Lys Val Lys Ile Lys Gly Thr Thr Tyr Gly
    480                 485                 490

atg tgt gac tct gcc ttc acc ttc agc aag aac cca act gac aca ggg     2439
Met Cys Asp Ser Ala Phe Thr Phe Ser Lys Asn Pro Thr Asp Thr Gly
495                 500                 505                 510

cac ggg aca gtg att gtg gaa ctg cag tat act gga agc aac gga ccc     2487
His Gly Thr Val Ile Val Glu Leu Gln Tyr Thr Gly Ser Asn Gly Pro
                515                 520                 525
```

```
tgc cga gtt ccc atc tcc gtg act gca aac ctc atg gat ttg aca ccg     2535
Cys Arg Val Pro Ile Ser Val Thr Ala Asn Leu Met Asp Leu Thr Pro
            530                 535                 540

gtt gga aga ttg gtc acg gtc aat ccc ttt ata agc aca ggg gga gcg     2583
Val Gly Arg Leu Val Thr Val Asn Pro Phe Ile Ser Thr Gly Gly Ala
        545                 550                 555

aac aac aag gtc atg atc gaa gtt gaa cca ccc ttt ggc gat tct tac     2631
Asn Asn Lys Val Met Ile Glu Val Glu Pro Pro Phe Gly Asp Ser Tyr
    560                 565                 570

atc gtc gtc gga aga ggc acc acc cag att aac tac cac tgg cac aaa     2679
Ile Val Val Gly Arg Gly Thr Thr Gln Ile Asn Tyr His Trp His Lys
575                 580                 585                 590

gag gga agc agc att ggg aag gct ttg gcg acc aca tgg aaa gga gcc     2727
Glu Gly Ser Ser Ile Gly Lys Ala Leu Ala Thr Thr Trp Lys Gly Ala
                595                 600                 605

caa cgg cta gcc gtc tta ggg gac aca gcg tgg gac ttt gga tct att     2775
Gln Arg Leu Ala Val Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Ile
            610                 615                 620

gga gga gtt ttc aat tca att ggc aaa gct gtc cac caa gtt ttc gga     2823
Gly Gly Val Phe Asn Ser Ile Gly Lys Ala Val His Gln Val Phe Gly
        625                 630                 635

gga gcg ttc agg act ctg ttc ggg gga atg tcc tgg atc aca cag ggg     2871
Gly Ala Phe Arg Thr Leu Phe Gly Gly Met Ser Trp Ile Thr Gln Gly
    640                 645                 650

cta ctt gga gct ctt ctc ctg tgg atg ggg ttg cag gcc cgc gac agg     2919
Leu Leu Gly Ala Leu Leu Leu Trp Met Gly Leu Gln Ala Arg Asp Arg
655                 660                 665                 670

agc atc tcg ctg act cta ctg gct gtc gga ggg att ctc atc ttt ctg     2967
Ser Ile Ser Leu Thr Leu Leu Ala Val Gly Gly Ile Leu Ile Phe Leu
                675                 680                 685

gca acc agc gtg caa gcc t gagcggccgc tcgagcatgc atctagaggg          3016
Ala Thr Ser Val Gln Ala
            690

ccctattcta tagtgtcacc taaatgctag agctcgctga tcagcctcga ctgtgccttc   3076
tagttgccag ccatctgttg tttgcccctc ccccgtgcct tccttgaccc tggaaggtgc   3136
cactcccact gtcctttcct aataaaatga ggaaattgca tcgcattgtc tgagtaggtg   3196
tcattctatt ctgggggggtg gggtggggca ggacagcaag ggggaggatt gggaagacaa   3256
tagcaggcat gctggggatg cggtgggctc tatggcttct gaggcggaaa gaacagctgc   3316
attaatgaat cggccaacgc gcggggagag gcggtttgcg tattgggcgc tcttccgctt   3376
cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagctcact   3436
caaaggcggt aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag   3496
caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata   3556
ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc   3616
cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg   3676
ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccttcgggg agcgtggcgc   3736
tttctcaatg ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc tccaagctgg   3796
gctgtgtgca cgaaccccccc gttcagcccg accgctgcgc cttatccggt aactatcgtc   3856
ttgagtccaa cccggtaaga cacgacttat cgccactggc agcagccact ggtaacagga   3916
ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtggtgg cctaactacg   3976
gctacactag aaggacagta tttggtatct gcgctctgct gaagccagtt accttcggaa   4036
```

aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggt ggtttttttg 4096
tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt 4156
ctacggggtc tgacgctcag tggaacgaaa actcacgtta agggattttg gtcatgagat 4216
tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct 4276
aaagtatata tgagtaaact tggtctgaca gttaccaatg cttaatcagt gaggcaccta 4336
tctcagcgat ctgtctattt cgttcatcca tagttgcctg actccccgtc gtgtagataa 4396
ctacgatacg ggagggctta ccatctggcc ccagtgctgc aatgataccg cgagacccac 4456
gctcaccggc tccagattta tcagcaataa accagccagc cggaagggcc gagcgcagaa 4516
gtggtcctgc aactttatcc gcctccatcc agtctattaa ttgttgccgg gaagctagag 4576
taagtagttc gccagttaat agtttgcgca acgttgttgc cattgctaca ggcatcgtgg 4636
tgtcacgctc gtcgtttggt atggcttcat tcagctccgg ttcccaacga tcaaggcgag 4696
ttacatgatc ccccatgttg tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg 4756
tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg cataattctc 4816
ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca accaagtcat 4876
tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata cgggataata 4936
ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa 4996
aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaacccact cgtgcaccca 5056
actgatcttc agcatctttt actttcacca gcgtttctgg gtgagcaaaa acaggaaggc 5116
aaaatgccgc aaaaaaggga ataagggcga cacggaaatg ttgaatactc atactcttcc 5176
tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg 5236
aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac 5296
ctgacgtc 5304

<210> 22
<211> 692
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 22

Met	Gly	Lys	Arg	Ser	Ala	Gly	Ser	Ile	Met	Trp	Leu	Ala	Ser	Leu	Ala
1				5					10					15	
Val	Val	Ile	Ala	Gly	Thr	Ser	Ala	Leu	Gln	Leu	Ser	Thr	Tyr	Gln	Gly
			20					25					30		
Lys	Val	Leu	Met	Ser	Ile	Asn	Lys	Thr	Asp	Ala	Gln	Ser	Ala	Ile	Asn
		35					40					45			
Ile	Pro	Ser	Ala	Asn	Gly	Ala	Asn	Thr	Cys	Ile	Val	Arg	Ala	Leu	Asp
	50					55					60				
Val	Gly	Val	Met	Cys	Lys	Asp	Asp	Ile	Thr	Tyr	Leu	Cys	Pro	Val	Leu
65					70					75					80
Ser	Ala	Gly	Asn	Asp	Pro	Glu	Asp	Ile	Asp	Cys	Trp	Cys	Asp	Val	Glu
				85					90					95	
Glu	Val	Trp	Val	His	Tyr	Gly	Arg	Cys	Thr	Arg	Met	Gly	His	Ser	Arg
			100					105					110		
Arg	Ser	Arg	Arg	Ser	Ile	Ser	Val	Gln	His	His	Gly	Asp	Ser	Thr	Leu
		115					120					125			
Ala	Thr	Lys	Asn	Thr	Pro	Trp	Leu	Asp	Thr	Val	Lys	Thr	Thr	Lys	Tyr
	130					135					140				
Leu	Thr	Lys	Val	Glu	Asn	Trp	Val	Leu	Arg	Asn	Pro	Gly	Tyr	Ala	Leu
145					150					155					160
Val	Ala	Leu	Ala	Ile	Gly	Trp	Met	Leu	Gly	Ser	Asn	Asn	Thr	Gln	Arg
				165					170					175	
Val	Val	Phe	Val	Ile	Met	Leu	Met	Leu	Ile	Ala	Pro	Ala	Tyr	Ser	Phe
			180					185					190		
Asn	Cys	Leu	Gly	Thr	Ser	Asn	Arg	Asp	Phe	Val	Glu	Gly	Ala	Ser	Gly
		195					200					205			
Ala	Thr	Trp	Ile	Asp	Leu	Val	Leu	Glu	Gly	Gly	Ser	Cys	Val	Thr	Val
	210					215					220				

