

VIS
Industria y Comercio
SUPERINTENDENCIA

Delegatura de propiedad Industrial
División de Nuevas Creaciones

SOLICITUD
PATENTE DE INVENCION

1. EXPEDIENTE No.

03-48706

4. TÍTULO

Proteínas de Infusión de GLP-1

CLASIFICACIÓN INTERNACIONAL

A61K 38/38 ; C07K
14/605

SOLICITANTE

ELI LILLY and Company

DOMICILIO

Indianapolis, IN 46285 Estados Unidos
de América.

APODERADO

Carlos D. Gorte

BOGOTÁ, D. C.,

junio 9 2003

PETITORIO

Industria y Comercio
SUPERINTENDENCIA

03 048706 00



SUPERINDUSTRIA Y COMERCIO

Radicación: 83048706 00000000

Folios: 429

Fecha (AMB): 2003-06-09 16:53:22

Trámite: 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

SOLICITUD DE :

☒

Patente de Invención

☐

Patente de Modelo de Utilidad

☐

Capítulo I.

☒

Capítulo II.

②

SOLICITANTE
(71)

Nombre: ELI LILLY AND COMPANY.

Dirección: Indianapolis, IN 46285 Estados Unidos de América.

Nacionalidad o Domicilio: ESTADOUNIDENSE

Lugar de Constitución: Indianapolis, Estados Unidos de América.

Teléfono:

Fax:

E-Mail:

IDENTIFICACION

C.C.

☐

NIT

☐

C.E.

☐

Otro

☐

Cual

Número

③

REPRESENTANTE
O APODERADO

Nombre: CARLOS R. OLARTE

Dirección: Avda 82 No. 10-62 Piso 6

Teléfono: 634-1500

Fax: 376-2211

E-Mail: patents.bogota@bakermct.com

IDENTIFICACION

C.C.

☒

NIT

☐

C.E.

☐

Otro

☐

Cual

Número 79.782.747 Bú TP 74.295

④

INVENTOR (ES)
(72)

Nombre: WOLFGANG GLAESNER, RADMILA MICANOVIC, SHENG-HUNG RAINBOW TSCHANG.

Dirección: 7512 Fielstone Court, Indianapolis, Indiana 46254, E.U.A.; 7126 White Oak Trail, Indianapolis, Indiana 46236, Estados Unidos de América; 4963 Riley Mews, Carmel, Indiana 46033, Estados Unidos de América.

Nacionalidad o Domicilio: ALEMANA; ESTADOUNIDENSE, ESTADOUNIDENSE

⑤ PCT (85)

Solicitud Internacional No. PCT/US01/43165

Fecha Noviembre 29, 2001

Publicación Internacional No. WO 02/46227

Fecha Junio 13, 2002

⑥ Título (54)

PROTEINAS DE INFUSION DE GLP-1

⑦

Clasificación Internacional (51) A61K 38/38 ; 38/26 ; C07K 14/605 , 14/765

⑧

Prioridad

SI

☒

NO

☐

(33) País de Origen

US

(32) Fecha

Diciembre 7, 2000

(31) Número de Solicitud

60/251,954

⑨

Comprobante de pago No. 33,799

Fecha (Mayo 15, 2003)

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

10

ANEXOS

2

- ☒ 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, al fuere el caso
- ☐ Documento de cesión del inventor al solicitante o a su causante
- ☐ Declaraciones
- ☒ Copia de la solicitud internacional. (Resumen, descripción, reivindicaciones, dibujos)
- ☒ 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: Búsqueda internacional.