```
Met Ala Pro Glu Lys Pro Thr Leu Asp Phe Lys Val Met Lys Met Glu
225                 230                 235                 240
Ala Thr Glu Leu Ala Thr Val Arg Glu Tyr Cys Tyr Glu Ala Thr Leu
                245                 250                 255
Asp Thr Leu Ser Thr Val Ala Arg Cys Pro Thr Thr Gly Glu Ala His
            260                 265                 270
Asn Thr Lys Arg Ser Asp Pro Thr Phe Val Cys Lys Arg Asp Val Val
        275                 280                 285
Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile
    290                 295                 300
Asp Thr Cys Ala Lys Phe Thr Cys Lys Asn Lys Ala Thr Gly Lys Thr
305                 310                 315                 320
Ile Leu Arg Glu Asn Ile Lys Tyr Glu Val Ala Ile Phe Val His Gly
                325                 330                 335
Ser Thr Asp Ser Thr Ser His Gly Asn Tyr Ser Glu Gln Ile Gly Lys
            340                 345                 350
Asn Gln Ala Ala Arg Phe Thr Ile Ser Pro Gln Ala Pro Ser Phe Thr
        355                 360                 365
Ala Asn Met Gly Glu Tyr Gly Thr Val Thr Ile Asp Cys Glu Ala Arg
    370                 375                 380
Ser Gly Ile Asn Thr Glu Asp Tyr Tyr Val Phe Thr Val Lys Glu Lys
385                 390                 395                 400
Ser Trp Leu Val Asn Arg Asp Trp Phe His Asp Leu Asn Leu Pro Trp
                405                 410                 415
Thr Ser Pro Ala Thr Thr Asp Trp Arg Asn Arg Glu Thr Leu Val Glu
            420                 425                 430
Phe Glu Glu Pro His Ala Thr Lys Gln Thr Val Val Ala Leu Gly Ser
        435                 440                 445
Gln Glu Gly Ala Leu His Thr Ala Leu Ala Gly Ala Ile Pro Ala Thr
    450                 455                 460
Val Ser Ser Ser Thr Leu Thr Leu Gln Ser Gly His Leu Lys Cys Arg
465                 470                 475                 480
Ala Lys Leu Asp Lys Val Lys Ile Lys Gly Thr Thr Tyr Gly Met Cys
                485                 490                 495
Asp Ser Ala Phe Thr Phe Ser Lys Asn Pro Thr Asp Thr Gly His Gly
            500                 505                 510
Thr Val Ile Val Glu Leu Gln Tyr Thr Gly Ser Asn Gly Pro Cys Arg
        515                 520                 525
Val Pro Ile Ser Val Thr Ala Asn Leu Met Asp Leu Thr Pro Val Gly
    530                 535                 540
Arg Leu Val Thr Val Asn Pro Phe Ile Ser Thr Gly Gly Ala Asn Asn
545                 550                 555                 560
Lys Val Met Ile Glu Val Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val
                565                 570                 575
Val Gly Arg Gly Thr Thr Gln Ile Asn Tyr His Trp His Lys Glu Gly
            580                 585                 590
Ser Ser Ile Gly Lys Ala Leu Ala Thr Thr Trp Lys Gly Ala Gln Arg
        595                 600                 605
Leu Ala Val Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Ile Gly Gly
    610                 615                 620
Val Phe Asn Ser Ile Gly Lys Ala Val His Gln Val Phe Gly Gly Ala
625                 630                 635                 640
Phe Arg Thr Leu Phe Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Leu
                645                 650                 655
Gly Ala Leu Leu Leu Trp Met Gly Leu Gln Ala Arg Asp Arg Ser Ile
            660                 665                 670
Ser Leu Thr Leu Leu Ala Val Gly Gly Ile Leu Ile Phe Leu Ala Thr
        675                 680                 685
Ser Val Gln Ala
    690
```

<210> 23
<211> 5271
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (910)...(2953)

<400> 23
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc 900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc 951
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser
1 5 10

ttg gca gtt gtc ata gct ggt aca agc gct gtg acc ttg gtg cgg aaa 999
Leu Ala Val Val Ile Ala Gly Thr Ser Ala Val Thr Leu Val Arg Lys
15 20 25 30

aac aga tgg ttg ctc cta aat gtg aca tct gag gac ctc ggg aaa aca 1047
Asn Arg Trp Leu Leu Leu Asn Val Thr Ser Glu Asp Leu Gly Lys Thr
35 40 45

ttc tct gtg ggc aca ggc aac tgc aca aca aac att ttg gaa gcc aag 1095
Phe Ser Val Gly Thr Gly Asn Cys Thr Thr Asn Ile Leu Glu Ala Lys
50 55 60

tac tgg tgc cca gac tca atg gaa tac aac tgt ccc aat ctc agt cca 1143
Tyr Trp Cys Pro Asp Ser Met Glu Tyr Asn Cys Pro Asn Leu Ser Pro
65 70 75

aga gag gag cca gat gac att gat tgc tgg tgc tat ggg gtg gaa aac 1191
Arg Glu Glu Pro Asp Asp Ile Asp Cys Trp Cys Tyr Gly Val Glu Asn
80 85 90

gtt aga gtc gca tat ggt aag tgt gac tca gca ggc agg tct agg agg 1239
Val Arg Val Ala Tyr Gly Lys Cys Asp Ser Ala Gly Arg Ser Arg Arg
95 100 105 110

tca aga agg gcc att gac ttg cct acg cat gaa aac cat ggt ttg aag 1287
Ser Arg Arg Ala Ile Asp Leu Pro Thr His Glu Asn His Gly Leu Lys
115 120 125

```
acc cgg caa gaa aaa tgg atg act gga aga atg ggt gaa agg caa ctc     1335
Thr Arg Gln Glu Lys Trp Met Thr Gly Arg Met Gly Glu Arg Gln Leu
            130                 135                 140

caa aag att gag aga tgg ttc gtg agg aac ccc ttt ttt gca gtg acg     1383
Gln Lys Ile Glu Arg Trp Phe Val Arg Asn Pro Phe Phe Ala Val Thr
        145                 150                 155

gct ctg acc att gcc tac ctt gtg gga agc aac atg acg caa cga gtc     1431
Ala Leu Thr Ile Ala Tyr Leu Val Gly Ser Asn Met Thr Gln Arg Val
    160                 165                 170

gtg att gcc cta ctg gtc ttg gct gtt ggt ccg gcc tac tca gct cac     1479
Val Ile Ala Leu Leu Val Leu Ala Val Gly Pro Ala Tyr Ser Ala His
175                 180                 185                 190

tgc att gga att act gac agg gat ttc att gag ggg gtg cat gga gga     1527
Cys Ile Gly Ile Thr Asp Arg Asp Phe Ile Glu Gly Val His Gly Gly
                195                 200                 205

act tgg gtt tca gct acc ctg gag caa gac aag tgt gtc act gtt atg     1575
Thr Trp Val Ser Ala Thr Leu Glu Gln Asp Lys Cys Val Thr Val Met
            210                 215                 220

gcc cct gac aag cct tca ttg gac atc tca cta gag aca gta gcc att     1623
Ala Pro Asp Lys Pro Ser Leu Asp Ile Ser Leu Glu Thr Val Ala Ile
        225                 230                 235

gat aga cct gct gag gtg agg aaa gtg tgt tac aat gca gtt ctc act     1671
Asp Arg Pro Ala Glu Val Arg Lys Val Cys Tyr Asn Ala Val Leu Thr
    240                 245                 250

cat gtg aag att aat gac aag tgc ccc agc act gga gag gcc cac cta     1719
His Val Lys Ile Asn Asp Lys Cys Pro Ser Thr Gly Glu Ala His Leu
255                 260                 265                 270

gct gaa gag aac gaa ggg gac aat gcg tgc aag cgc act tat tct gat     1767
Ala Glu Glu Asn Glu Gly Asp Asn Ala Cys Lys Arg Thr Tyr Ser Asp
                275                 280                 285

aga ggc tgg ggc aat ggc tgt ggc cta ttt ggg aaa ggg agc att gtg     1815
Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Val
            290                 295                 300

gca tgc gcc aaa ttc act tgt gcc aaa tcc atg agt ttg ttt gag gtt     1863
Ala Cys Ala Lys Phe Thr Cys Ala Lys Ser Met Ser Leu Phe Glu Val
        305                 310                 315

gat cag acc aaa att cag tat gtc atc aga gca caa ttg cat gta ggg     1911
Asp Gln Thr Lys Ile Gln Tyr Val Ile Arg Ala Gln Leu His Val Gly
    320                 325                 330

gcc aag cag gaa aat tgg act acc gac att aag act ctc aag ttt gat     1959
Ala Lys Gln Glu Asn Trp Thr Thr Asp Ile Lys Thr Leu Lys Phe Asp
335                 340                 345                 350

gcc ctg tca ggc tcc cag gaa gtc gag ttc att ggg tat gga aaa gct     2007
Ala Leu Ser Gly Ser Gln Glu Val Glu Phe Ile Gly Tyr Gly Lys Ala
                355                 360                 365
```

```
aca ctg gaa tgc cag gtg caa act gcg gtg gac ttt ggt aac agt tac     2055
Thr Leu Glu Cys Gln Val Gln Thr Ala Val Asp Phe Gly Asn Ser Tyr
            370                 375                 380

atc gct gag atg gaa aca gag agc tgg ata gtg gac aga cag tgg gcc     2103
Ile Ala Glu Met Glu Thr Glu Ser Trp Ile Val Asp Arg Gln Trp Ala
        385                 390                 395

cag gac ttg acc ctg cca tgg cag agt gga agt ggc ggg gtg tgg aga     2151
Gln Asp Leu Thr Leu Pro Trp Gln Ser Gly Ser Gly Gly Val Trp Arg
    400                 405                 410

gag atg cat cat ctt gtc gaa ttt gaa cct ccg cat gcc gcc act atc     2199
Glu Met His His Leu Val Glu Phe Glu Pro Pro His Ala Ala Thr Ile
415                 420                 425                 430

aga gta ctg gcc ctg gga aac cag gaa ggc tcc ttg aaa aca gct ctt     2247
Arg Val Leu Ala Leu Gly Asn Gln Glu Gly Ser Leu Lys Thr Ala Leu
                435                 440                 445

act ggc gca atg agg gtt aca aag gac aca aat gac aac aac ctt tac     2295
Thr Gly Ala Met Arg Val Thr Lys Asp Thr Asn Asp Asn Asn Leu Tyr
            450                 455                 460

aaa cta cat ggt gga cat gtt tct tgc aga gtg aaa ttg tca gct ttg     2343
Lys Leu His Gly Gly His Val Ser Cys Arg Val Lys Leu Ser Ala Leu
        465                 470                 475

aca ctc aag ggg aca tcc tac aaa ata tgc act gac aaa atg ttt ttt     2391
Thr Leu Lys Gly Thr Ser Tyr Lys Ile Cys Thr Asp Lys Met Phe Phe
    480                 485                 490

gtc aag aac cca act gac act ggc cat ggc act gtt gtg atg cag gtg     2439
Val Lys Asn Pro Thr Asp Thr Gly His Gly Thr Val Val Met Gln Val
495                 500                 505                 510

aaa gtg tca aaa gga gcc ccc tgc agg att cca gtg ata gta gct gat     2487
Lys Val Ser Lys Gly Ala Pro Cys Arg Ile Pro Val Ile Val Ala Asp
                515                 520                 525

gat ctt aca gcg gca atc aat aaa ggc att ttg gtt aca gtt aac ccc     2535
Asp Leu Thr Ala Ala Ile Asn Lys Gly Ile Leu Val Thr Val Asn Pro
            530                 535                 540

atc gcc tca acc aat gat gat gaa gtg ctg att gag gtg aac cca cct     2583
Ile Ala Ser Thr Asn Asp Asp Glu Val Leu Ile Glu Val Asn Pro Pro
        545                 550                 555

ttt gga gac agc tac att atc gtt ggg aga gga gat tca cgt ctc act     2631
Phe Gly Asp Ser Tyr Ile Ile Val Gly Arg Gly Asp Ser Arg Leu Thr
    560                 565                 570

tac cag tgg cac aaa gag gga agc tca ata gga aag ttg ttc act cag     2679
Tyr Gln Trp His Lys Glu Gly Ser Ser Ile Gly Lys Leu Phe Thr Gln
575                 580                 585                 590

acc atg aaa ggc gtg gaa cgc ctg gcc gtc atg gga gac acc gcc tgg     2727
Thr Met Lys Gly Val Glu Arg Leu Ala Val Met Gly Asp Thr Ala Trp
                595                 600                 605
```

```
gat ttc agc tcc gct gga ggg ttc ttc act tcg gtt ggg aaa gga att     2775
Asp Phe Ser Ser Ala Gly Gly Phe Phe Thr Ser Val Gly Lys Gly Ile
            610                 615                 620

cat acg gtg ttt ggc tct gcc ttt cag ggg cta ttt ggc ggc ttg aac     2823
His Thr Val Phe Gly Ser Ala Phe Gln Gly Leu Phe Gly Gly Leu Asn
        625                 630                 635

tgg ata aca aag gtc atc atg ggg gcg gta ctt ata tgg gtt ggc atc     2871
Trp Ile Thr Lys Val Ile Met Gly Ala Val Leu Ile Trp Val Gly Ile
    640                 645                 650

aac aca aga aac atg aca atg tcc atg agc atg atc ttg gta gga gtg     2919
Asn Thr Arg Asn Met Thr Met Ser Met Ser Met Ile Leu Val Gly Val
655                 660                 665                 670

atc atg atg ttt ttg tct cta gga gtt ggg gcg t gagcggccgc           2963
Ile Met Met Phe Leu Ser Leu Gly Val Gly Ala
                675                 680

tcgagcatgc atctagaggg ccctattcta tagtgtcacc taaatgctag agctcgctga   3023
tcagcctcga ctgtgccttc tagttgccag ccatctgttg tttgcccctc ccccgtgcct   3083
tccttgaccc tggaaggtgc cactcccact gtcctttcct aataaaatga ggaaattgca   3143
tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg gggtgggggca ggacagcaag   3203
ggggaggatt gggaagacaa tagcaggcat gctgggggatg cggtgggctc tatggcttct   3263
gaggcggaaa gaacagctgc attaatgaat cggccaacgc gcgggggagag gcggtttgcg   3323
tattgggcgc tcttccgctt cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg   3383
gcgagcggta tcagctcact caaaggcggt aatacggtta tccacagaat caggggataa   3443
cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc   3503
gttgctggcg ttttttccata ggctccgccc ccctgacgag catcacaaaa atcgacgctc   3563
aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag   3623
ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggatacctgt ccgcctttct   3683
ccccttcggga agcgtggcgc tttctcaatg ctcacgctgt aggtatctca gttcggtgta   3743
ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccg accgctgcgc   3803
cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat cgccactggc   3863
agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt   3923
gaagtggtgg cctaactacg gctacactag aaggacagta tttggtatct gcgctctgct   3983
gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc   4043
tggtagcggt ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca   4103
agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta   4163
agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa   4223
atgaagtttt aaatcaatct aaagtatata tgagtaaact tggtctgaca gttaccaatg   4283
cttaatcagt gaggcaccta tctcagcgat ctgtctattt cgttcatcca tagttgcctg   4343
actccccgtc gtgtagataa ctacgatacg ggagggctta ccatctggcc ccagtgctgc   4403
aatgataccg cgagacccac gctcaccggc tccagattta tcagcaataa accagccagc   4463
cggaagggcc gagcgcagaa gtggtcctgc aactttatcc gcctccatcc agtctattaa   4523
ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca acgttgttgc   4583
cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat tcagctccgg   4643
ttcccaacga tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag cggttagctc   4703
cttcggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat   4763
ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt ctgtgactgg   4823
tgagtactca accaagtcat tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc   4883
ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg   4943
aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat   5003
gtaacccact cgtgcaccca actgatcttc agcatctttt actttcacca gcgtttctgg   5063
gtgagcaaaa acaggaaggc aaaatgccgc aaaaaagggа ataagggcga cacggaaatg   5123
ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg gttattgtct   5183
catgagcgga tacatatttg aatgtattta gaaaaataaa caaatagggg ttccgcgcac   5243
atttccccga aaagtgccac ctgacgtc                                      5271
```

<210> 24

<211> 681
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 24

```
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
 1               5                  10                  15
Val Val Ile Ala Gly Thr Ser Ala Val Thr Leu Val Arg Lys Asn Arg
            20                  25                  30
Trp Leu Leu Leu Asn Val Thr Ser Glu Asp Leu Gly Lys Thr Phe Ser
        35                  40                  45
Val Gly Thr Gly Asn Cys Thr Thr Asn Ile Leu Glu Ala Lys Tyr Trp
    50                  55                  60
Cys Pro Asp Ser Met Glu Tyr Asn Cys Pro Asn Leu Ser Pro Arg Glu
65                  70                  75                  80
Glu Pro Asp Asp Ile Asp Cys Trp Cys Tyr Gly Val Glu Asn Val Arg
                85                  90                  95
Val Ala Tyr Gly Lys Cys Asp Ser Ala Gly Arg Ser Arg Arg Ser Arg
            100                 105                 110
Arg Ala Ile Asp Leu Pro Thr His Glu Asn His Gly Leu Lys Thr Arg
        115                 120                 125
Gln Glu Lys Trp Met Thr Gly Arg Met Gly Glu Arg Gln Leu Gln Lys
    130                 135                 140
Ile Glu Arg Trp Phe Val Arg Asn Pro Phe Phe Ala Val Thr Ala Leu
145                 150                 155                 160
Thr Ile Ala Tyr Leu Val Gly Ser Asn Met Thr Gln Arg Val Val Ile
                165                 170                 175
Ala Leu Leu Val Leu Ala Val Gly Pro Ala Tyr Ser Ala His Cys Ile
            180                 185                 190
Gly Ile Thr Asp Arg Asp Phe Ile Glu Gly Val His Gly Gly Thr Trp
        195                 200                 205
Val Ser Ala Thr Leu Glu Gln Asp Lys Cys Val Thr Val Met Ala Pro
    210                 215                 220
Asp Lys Pro Ser Leu Asp Ile Ser Leu Glu Thr Val Ala Ile Asp Arg
225                 230                 235                 240
Pro Ala Glu Val Arg Lys Val Cys Tyr Asn Ala Val Leu Thr His Val
                245                 250                 255
Lys Ile Asn Asp Lys Cys Pro Ser Thr Gly Glu Ala His Leu Ala Glu
            260                 265                 270
Glu Asn Glu Gly Asp Asn Ala Cys Lys Arg Thr Tyr Ser Asp Arg Gly
        275                 280                 285
Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Ile Val Ala Cys
    290                 295                 300
Ala Lys Phe Thr Cys Ala Lys Ser Met Ser Leu Phe Glu Val Asp Gln
305                 310                 315                 320
Thr Lys Ile Gln Tyr Val Ile Arg Ala Gln Leu His Val Gly Ala Lys
                325                 330                 335
Gln Glu Asn Trp Thr Thr Asp Ile Lys Thr Leu Lys Phe Asp Ala Leu
            340                 345                 350
Ser Gly Ser Gln Glu Val Glu Phe Ile Gly Tyr Gly Lys Ala Thr Leu
        355                 360                 365
Glu Cys Gln Val Gln Thr Ala Val Asp Phe Gly Asn Ser Tyr Ile Ala
    370                 375                 380
Glu Met Glu Thr Glu Ser Trp Ile Val Asp Arg Gln Trp Ala Gln Asp
385                 390                 395                 400
Leu Thr Leu Pro Trp Gln Ser Gly Ser Gly Gly Val Trp Arg Glu Met
                405                 410                 415
```