11

FIGURA CARACTERISTICA

12

Solicito la concesión de la patente,

NOMBRE: CARLOS B. OLARTE

FIRMA:

C.C. 19.782.747 de Btá T.P. 74295

43

SUPERINDUSTRIA Y COMERCIO
Radicación : 83046706 00000000 Folios: 429
Fecha (HDD): 2003-05-09 16:53:52
Trámite : 011 PCT-PATENT 379 FASE NACI 411 PRESENTAC
Dependencia: 2020 DIVISION DE NUEVAS CREACIONES

NIT : 800176089-2

RECIBO OFICIAL DE CAJA : 03 - 33,799
FECHA : MAYO 15 DE 2003

***** CONSIGNACION *****

DEPOSITANTE	TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PGO	Vr. PAGO
JORGE LUIS BOMEZ	CONSIGNACION	BANCO POPULAR	050-00110-6	13227941	15/05/2003	25,300,000.00

***** CONCEPTO *****

CANT.	RENTISTICO	CONCEPTO	TOTAL CONCEPTO
1	50005-01-01 SOLICITUDES	1 TRAMITES DE SOL. DE PATENTE DE IN	400,000.00
		TOTAL :	\$ 400,000.00

SON: CUATROCIENTOS MIL PESOS

RESPONSABLE : 

RECIBO DE CAJA APLICADO AL EXPEDIENTE No. _____

SECRETARY OF STATE
STATE OF INDIANA



Todd Rokita
Secretary of State

APOSTILLE

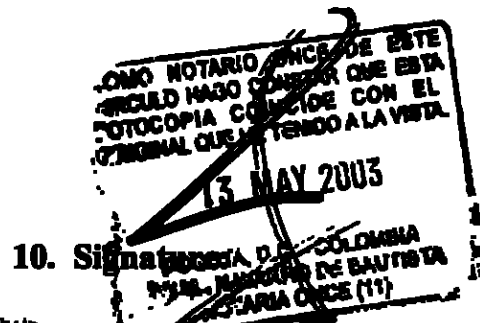
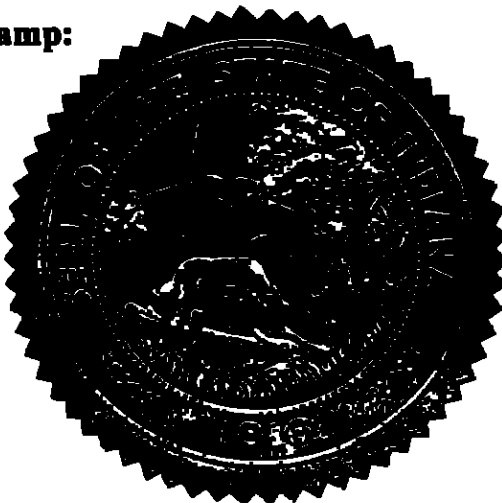
(Convention de la Haye du 5 Octobre 1961)

1. Country: United States of America
2. This public document has been signed by *Julie A. Worland*
3. acting in the capacity of Notary Public in & for *Morgan County*
4. and bears the seal/stamp of notary public in & for the State of Indiana

Certified

5. at Indianapolis, Indiana
6. this *Eighth* day of *January*, 2003
7. by Secretary of State of Indiana
8. No. A2002- 31613

9. Seal/Stamp:



10. Signature of Secretary of State

Todd Rokita

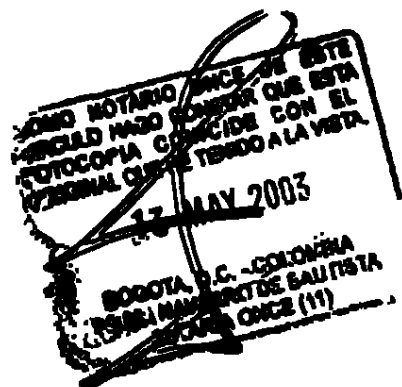
**Todd Rokita
Indiana Secretary of State**

5
7
IN WITNESS WHEREOF, this Power of Attorney has been signed by
Robert A. Armitage, Senior Vice President and General Counsel this 7th day of
January, 2003 at Indianapolis, Indiana USA.

SUBSCRIBED AND SWORN to be before me this 7th day of January,
2003.