```
His His Leu Val Glu Phe Glu Pro Pro His Ala Ala Thr Ile Arg Val
            420                 425                 430
Leu Ala Leu Gly Asn Gln Glu Gly Ser Leu Lys Thr Ala Leu Thr Gly
        435                 440                 445
Ala Met Arg Val Thr Lys Asp Thr Asn Asp Asn Asn Leu Tyr Lys Leu
    450                 455                 460
His Gly Gly His Val Ser Cys Arg Val Lys Leu Ser Ala Leu Thr Leu
465                 470                 475                 480
Lys Gly Thr Ser Tyr Lys Ile Cys Thr Asp Lys Met Phe Phe Val Lys
                485                 490                 495
Asn Pro Thr Asp Thr Gly His Gly Thr Val Val Met Gln Val Lys Val
            500                 505                 510
Ser Lys Gly Ala Pro Cys Arg Ile Pro Val Ile Val Ala Asp Asp Leu
        515                 520                 525
Thr Ala Ala Ile Asn Lys Gly Ile Leu Val Thr Val Asn Pro Ile Ala
    530                 535                 540
Ser Thr Asn Asp Asp Glu Val Leu Ile Glu Val Asn Pro Pro Phe Gly
545                 550                 555                 560
Asp Ser Tyr Ile Ile Val Gly Arg Gly Asp Ser Arg Leu Thr Tyr Gln
                565                 570                 575
Trp His Lys Glu Gly Ser Ser Ile Gly Lys Leu Phe Thr Gln Thr Met
            580                 585                 590
Lys Gly Val Glu Arg Leu Ala Val Met Gly Asp Thr Ala Trp Asp Phe
        595                 600                 605
Ser Ser Ala Gly Gly Phe Phe Thr Ser Val Gly Lys Gly Ile His Thr
    610                 615                 620
Val Phe Gly Ser Ala Phe Gln Gly Leu Phe Gly Gly Leu Asn Trp Ile
625                 630                 635                 640
Thr Lys Val Ile Met Gly Ala Val Leu Ile Trp Val Gly Ile Asn Thr
                645                 650                 655
Arg Asn Met Thr Met Ser Met Ser Met Ile Leu Val Gly Val Ile Met
            660                 665                 670
Met Phe Leu Ser Leu Gly Val Gly Ala
        675                 680
```

<210> 25
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-35
<223> POW 454

<400> 25
aaaagaaaaa gcgctaccac catccaccgg gacag 35

<210> 26
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature

<222> 1-41
<223> CPOW 2417

<400> 26
actgttaccc tcaaccccgt actcgccggc gaaaaagaaa a 41

<210> 27
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<223> Modified JE Signal

<400> 27
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
1 5 10 15
Val Val Ile Ala Gly Thr Ser Ala
20

<210> 28
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-36
<223> YF 482

<400> 28
aaaagaaaaa gcgctgtgac cttggtgcgg aaaaac 36

<210> 29
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-41
<223> CYF 2433

<400> 29
acagagatcc tcaaccccgc actcgccggc gaaaaagaaa a 41

<210> 30
<211> 41
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-41
<223> SLE 463

<400> 30
aaaagaaaaa gcgctttgca gttatcaacc tatcagggga a 41

<210> 31
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> misc_feature
<222> 1-40
<223> CSLE 2477

<400> 31
accgttggtc gcacgttcgg actcgccggc gaaaaagaaa 40

<210> 32
<211> 39
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 32
Leu Asp Thr Ile Asn Arg Arg Pro Ser Lys Lys Arg Gly Gly Thr Arg
1 5 10 15
Ser Leu Leu Gly Leu Ala Ala Leu Ile Gly Leu Ala Ser Ser Leu Gln
20 25 30
Leu Leu Ser Thr Tyr Gln Gly
35

<210> 33
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 33
Met Trp Leu Ala Ser Leu Ala Val Val Ile Ala Cys Ala Gly Ala Met
1 5 10 15
Lys Leu Ser Asn Phe Gln Gly Lys
20

<210> 34
<211> 30
<212> PRT

<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 34
Met Asn Glu Gly Ser Ile Met Trp Leu Ala Ser Leu Ala Val Val Ile
1 5 10 15
Ala Cys Ala Gly Ala Met Lys Leu Ser Asn Phe Gln Gly Lys
20 25 30

<210> 35
<211> 39
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 35
Met Gly Arg Lys Gln Asn Lys Arg Gly Gly Asn Glu Gly Ser Ile Met
1 5 10 15
Trp Leu Ala Ser Leu Ala Val Val Ile Ala Cys Ala Gly Ala Met Lys
20 25 30
Leu Ser Asn Phe Gln Gly Lys
35

<210> 36
<211> 34
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 36
Met Ser Lys Lys Arg Gly Gly Ser Glu Thr Ser Val Leu Met Val Ile
1 5 10 15
Phe Met Leu Ile Gly Phe Ala Ala Ala Leu Lys Leu Ser Asn Phe Gln
20 25 30
Gly Lys

<210> 37
<211> 33
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 37
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
1 5 10 15
Val Val Ile Ala Cys Ala Gly Ala Val Thr Leu Ser Asn Phe Gln Gly
20 25 30
Lys

<210> 38
<211> 46
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 38
Met Asn Val Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn
1 5 10 15
Ile Leu Asn Arg Arg Arg Arg Thr Ala Gly Met Ile Ile Met Leu Ile
20 25 30
Pro Thr Val Met Ala Phe His Leu Thr Thr Arg Asn Gly Glu
35 40 45

<210> 39
<211> 40
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 39
Met Val Gly Leu Gln Lys Arg Gly Lys Arg Arg Ser Ala Thr Asp Trp
1 5 10 15
Met Ser Trp Leu Leu Val Ile Thr Leu Leu Gly Met Thr Leu Ala Ala
20 25 30
Thr Val Arg Lys Glu Arg Gly Asp
35 40

<210> 40
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 40
Met Gly Trp Leu Leu Val Val Val Leu Leu Gly Val Thr Leu Ala Ala
1 5 10 15
Thr Val Arg Lys Glu Arg Gly Asp
20

<210> 41
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 41

```
Met Ser Trp Leu Leu Val Ile Thr Leu Leu Gly Met Thr Ile Ala Ala
 1               5                  10                  15
Thr Val Arg Lys Glu Arg Gly Asp
            20
```

<210> 42
<211> 5292
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (910)...(2964)

<400> 42

```
gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg      60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg     120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc     180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt     240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata     300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc     360
ccccgcccat tgacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc     420
attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt     480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt     540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca     600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg     660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc     720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg     780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca     840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc     900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc      951
          Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser
           1               5                  10

ttg gca gtt gtc ata gct tgt gca ggc gcc ttc cat tta acc aca cgt       999
Leu Ala Val Val Ile Ala Cys Ala Gly Ala Phe His Leu Thr Thr Arg
 15                  20                  25                  30

aac gga gaa cca cac atg atc gtc agc aga caa gag aaa ggg aaa agt      1047
Asn Gly Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly Lys Ser
                 35                  40                  45

ctt ctg ttt aaa aca gag gat ggc gtg aac atg tgt acc ctc atg gcc      1095
Leu Leu Phe Lys Thr Glu Asp Gly Val Asn Met Cys Thr Leu Met Ala
             50                  55                  60

atg gac ctt ggt gaa ttg tgt gaa gac aca atc acg tac aag tgt ccc      1143
Met Asp Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys Cys Pro
         65                  70                  75

ctt ctc agg cag aat gag cca gaa gac ata gac tgt tgg tgc aac tct      1191
Leu Leu Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Ser
     80                  85                  90

acg tcc acg tgg gta act tat ggg acg tgt acc acc atg gga gaa cat      1239
Thr Ser Thr Trp Val Thr Tyr Gly Thr Cys Thr Thr Met Gly Glu His
 95                 100                 105                 110
```

```
aga aga gaa aaa aga tca gtg gca ctc gtt cca cat gtg gga atg gga     1287
Arg Arg Glu Lys Arg Ser Val Ala Leu Val Pro His Val Gly Met Gly
                115                 120                 125

ctg gag aca cga act gaa aca tgg atg tca tca gaa ggg gcc tgg aaa     1335
Leu Glu Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys
            130                 135                 140

cat gtc cag aga att gaa act tgg atc ttg aga cat cca ggc ttc acc     1383
His Val Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr
        145                 150                 155

atg atg gca gca atc ctg gca tac acc ata gga acg aca cat ttc caa     1431
Met Met Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln
    160                 165                 170

aga gcc ctg att ttc atc tta ctg aca gct gtc act cct tca atg aca     1479
Arg Ala Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr
175                 180                 185                 190

atg cgt tgc ata gga atg tca aat aga gac ttt gtg gaa ggg gtt tca     1527
Met Arg Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser
                195                 200                 205

gga gga agc tgg gtt gac ata gtc tta gaa cat gga agc tgt gtg acg     1575
Gly Gly Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr
            210                 215                 220

acg atg gca aaa aac aaa cca aca ttg gat ttt gaa ctg ata aaa aca     1623
Thr Met Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr
        225                 230                 235

gaa gcc aaa cag cct gcc acc cta agg aag tac tgt ata gag gca aag     1671
Glu Ala Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys
    240                 245                 250

cta acc aac aca aca aca gaa tct cgc tgc cca aca caa ggg gaa ccc     1719
Leu Thr Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro
255                 260                 265                 270

agc cta aat gaa gag cag gac aaa agg ttc gtc tgc aaa cac tcc atg     1767
Ser Leu Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met
                275                 280                 285

gta gac aga gga tgg gga aat gga tgt gga cta ttt gga aag gga ggc     1815
Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly
            290                 295                 300

att gtg acc tgt gct atg ttc aga tgc aaa aag aac atg gaa gga aaa     1863
Ile Val Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys
        305                 310                 315

gtt gtg caa cca gaa aac ttg gaa tac acc att gtg ata aca cct cac     1911
Val Val Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His
    320                 325                 330

tca ggg gaa gag cat gca gtc gga aat gac aca gga aaa cat ggc aag     1959
Ser Gly Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys
335                 340                 345                 350
```

```
gaa atc aaa ata aca cca cag agt tcc atc aca gaa gca gaa ttg aca      2007
Glu Ile Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr
                355                 360                 365

ggt tat ggc act gtc aca atg gag tgc tct cca aga acg ggc ctc gac      2055
Gly Tyr Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp
            370                 375                 380

ttc aat gag atg gtg ttg ttg cag atg gaa aat aaa gct tgg ctg gtg      2103
Phe Asn Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val
        385                 390                 395

cac agg caa tgg ttc cta gac ctg ccg tta cca tgg ttg ccc gga gcg      2151
His Arg Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala
    400                 405                 410

gac aca caa ggg tca aat tgg ata cag aaa gag aca ttg gtc act ttc      2199
Asp Thr Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe
415                 420                 425                 430

aaa aat ccc cat gcg aag aaa cag gat gtt gtt gtt tta gga tcc caa      2247
Lys Asn Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln
                435                 440                 445

gaa ggg gcc atg cac aca gca ctt aca ggg gcc aca gaa atc caa atg      2295
Glu Gly Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met
            450                 455                 460

tca tca gga aac tta ctc ttc aca gga cat ctc aag tgc agg ctg aga      2343
Ser Ser Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg
        465                 470                 475

atg gac aag cta cag ctc aaa gga atg tca tac tct atg tgc aca gga      2391
Met Asp Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly
    480                 485                 490

aag ttt aaa gtt gtg aag gaa ata gca gaa aca caa cat gga aca ata      2439
Lys Phe Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile
495                 500                 505                 510

gtt atc aga gtg caa tat gaa ggg gac ggc tct cca tgc aag atc cct      2487
Val Ile Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro
                515                 520                 525

ttt gag ata atg gat ttg gaa aaa aga cat gtc tta ggt cgc ctg att      2535
Phe Glu Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile
            530                 535                 540

aca gtc aac cca att gtg aca gaa aaa gat agc cca gtc aac ata gaa      2583
Thr Val Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu
        545                 550                 555

gca gaa cct cca ttc gga gac agc tac atc atc ata gga gta gag ccg      2631
Ala Glu Pro Pro Phe Gly Asp Ser Tyr Ile Ile Ile Gly Val Glu Pro
    560                 565                 570

gga caa ctg aag ctc aac tgg ttt aag aaa gga agt tct atc ggc caa      2679
Gly Gln Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Ser Ile Gly Gln
575                 580                 585                 590
```

```
atg ttt gag aca aca atg agg ggg gcg aag aga atg gcc att tta ggt     2727
Met Phe Glu Thr Thr Met Arg Gly Ala Lys Arg Met Ala Ile Leu Gly
                595                 600                 605

gac aca gcc tgg gat ttt gga tcc ttg gga gga gtg ttt aca tct ata     2775
Asp Thr Ala Trp Asp Phe Gly Ser Leu Gly Gly Val Phe Thr Ser Ile
            610                 615                 620

gga aag gct ctc cac caa gtc ttt gga gca atc tat gga get gcc ttc     2823
Gly Lys Ala Leu His Gln Val Phe Gly Ala Ile Tyr Gly Ala Ala Phe
        625                 630                 635

agt ggg gtt tca tgg act atg aaa atc ctc ata gga gtc att atc aca     2871
Ser Gly Val Ser Trp Thr Met Lys Ile Leu Ile Gly Val Ile Ile Thr
    640                 645                 650

tgg ata gga atg aat tca cgc agc gcc tca ctg tct gtg aca cta gta     2919
Trp Ile Gly Met Asn Ser Arg Ser Thr Ser Leu Ser Val Thr Leu Val
655                 660                 665                 670

ttg gtg gga att gtg aca ctg tat ttg gga gtc atg gtg cag gcc         2964
Leu Val Gly Ile Val Thr Leu Tyr Leu Gly Val Met Val Gln Ala
                675                 680                 685

taattagttg agcggccgct cgagcatgca tctagagggc cctatttcat agtgtcacct   3024
aaatgctaga gctcgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt   3084
ttgcccctcc cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcctttccta   3144
ataaaatgag gaaattgcat cgcattgtct gagtaggtgt cattctattc tgggggggtgg   3204
ggtggggcag gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc   3264
ggtgggctct atggcttctg aggcggaaag aaccagctgc attaatgaat cggccaacgc   3324
gcggggagag gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg   3384
cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta   3444
tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc   3504
aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc cctgacgag    3564
catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac   3624
caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc   3684
ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt   3744
aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc    3804
gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga   3864
cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta   3924
ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag aagaacagta   3984
tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga   4044
tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg   4104
cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag   4164
tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc   4224
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact   4284
tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt   4344
cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta   4404
ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta   4464
tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc   4524
gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat   4584
agtttgcgca acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt   4644
atggcttcat tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg   4704
tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca   4764
gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta   4824
agatgctttt ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg   4884
cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact   4944
ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg   5004
ctgttgagat ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt   5064
actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaagggа   5124
ataagggcga cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc   5184
```

atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa 5244
caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgtc 5292

<210> 43
<211> 685
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 43

```
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
 1               5                  10                  15
Val Val Ile Ala Cys Ala Gly Ala Phe His Leu Thr Thr Arg Asn Gly
            20                  25                  30
Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly Lys Ser Leu Leu
        35                  40                  45
Phe Lys Thr Glu Asp Gly Val Asn Met Cys Thr Leu Met Ala Met Asp
    50                  55                  60
Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys Cys Pro Leu Leu
65                  70                  75                  80
Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Ser Thr Ser
                85                  90                  95
Thr Trp Val Thr Tyr Gly Thr Cys Thr Thr Met Gly Glu His Arg Arg
            100                 105                 110
Glu Lys Arg Ser Val Ala Leu Val Pro His Val Gly Met Gly Leu Glu
        115                 120                 125
Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys His Val
    130                 135                 140
Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr Met Met
145                 150                 155                 160
Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln Arg Ala
                165                 170                 175
Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr Met Arg
            180                 185                 190
Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser Gly Gly
        195                 200                 205
Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr Thr Met
    210                 215                 220
Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr Glu Ala
225                 230                 235                 240
Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys Leu Thr
                245                 250                 255
Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro Ser Leu
            260                 265                 270
Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met Val Asp
        275                 280                 285
Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly Ile Val
    290                 295                 300
Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys Val Val
305                 310                 315                 320
Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His Ser Gly
                325                 330                 335
Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys Glu Ile
            340                 345                 350
Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr
        355                 360                 365
Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn
    370                 375                 380
```

```
Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val His Arg
385                 390                 395                 400
Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala Asp Thr
                405                 410                 415
Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe Lys Asn
            420                 425                 430
Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln Glu Gly
        435                 440                 445
Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met Ser Ser
    450                 455                 460
Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg Met Asp
465                 470                 475                 480
Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly Lys Phe
                485                 490                 495
Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile Val Ile
            500                 505                 510
Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu
        515                 520                 525
Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile Thr Val
    530                 535                 540
Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu Ala Glu
545                 550                 555                 560
Pro Pro Phe Gly Asp Ser Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln
                565                 570                 575
Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Ser Ile Gly Gln Met Phe
            580                 585                 590
Glu Thr Thr Met Arg Gly Ala Lys Arg Met Ala Ile Leu Gly Asp Thr
        595                 600                 605
Ala Trp Asp Phe Gly Ser Leu Gly Gly Val Phe Thr Ser Ile Gly Lys
    610                 615                 620
Ala Leu His Gln Val Phe Gly Ala Ile Tyr Gly Ala Ala Phe Ser Gly
625                 630                 635                 640
Val Ser Trp Thr Met Lys Ile Leu Ile Gly Val Ile Ile Thr Trp Ile
                645                 650                 655
Gly Met Asn Ser Arg Ser Thr Ser Leu Ser Val Thr Leu Val Leu Val
            660                 665                 670
Gly Ile Val Thr Leu Tyr Leu Gly Val Met Val Gln Ala
        675                 680                 685
```

<210> 44
<211> 5293
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<221> CDS
<222> (910)...(2964)

<400> 44

```
gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg     60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg    120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc    180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt    240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata    300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc    360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc    420
attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt    480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt    540
```

```
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca    600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg    660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc    720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg    780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca    840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc    900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc    951
          Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser
           1               5                  10

ttg gca gtt gtc ata gct tgt gca ggc gcc ttc cat tta acc aca cgt      999
Leu Ala Val Val Ile Ala Cys Ala Gly Ala Phe His Leu Thr Thr Arg
 15                  20                  25                  30

aac gga gaa cca cac atg atc gtc agc aga caa gag aaa ggg aaa agt     1047
Asn Gly Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly Lys Ser
                35                  40                  45

ctt ctg ttt aaa aca gag gat ggc gtg aac atg tgt acc ctc atg gcc     1095
Leu Leu Phe Lys Thr Glu Asp Gly Val Asn Met Cys Thr Leu Met Ala
            50                  55                  60

atg gac ctt ggt gaa ttg tgt gaa gac aca atc acg tac aag tgt ccc     1143
Met Asp Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys Cys Pro
        65                  70                  75

ctt ctc agg cag aat gag cca gaa gac ata gac tgt tgg tgc aac tct     1191
Leu Leu Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Ser
    80                  85                  90

acg tcc acg tgg gta act tat ggg acg tgt acc acc atg gga gaa cat     1239
Thr Ser Thr Trp Val Thr Tyr Gly Thr Cys Thr Thr Met Gly Glu His
 95                  100                 105                 110

aga aga gaa aaa aga tca gtg gca ctc gtt cca cat gtg gga atg gga     1287
Arg Arg Glu Lys Arg Ser Val Ala Leu Val Pro His Val Gly Met Gly
                115                 120                 125

ctg gag aca cga act gaa aca tgg atg tca tca gaa ggg gcc tgg aaa     1335
Leu Glu Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys
            130                 135                 140

cat gtc cag aga att gaa act tgg atc ttg aga cat cca ggc ttc acc     1383
His Val Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr
        145                 150                 155

atg atg gca gca atc ctg gca tac acc ata gga acg aca cat ttc caa     1431
Met Met Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln
    160                 165                 170

aga gcc ctg att ttc atc tta ctg aca gct gtc act cct tca atg aca     1479
Arg Ala Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr
175                 180                 185                 190

atg cgt tgc ata gga atg tca aat aga gac ttt gtg gaa ggg gtt tca     1527
Met Arg Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser
                195                 200                 205

gga gga agc tgg gtt gac ata gtc tta gaa cat gga agc tgt gtg acg     1575
Gly Gly Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr
            210                 215                 220
```