Julie Ann Worland
NOTARY PUBLIC

JULIE ANN WORLAND
MORGAN COUNTY
MY COMMISSION EXPIRES:
JULY 22, 2009



6
8

PODER Y CERTIFICACIÓN NOTARIAL

(nota traductor: Han sido diligenciados y suscritos en la columna izquierda (inglés) que corresponde a la columna derecha en español)

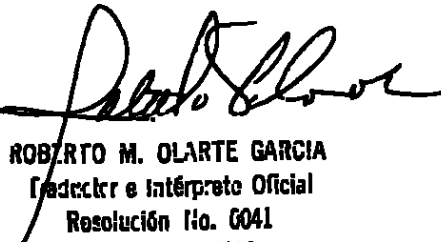
(nota traductor: punto 6- Acta del Directorio con fecha 29 de agosto de 1986 y Certificado de Constitución con fecha 17 de enero de 1901.)

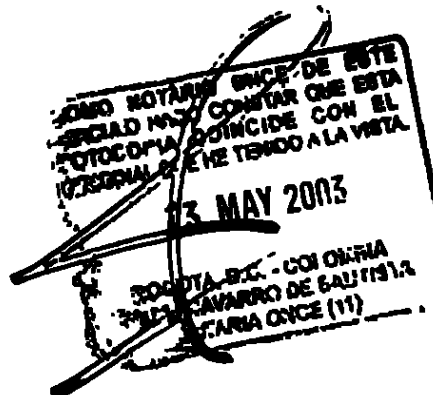
EN TESTIMONIO DE LO CUAL, este Poder ha sido suscrito por Robert A. Armitage, Vicepresidente Ejecutivo y Asesor Jurídico, el 7 de enero de 2003 en Indianápolis, Indiana, Estados Unidos de América.

SUSCRITO Y JURAMENTADO ante mí el 7 de enero de 2003.

(firmado: Julie Ann Worland)
NOTARIO PUBLICO

JULIE ANN WORLAND
CONDADO MORGAN
MI COMISION VENCE:
20 DE JULIO DE 2008


ROBERTO M. OLARTE GARCIA
Traductor e intérprete Oficial
Resolución No. 0041
Minijusticia 1996



79

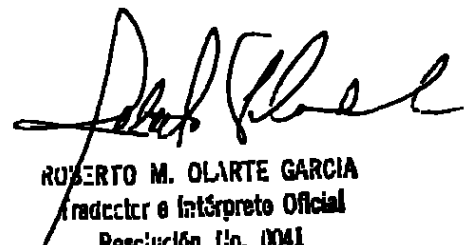
SECRETARÍA DE ESTADO
ESTADO DE INDIANA

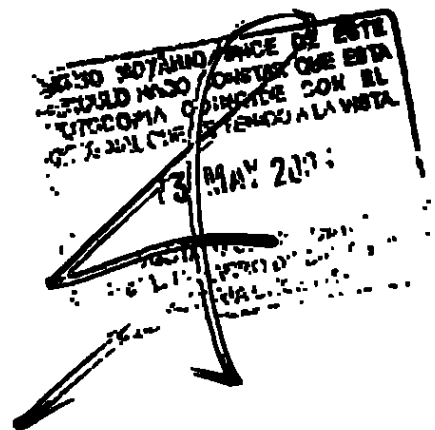
APOSTILLA
(Convención de la Haya del 5 de octubre de 1961)

1. País: ESTADOS UNIDOS DE AMERICA
2. Este documento público ha sido suscrito por: Julie A. Worland
3. en calidad de Notario Público para el Condado de Morgan
4. y lleva el sello de Notario Público del Estado de Indiana

Certificó

5. en la ciudad de Indianápolis, Indiana
6. el ocho de enero de 2003,
7. por el Secretario del Estado de Indiana
8. No. A2002-31613
9. Sello: Sello del Estado de Indiana - 1816
10. Firma: (Todd Rokita)
Todd Rokita
Secretario del Estado de Indiana


ROBERTO M. OLARTE GARCIA
Traductor e Intérprete Oficial
Resolución No. 0041
Minijusticia 1996



RAISBECK, LARA, RODRIGUEZ & RUEDA
MEMBERS OF THE FIRM
BAKER & MCKENZIE

ATTORNEYS AT LAW

AV 82 No. 10-62, 8TH FLOOR
BOGOTA, D.C. COLOMBIA
TELEPHONE (57-1) 634-1500
FACSIMILE (57-1) 376-2211
NIT 860.068.840-3
OFFICE.BOGOTA@BAKERNET.COM
WWW.BAKERNET.COM

8 to

POWER OF ATTORNEY

The undersigned.