```
acg atg gca aaa aac aaa cca aca ttg gat ttt gaa ctg ata aaa aca      1623
Thr Met Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr
        225                 230                 235

gaa gcc aaa cag cct gcc acc cta agg aag tac tgt ata gag gca aag      1671
Glu Ala Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys
    240                 245                 250

cta acc aac aca aca aca gaa tct cgc tgc cca aca caa ggg gaa ccc      1719
Leu Thr Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro
255                 260                 265                 270

agc cta aat gaa gag cag gac aaa agg ttc gtc tgc aaa cac tcc atg      1767
Ser Leu Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met
                275                 280                 285

gta gac aga gga tgg gga aat gga tgt gga cta ttt gga aag gga ggc      1815
Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly
            290                 295                 300

att gtg acc tgt gct atg ttc aga tgc aaa aag aac atg gaa gga aaa      1863
Ile Val Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys
        305                 310                 315

gtt gtg caa cca gaa aac ttg gaa tac acc att gtg ata aca cct cac      1911
Val Val Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His
    320                 325                 330

tca ggg gaa gag cat gca gtc gga aat gac aca gga aaa cat ggc aag      1959
Ser Gly Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys
335                 340                 345                 350

gaa atc aaa ata aca cca cag agt tcc atc aca gaa gca gaa ttg aca      2007
Glu Ile Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr
                355                 360                 365

ggt tat ggc act gtc aca atg gag tgc tct cca aga acg ggc ctc gac      2055
Gly Tyr Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp
            370                 375                 380

ttc aat gag atg gtg ttg ttg cag atg gaa aat aaa gct tgg ctg gtg      2103
Phe Asn Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val
        385                 390                 395

cac agg caa tgg ttc cta gac ctg ccg tta cca tgg ttg ccc gga gcg      2151
His Arg Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala
    400                 405                 410

gac aca caa ggg tca aat tgg ata cag aaa gag aca ttg gtc act ttc      2199
Asp Thr Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe
415                 420                 425                 430

aaa aat ccc cat gcg aag aaa cag gat gtt gtt gtt tta gga tcc caa      2247
Lys Asn Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln
                435                 440                 445

gaa ggg gcc atg cac aca gca ctt aca ggg gcc aca gaa atc caa atg      2295
Glu Gly Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met
            450                 455                 460
```

```
tca tca gga aac tta ctc ttc aca gga cat ctc aag tgc agg ctg aga     2343
Ser Ser Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg
        465                 470                 475

atg gac aag cta cag ctc aaa gga atg tca tac tct atg tgc aca gga     2391
Met Asp Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly
    480                 485                 490

aag ttt aaa gtt gtg aag gaa ata gca gaa aca caa cat gga aca ata     2439
Lys Phe Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile
495                 500                 505                 510

gtt atc aga gtg caa tat gaa ggg gac ggc tct cca tgc aag atc cct     2487
Val Ile Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro
                515                 520                 525

ttt gag ata atg gat ttg gaa aaa aga cat gtc tta ggt cgc ctg att     2535
Phe Glu Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile
            530                 535                 540

aca gtc aac cca att gtg aca gaa aaa gat agc cca gtc aac ata gaa     2583
Thr Val Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu
        545                 550                 555

gca gaa cct cca ttc gga gac agc cac atc atc ata gga gta gag ccg     2631
Ala Glu Pro Pro Phe Gly Asp Ser His Ile Ile Ile Gly Val Glu Pro
    560                 565                 570

gga caa ctg aag ctc aac tgg ttt aag aaa gga agt tct atc ggc caa     2679
Gly Gln Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Ser Ile Gly Gln
575                 580                 585                 590

atg ttt gag aca aca atg agg ggg gcg aag aga atg gcc att tta ggt     2727
Met Phe Glu Thr Thr Met Arg Gly Ala Lys Arg Met Ala Ile Leu Gly
                595                 600                 605

gac aca gcc tgg gat ttt gga tcc ttg gga gga gtg ttt aca tct ata     2775
Asp Thr Ala Trp Asp Phe Gly Ser Leu Gly Gly Val Phe Thr Ser Ile
            610                 615                 620

gga aag gct ctc cac caa gtg ttt ggt ggt gcc ttc aga aca ctc ttt     2823
Gly Lys Ala Leu His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe
        625                 630                 635

ggg gga atg tct tgg atc aca caa ggg cta atg ggt gcc cta ctg ctc     2871
Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu
    640                 645                 650

tgg atg ggc gtc aac gca cga gac cga tca att gct ttg gcc ttc tta     2919
Trp Met Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu
655                 660                 665                 670

gcc aca ggg ggt gtg ctc gtg ttc tta gcg acc aat gtg cat gct          2964
Ala Thr Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala
                675                 680                 685

taattagttt gggcggccgc tcgagcatgc atctagaggg ccctattcta tagtgtcacc   3024
taaatgctag agctcgctga tcagcctcga ctgtgccttc tagttgccag ccatctgttg   3084
tttgcccctc ccccgtgcct tccttgaccc tggaaggtgc cactcccact gtcctttcct   3144
ataaaaatga ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg   3204
gggtggggca ggacagcaag ggggaggatt gggaagacaa tagcaggcat gctggggatg   3264
```

```
cggtgggctc tatggcttct gaggcggaaa gaaccagctg cattaatgaa tcggccaacg   3324
cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca ctgactcgct   3384
gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt   3444
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc   3504
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga   3564
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata   3624
ccaggcgttt cccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac   3684
cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg   3744
taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc   3804
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag   3864
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt   3924
aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta gaagaacagt   3984
atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg   4044
atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac   4104
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca   4164
gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac   4224
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac   4284
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt   4344
tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac gggagggctt   4404
accatctggc cccagtgctg caatgatacc gcgagaccca cgctcaccgg ctccagattt   4464
atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc   4524
cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa   4584
tagtttgcgc aacgttgttg ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg   4644
tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat cccccatgtt   4704
gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc   4764
agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt   4824
aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg   4884
gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac   4944
tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc   5004
gctgttgaga tccagttcga tgtaacccac tcgtgcaccc aactgatctt cagcatcttt   5064
tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg   5124
aataagggcg acacggaaat gttgaatact catactcttc cttttcaat attattgaag   5184
catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa   5244
acaaatagggg gttccgcgca catttccccg aaaagtgcca cctgacgtc              5293
```

<210> 45
<211> 685
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 45

```
Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser Leu Ala
 1               5                  10                  15
Val Val Ile Ala Cys Ala Gly Ala Phe His Leu Thr Thr Arg Asn Gly
            20                  25                  30
Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly Lys Ser Leu Leu
        35                  40                  45
Phe Lys Thr Glu Asp Gly Val Asn Met Cys Thr Leu Met Ala Met Asp
    50                  55                  60
Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys Cys Pro Leu Leu
65                  70                  75                  80
Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Ser Thr Ser
                85                  90                  95
Thr Trp Val Thr Tyr Gly Thr Cys Thr Thr Met Gly Glu His Arg Arg
            100                 105                 110
Glu Lys Arg Ser Val Ala Leu Val Pro His Val Gly Met Gly Leu Glu
        115                 120                 125
```

Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys His Val
130 135 140
Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr Met Met
145 150 155 160
Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln Arg Ala
165 170 175
Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr Met Arg
180 185 190
Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser Gly Gly
195 200 205
Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr Thr Met
210 215 220
Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr Glu Ala
225 230 235 240
Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys Leu Thr
245 250 255
Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro Ser Leu
260 265 270
Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met Val Asp
275 280 285
Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly Ile Val
290 295 300
Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys Val Val
305 310 315 320
Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His Ser Gly
325 330 335
Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys Glu Ile
340 345 350
Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr
355 360 365
Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn
370 375 380
Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val His Arg
385 390 395 400
Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala Asp Thr
405 410 415
Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe Lys Asn
420 425 430
Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln Glu Gly
435 440 445
Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met Ser Ser
450 455 460
Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg Met Asp
465 470 475 480
Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly Lys Phe
485 490 495
Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile Val Ile
500 505 510
Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu
515 520 525
Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile Thr Val
530 535 540
Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu Ala Glu
545 550 555 560
Pro Pro Phe Gly Asp Ser His Ile Ile Ile Gly Val Glu Pro Gly Gln
565 570 575
Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Ser Ile Gly Gln Met Phe
580 585 590
Glu Thr Thr Met Arg Gly Ala Lys Arg Met Ala Ile Leu Gly Asp Thr
595 600 605

```
Ala Trp Asp Phe Gly Ser Leu Gly Gly Val Phe Thr Ser Ile Gly Lys
    610                 615                 620
Ala Leu His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly
625                 630                 635                 640
Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met
                645                 650                 655
Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr
            660                 665                 670
Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala
        675                 680                 685
```

<210> 46
<211> 5293
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note = synthetic construct

<221> CDS
<222> (910)...(2964)

<400> 46

```
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg      60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg     120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc     180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt     240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata     300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc     360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc     420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt     480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt     540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca     600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg     660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc     720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg     780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca     840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gcttggtacc     900
gccgccgcc atg ggc aag agg tcc gcc ggc tca atc atg tgg ctc gcg agc      951
          Met Gly Lys Arg Ser Ala Gly Ser Ile Met Trp Leu Ala Ser
            1               5                   10

ttg gca gtt gtc ata gct tgt gca ggc gcc ttc cat tta acc aca cgt       999
Leu Ala Val Val Ile Ala Cys Ala Gly Ala Phe His Leu Thr Thr Arg
 15                  20                  25                  30

aac gga gaa cca cac atg atc gtc agc aga caa gag aaa ggg aaa agt      1047
Asn Gly Glu Pro His Met Ile Val Ser Arg Gln Glu Lys Gly Lys Ser
                35                  40                  45

ctt ctg ttt aaa aca gag gat ggc gtg aac atg tgt acc ctc atg gcc      1095
Leu Leu Phe Lys Thr Glu Asp Gly Val Asn Met Cys Thr Leu Met Ala
            50                  55                  60

atg gac ctt ggt gaa ttg tgt gaa gac aca atc acg tac aag tgt ccc      1143
Met Asp Leu Gly Glu Leu Cys Glu Asp Thr Ile Thr Tyr Lys Cys Pro
        65                  70                  75
```

```
ctt ctc agg cag aat gag cca gaa gac ata gac tgt tgg tgc aac tct     1191
Leu Leu Arg Gln Asn Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Ser
     80                  85                  90

acg tcc acg tgg gta act tat ggg acg tgt acc acc atg gga gaa cat     1239
Thr Ser Thr Trp Val Thr Tyr Gly Thr Cys Thr Thr Met Gly Glu His
 95                 100                 105                 110

aga aga gaa aaa aga tca gtg gca ctc gtt cca cat gtg gga atg gga     1287
Arg Arg Glu Lys Arg Ser Val Ala Leu Val Pro His Val Gly Met Gly
                115                 120                 125

ctg gag aca cga act gaa aca tgg atg tca tca gaa ggg gcc tgg aaa     1335
Leu Glu Thr Arg Thr Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys
            130                 135                 140

cat gtc cag aga att gaa act tgg atc ttg aga cat cca ggc ttc acc     1383
His Val Gln Arg Ile Glu Thr Trp Ile Leu Arg His Pro Gly Phe Thr
        145                 150                 155

atg atg gca gca atc ctg gca tac acc ata gga acg aca cat ttc caa     1431
Met Met Ala Ala Ile Leu Ala Tyr Thr Ile Gly Thr Thr His Phe Gln
    160                 165                 170

aga gcc ctg att ttc atc tta ctg aca gct gtc act cct tca atg aca     1479
Arg Ala Leu Ile Phe Ile Leu Leu Thr Ala Val Thr Pro Ser Met Thr
175                 180                 185                 190

atg cgt tgc ata gga atg tca aat aga gac ttt gtg gaa ggg gtt tca     1527
Met Arg Cys Ile Gly Met Ser Asn Arg Asp Phe Val Glu Gly Val Ser
                195                 200                 205

gga gga agc tgg gtt gac ata gtc tta gaa cat ggg agc tgt gtg acg     1575
Gly Gly Ser Trp Val Asp Ile Val Leu Glu His Gly Ser Cys Val Thr
            210                 215                 220

acg atg gca aaa aac aaa cca aca ttg gat ttt gaa ctg ata aaa aca     1623
Thr Met Ala Lys Asn Lys Pro Thr Leu Asp Phe Glu Leu Ile Lys Thr
        225                 230                 235

gaa gcc aaa cag cct gcc acc cta agg aag tac tgt ata gag gca aag     1671
Glu Ala Lys Gln Pro Ala Thr Leu Arg Lys Tyr Cys Ile Glu Ala Lys
    240                 245                 250

cta acc aac aca aca aca gaa tct cgc tgc cca aca caa ggg gaa ccc     1719
Leu Thr Asn Thr Thr Thr Glu Ser Arg Cys Pro Thr Gln Gly Glu Pro
255                 260                 265                 270

agc cta aat gaa gag cag gac aaa agg ttc gtc tgc aaa cac tcc atg     1767
Ser Leu Asn Glu Glu Gln Asp Lys Arg Phe Val Cys Lys His Ser Met
                275                 280                 285

gta gac aga gga tgg gga aat gga tgt gga cta ttt gga aag gga ggc     1815
Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly
            290                 295                 300

att gtg acc tgt gct atg ttc aga tgc aaa aag aac atg gaa gga aaa     1863
Ile Val Thr Cys Ala Met Phe Arg Cys Lys Lys Asn Met Glu Gly Lys
        305                 310                 315
```

```
gtt gtg caa cca gaa aac ttg gaa tac acc att gtg ata aca cct cac      1911
Val Val Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His
    320                 325                 330

tca ggg gaa gag cat gca gtc gga aat gac aca gga aaa cat ggc aag      1959
Ser Gly Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys
335                 340                 345                 350

gaa atc aaa ata aca cca cag agt tcc atc aca gaa gca gaa ttg aca      2007
Glu Ile Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr
                355                 360                 365

ggt tat ggc act gtc aca atg gag tgc tct cca aga acg ggc ctc gac      2055
Gly Tyr Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp
            370                 375                 380

ttc aat gag atg gtg ttg ttg cag atg gaa aat aaa gct tgg ctg gtg      2103
Phe Asn Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val
        385                 390                 395

cac agg caa tgg ttc cta gac ctg ccg tta cca tgg ttg ccc gga gcg      2151
His Arg Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala
    400                 405                 410

gac aca caa ggg tca aat tgg ata cag aaa gag aca ttg gtc act ttc      2199
Asp Thr Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe
415                 420                 425                 430

aaa aat ccc cat gcg aag aaa cag gat gtt gtt gtt tta gga tcc caa      2247
Lys Asn Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln
                435                 440                 445

gaa ggg gcc atg cac aca gca ctt aca ggg gcc aca gaa atc caa atg      2295
Glu Gly Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met
            450                 455                 460

tca tca gga aac tta ctc ttc aca gga cat ctc aag tgc agg ctg aga      2343
Ser Ser Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg
        465                 470                 475

atg gac aag cta cag ctc aaa gga atg tca tac tct atg tgc aca gga      2391
Met Asp Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly
    480                 485                 490

aag ttt aaa gtt gtg aag gaa ata gca gaa aca caa cat gga aca ata      2439
Lys Phe Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile
495                 500                 505                 510

gtt atc aga gtg caa tat gaa gug gac ggc tct cca tgc aag atc cct      2487
Val Ile Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro
                515                 520                 525

ttt gag ata atg gat ttg gaa aaa aga cat gtc tta ggt cgc ctg att      2535
Phe Glu Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile
            530                 535                 540

aca gtc aac cca att gtg aca gaa aaa gat agc cca gtc aac ata gaa      2583
Thr Val Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu
        545                 550                 555
```

```
gca gaa cct cca ttc gga gac agc tac atc atc ata gga gta gag ccg     2631
Ala Glu Pro Pro Phe Gly Asp Ser Tyr Ile Ile Ile Gly Val Glu Pro
    560                 565                 570

gga caa ctg aag ctc aac tgg ttt aag aaa gga agc acg ctg ggc aag     2679
Gly Gln Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Thr Leu Gly Lys
575                 580                 585                 590

gcc ttt tca aca act ttg aag gga gct caa aga ctg gca gcg ttg ggc     2727
Ala Phe Ser Thr Thr Leu Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly
                595                 600                 605

gac aca gcc tgg gac ttt ggc tct att gga ggg gtc ttc aac tcc ata     2775
Asp Thr Ala Trp Asp Phe Gly Ser Ile Gly Gly Val Phe Asn Ser Ile
            610                 615                 620

gga aaa gcc gtt cac caa gtg ttt ggt ggt gcc ttc aga aca ctc ttt     2823
Gly Lys Ala Val His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe
        625                 630                 635

ggg gga atg tct tgg atc aca caa ggg cta atg ggt gcc cta ctg ctc     2871
Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu
    640                 645                 650

tgg atg ggc gtc aac gca cga gac cga tca att gct ttg gcc ttc tta     2919
Trp Met Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu
655                 660                 665                 670

gcc aca ggg ggt gtg ctc gtg ttc tta gcg acc aat gtg cat gct         2964
Ala Thr Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala
                675                 680                 685