(name of the company) **Eli Lilly and Company**

located in (full street address) **Lilly Corporate Center**

(city, state) **Indianapolis, Indiana U.S.A.**

declares that it hereby grants to

CARLOS REINALDO OLARTE AND/OR

ALVARO CORREA ORDOÑEZ

a full and sufficient power of attorney to apply to the proper national officers and authorities for the issue of:

Patents of Invention, Patents for Utility Models, Industrial Design Registrations, mergers, changes of name, assignments, filing and prosecution of oppositions, changes of address.

to which and it empowers them to take all steps necessary before said authorities for the objects stated, to file petitions, prepare descriptions, make protests, declarations, appeals and objections, pay imposts, apply for certifications, receive documents and securities, abandon applications and collect refunds. To the said mandatories sufficient power is granted hereby to receive notifications for and on behalf of the Company, to confer powers on those who are to represent the company judicially and extrajudicially, to answer in court in the matter of all claims and demands that may be presented in connection with the patents, and to do everything that may be required, before administrative or judicial officers of any class giving them likewise power to substitute these presents and revoke such substitutions if necessary.

Done and subscribed in

Indianapolis, Indiana U.S.A.

On the (day) **1st** day of (month) **January**

2003

(signature)

Robert A. Armitage
Senior Vice President
and General Counsel

RA

PODER

La abajo firmada.

domiciliada en _____

declara por el presente otorgar a

CARLOS REINALDO OLARTE Y/O

ALVARO CORREA ORDOÑEZ

poder especial amplio y bastante para recabar de las oficinas y autoridades nacionales que correspondan, la obtención de:

Patentes de invención, Patentes de Modelo de Utilidad, Registro de Diseños Industriales, fusiones, cambios de nombre, cesiones, presentación y trámite de oposiciones, cambios de dirección.

a cuyo efecto los faculta para dar ante dichas autoridades, todos los pasos necesarios al objeto indicado, elevar solicitudes, formular descripciones, protestas, declaraciones, apelaciones y reclamos, oblar impuestos, solicitar testimonios, recibir documentos y valores, deslitar y percibir. Se concede a los expresados mandatarios poder bastante para recibir notificaciones a nombre de la compañía, otorgar poderes a quienes vayan a representar a la compañía judicial y extrajudicialmente, responder en juicio a todas las reclamaciones o demandas que por motivo de las patentes se presentaren y hacer cuanto fuere necesario, ante las autoridades administrativas o judiciales de cualquier orden, dándose asimismo facultad para sustituir el presente y en caso necesario revocar dichas sustituciones.

Dado y firmado en

A los _____

días del mes de _____

2002

RECIBO DE ESTE PODER
OTORGADO POR LA COMPAÑIA
EL 1 DE ENERO DE 2003

AV 82 No. 10-62, 6TH FLOOR
BOGOTA, D.C. COLOMBIA
TELEPHONE (57-1) 834-1500
FACSIMILE (57-1) 376-2211
NIT 860.088.840-3
OFFICE.BOGOTA@BAKERNET.COM
WWW.BAKERNET.COM

119

NOTARIAL CERTIFICATION

CERTIFICACION NOTARIAL

On this 7th day of (month) January

A los _____ días del mes de _____

2003 the undersigned Notary Julie Ann Worland
(name of the notary)

de 2002, el suscrito Notario

CERTIFIES

DA FE

1. That (name of the officer) Robert A. Armitage

1. Que _____

of legal age and domiciled in (city, state) Indianapolis,
Indiana

mayor y domiciliado en _____

set his hand on the foregoing document.

suscribió de su puño y letra el documento que antecede.