taattagttt gagcggccgc tcgagcatgc atctagaggg ccctattcta tagtgtcacc   3024
taaatgctag agctcgctga tcagcctcga ctgtgccttc tagttgccag ccatctgttg   3084
tttgcccctc ccccgtgcct tccttgaccc tggaaggtgc cactcccact gtcctttcct   3144
aataaaatga ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg  3204
gggtggggca ggacagcaag ggggaggatt gggaagacaa tagcaggcat gctggggatg   3264
cggtgggctc tatggcttct gaggcggaaa gaaccagctg cattaatgaa tcggccaacg   3324
cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca ctgactcgct   3384
gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt   3444
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc   3504
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga   3564
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata   3624
ccaggcgttt cccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac   3684
cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcaat gctcacgctg   3744
taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc   3804
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag   3864
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt   3924
aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta gaaggacagt   3984
atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg   4044
atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac   4104
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca   4164
gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac   4224
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac   4284
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt   4344
tcgttcatcc atagttgcct gactcccccgt cgtgtagata actacgatac gggagggctt   4404
accatctggc cccagtgctg caatgatacc gcgagaccca cgctcaccgg ctccagattt   4464
atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc   4524
cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa   4584
tagtttgcgc aacgttgttg ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg   4644
tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat cccccatgtt   4704
```

gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc 4764
agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt 4824
aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg 4884
gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac 4944
tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc 5004
gctgttgaga tccagttcga tgtaacccac tcgtgcaccc aactgatctt cagcatcttt 5064
tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg 5124
aataagggcg acacggaaat gttgaatact catactcttc cttttcaat attattgaag 5184
catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa 5244
acaaatagag gttccgcgca catttccccg aaaagtgcca cctgacgtc 5293

<210> 47
<211> 685
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 47

Met	Gly	Lys	Arg	Ser	Ala	Gly	Ser	Ile	Met	Trp	Leu	Ala	Ser	Leu	Ala
1				5				10					15		
Val	Val	Ile	Ala	Cys	Ala	Gly	Ala	Phe	His	Leu	Thr	Thr	Arg	Asn	Gly
			20					25					30		
Glu	Pro	His	Met	Ile	Val	Ser	Arg	Gln	Glu	Lys	Gly	Lys	Ser	Leu	Leu
		35					40					45			
Phe	Lys	Thr	Glu	Asp	Gly	Val	Asn	Met	Cys	Thr	Leu	Met	Ala	Met	Asp
	50					55					60				
Leu	Gly	Glu	Leu	Cys	Glu	Asp	Thr	Ile	Thr	Tyr	Lys	Cys	Pro	Leu	Leu
65					70					75					80
Arg	Gln	Asn	Glu	Pro	Glu	Asp	Ile	Asp	Cys	Trp	Cys	Asn	Ser	Thr	Ser
				85					90					95	
Thr	Trp	Val	Thr	Tyr	Gly	Thr	Cys	Thr	Thr	Met	Gly	Glu	His	Arg	Arg
			100					105					110		
Glu	Lys	Arg	Ser	Val	Ala	Leu	Val	Pro	His	Val	Gly	Met	Gly	Leu	Glu
		115					120					125			
Thr	Arg	Thr	Glu	Thr	Trp	Met	Ser	Ser	Glu	Gly	Ala	Trp	Lys	His	Val
	130					135					140				
Gln	Arg	Ile	Glu	Thr	Trp	Ile	Leu	Arg	His	Pro	Gly	Phe	Thr	Met	Met
145					150					155					160
Ala	Ala	Ile	Leu	Ala	Tyr	Thr	Ile	Gly	Thr	Thr	His	Phe	Gln	Arg	Ala
				165					170					175	
Leu	Ile	Phe	Ile	Leu	Leu	Thr	Ala	Val	Thr	Pro	Ser	Met	Thr	Met	Arg
			180					185					190		
Cys	Ile	Gly	Met	Ser	Asn	Arg	Asp	Phe	Val	Glu	Gly	Val	Ser	Gly	Gly
		195					200					205			
Ser	Trp	Val	Asp	Ile	Val	Leu	Glu	His	Gly	Ser	Cys	Val	Thr	Thr	Met
	210					215					220				
Ala	Lys	Asn	Lys	Pro	Thr	Leu	Asp	Phe	Glu	Leu	Ile	Lys	Thr	Glu	Ala
225					230					235					240
Lys	Gln	Pro	Ala	Thr	Leu	Arg	Lys	Tyr	Cys	Ile	Glu	Ala	Lys	Leu	Thr
				245					250					255	
Asn	Thr	Thr	Thr	Glu	Ser	Arg	Cys	Pro	Thr	Gln	Gly	Glu	Pro	Ser	Leu
			260					265					270		
Asn	Glu	Glu	Gln	Asp	Lys	Arg	Phe	Val	Cys	Lys	His	Ser	Met	Val	Asp
		275					280					285			
Arg	Gly	Trp	Gly	Asn	Gly	Cys	Gly	Leu	Phe	Gly	Lys	Gly	Gly	Ile	Val
	290					295					300				
Thr	Cys	Ala	Met	Phe	Arg	Cys	Lys	Lys	Asn	Met	Glu	Gly	Lys	Val	Val
305					310					315					320

```
Gln Pro Glu Asn Leu Glu Tyr Thr Ile Val Ile Thr Pro His Ser Gly
                325                 330                 335
Glu Glu His Ala Val Gly Asn Asp Thr Gly Lys His Gly Lys Glu Ile
            340                 345                 350
Lys Ile Thr Pro Gln Ser Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr
        355                 360                 365
Gly Thr Val Thr Met Glu Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn
    370                 375                 380
Glu Met Val Leu Leu Gln Met Glu Asn Lys Ala Trp Leu Val His Arg
385                 390                 395                 400
Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Leu Pro Gly Ala Asp Thr
                405                 410                 415
Gln Gly Ser Asn Trp Ile Gln Lys Glu Thr Leu Val Thr Phe Lys Asn
            420                 425                 430
Pro His Ala Lys Lys Gln Asp Val Val Val Leu Gly Ser Gln Glu Gly
        435                 440                 445
Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Met Ser Ser
    450                 455                 460
Gly Asn Leu Leu Phe Thr Gly His Leu Lys Cys Arg Leu Arg Met Asp
465                 470                 475                 480
Lys Leu Gln Leu Lys Gly Met Ser Tyr Ser Met Cys Thr Gly Lys Phe
                485                 490                 495
Lys Val Val Lys Glu Ile Ala Glu Thr Gln His Gly Thr Ile Val Ile
            500                 505                 510
Arg Val Gln Tyr Glu Gly Asp Gly Ser Pro Cys Lys Ile Pro Phe Glu
        515                 520                 525
Ile Met Asp Leu Glu Lys Arg His Val Leu Gly Arg Leu Ile Thr Val
    530                 535                 540
Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile Glu Ala Glu
545                 550                 555                 560
Pro Pro Phe Gly Asp Ser Tyr Ile Ile Ile Gly Val Glu Pro Gly Gln
                565                 570                 575
Leu Lys Leu Asn Trp Phe Lys Lys Gly Ser Thr Leu Gly Lys Ala Phe
            580                 585                 590
Ser Thr Thr Leu Lys Gly Ala Gln Arg Leu Ala Ala Leu Gly Asp Thr
        595                 600                 605
Ala Trp Asp Phe Gly Ser Ile Gly Gly Val Phe Asn Ser Ile Gly Lys
    610                 615                 620
Ala Val His Gln Val Phe Gly Gly Ala Phe Arg Thr Leu Phe Gly Gly
625                 630                 635                 640
Met Ser Trp Ile Thr Gln Gly Leu Met Gly Ala Leu Leu Leu Trp Met
                645                 650                 655
Gly Val Asn Ala Arg Asp Arg Ser Ile Ala Leu Ala Phe Leu Ala Thr
            660                 665                 670
Gly Gly Val Leu Val Phe Leu Ala Thr Asn Val His Ala
        675                 680                 685

<210> 48
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
      synthetic construct

<400> 48
tgtgcaggcg ccttccattt aaccacacgt aacg                                34

<210> 49
<211> 40
```

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 49
tcgagcggcc gctcaactaa ttaggcctgc accatgactc 40

<210> 50
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 50
cttatcgaaa ttaatacgac tcactatagg 30

<210> 51
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 51
atagattgct ccaaacactt ggtgg 25

<210> 52
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 52
actccatagg aaaagccgtt cacc 24

<210> 53
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 53
gcgagctcta gcatttaggt gacactatag 30

<210> 54
<211> 33
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note = synthetic construct

<400> 54
ctccaccaag tgtttggtgg tgccttcaga aca 33

<210> 55
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note = synthetic construct

<400> 55
Leu His Gln Val Phe Gly Gly Ala Phe Arg Thr
1 5 10

<210> 56
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note = synthetic construct

<400> 56
cttatcgaaa ttaatacgac tcactatagg 30

<210> 57
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note = synthetic construct

<400> 57
gaattcgtct cacttccttt cttaaaccag ttgagcttc 39

<210> 58
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note = synthetic construct

<400> 58
ggaattcgtc tcggaagcac gctgggcaag g 31

<210> 59
<211> 30
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 59
gcgagctcta gcatttaggt gacactatag 30

<210> 60
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 60
aactggttta agaaaggaag cacgctgggc gcc 33

<210> 61
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of artificial sequence; note =
synthetic construct

<400> 61
Asn Trp Lys Lys Gly Ser Thr Leu Gly Lys Ala
1 5 10

423

SUPERINDUSTRIA Y COMERCIO
Radicacion : 03097535 00000000 Folios: 462
Fecha (AMD): 2003-11-04 16:26:37
Tramite : 011 PCT-PATENT 379 FASE NACI 411 PRESENTACION
Dependencia: 2020 DIVISION DE NUEVAS CREACIONES

REQUISITOS NECESARIOS PARA P TRA

		USUARIO	RADICADOR
PATENTE DE INVENCIÓN	MODELO DE UTILIDAD	( )	( )
- Indicación de que se solicita la concesión de una patente.		( )	( )
- Datos de identificación del solicitante o de la persona que se presenta la solicitud.		( )	( )
- Descripción de la invención.		( )	( )
- Dibujos de ser estos pertinentes.		( )	( )
- Comprobante de Pago de las tasas establecidas.			
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 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 de trazado		( )	( )
Comprobante de pago de las tasas establecidas.		( )	( )
COMPLETA ( )	INCOMPLETA	( )	( )

Firma Responsable:

Eduar Nuevo

Fecha: 4-11-2003

Folio 424
Extracto Publicado

Industria y Comercio
SUPERINTENDENCIA

424

2020
Bogotá, D.C.,

Doctor:
Álvaro Correa Ordóñez
Apoderado
Government of the United States of America as represented by the Secretary of the Department of Health and Human Services, Centers for Disease Control and Prevention

ASUNTO:	Radicación N°	03-97535	
	Solicitud internacional	PCT/US2002/10764	
	Fecha inicio fase nacional	04/11/2003	
	Trámite	011	10662
	Evento	379	
	Actuación	430	
	Folios	001	

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:

- 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).

- La solicitud no contiene el poder debidamente otorgado, con el lleno de los requisitos exigidos por los artículos 63 y subsiguientes del Código de Procedimiento Civil. Adicionalmente, deberá presentar el documento que acredita la existencia y representación legal de la sociedad solicitante, cuando ésta fuere persona jurídica.

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 Única N° 10 de 2001.

ALIX CARMENZA CESPEDES DE VERGEL
Jefe de la División de Nuevas Creaciones.

05 DIC. 2003

El anterior requerimiento fue notificado por fijación en lista de fecha,_____.

202052