2. That the grantor is to me personally know and that he/she has
legal capacity to execute the instrument.

2. Que conozco al otorgante y que éste tiene capacidad legal
para el otorgamiento.

3. That the grantor has executed the foregoing document as

3. Que el otorgante ha suscrito el referido documento en su
carácter de

(officer's title/position) Senior Vice President and
General Counsel

of the judicial person Eli Lilly and Company
(name of the company)

de la persona jurídica

4. That the judicial person Eli Lilly and Company

4. Que la persona jurídica

(name of the company)

having its home office in (full street address) Lilly Corporate
Center, Indianapolis, Indiana 46285 U.S.A.

con sede en

is duly organized, has present legal existence and that the purposes for
which the power of attorney is granted are within the scope of its corporate
purposes.

está legalmente constituida, que tiene existencia legal actual y que
el acto para el cual se ha otorgado el poder está comprendido entre
los que constituyen su objeto social.

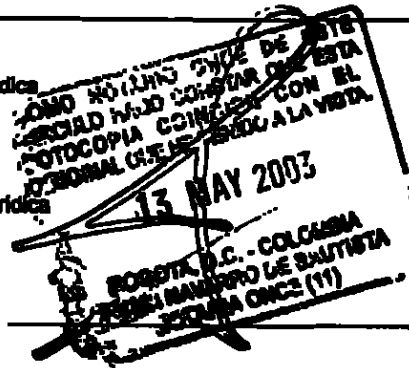
5. That the grantor has in fact the referred to authority and that his/her
representation is legal.

5. Que el otorgante tiene efectivamente la representación que
ostenta, y que esta es legítima.

6. That I have based certifications 3, 4 and 5 above on the following authentic
documents for such purpose exhibited to me and which I mention specifically,
giving their dates and their origin or source:

6. Que he basado las certificaciones 3, 4 y 5 anteriores en los siguientes
documentos auténticos que al efecto se me exhibieron y que
menciono específicamente con expresión de sus fechas y de su
origen o procedencia:

(provide name of documents and their dates) Board Meeting
Minutes dated August 29, 1986 and
Certificate of Incorporation dated
January 17, 1901.



(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
13 June 2002 (13.06.2002)

PCT

(10) International Publication Number
WO 02/46227 A2(51) International Patent Classification: C07K 14/605,
14/765, C12N 15/62, A61K 38/38, 38/26, C07K 19/00,
A61P 3/10

(21) International Application Number: PCT/US01/43165

(22) International Filing Date:
29 November 2001 (29.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/251,954 7 December 2000 (07.12.2000) US(71) Applicant (for all designated States except US): ELLI
LILLY AND COMPANY [US/US]; Lilly Corporate
Center, Indianapolis, IN 46285 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): GLAESNER,
Wolfgang [DE/US]; 7512 Fieldstone Court, Indianapolis,
IN 46254 (US). MICANOVIC, Radmila [US/US];
7126 White Oak Trail, Indianapolis, IN 46236 (US).
TSCHANG, Sheung-Hung, Rainbow [US/US]; 4963
Riley Mews, Carmel, IN 46033 (US).(74) Agents: STEWART, Mark, J. et al.; Eli Lilly and Com-
pany, Lilly Corporate Center, Drop Code 1104, Indianapo-
lis, IN 46285 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AT
(utility model), AU, AZ, BA, BB, BG, BR, BY, BZ, CA,
CH, CN, CO, CR, CU, CZ, CZ (utility model), DE, DK,
DK (utility model), DM, DZ, EC, EE, EE (utility model),
ES, FI, FI (utility model), GB, GD, GE, GH, GM, HR, HU,
ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, I.K., LR, I.S.,
LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO,
NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (utility
model), SL, TJ, TM, 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,
NB, SN, TD, TG).

Declarations under Rule 4.17:

— as to applicant's entitlement to apply for and be granted
a patent (Rule 4.17(II)) for the following designations 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, I.K., LR, LS, LT, LU, LV, MA, MD, MG, MK,
MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE,
SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU,
ZA, ZW, 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)— as to the applicant's entitlement to claim the priority of the
earlier application (Rule 4.17(III)) for the following desig-
nations 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, I.K., LR, LS, LT, LU, LV,
MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL,
PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ,
UA, UG, UZ, VN, YU, ZA, ZW, 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:

— without international search report and to be republished
upon receipt of that reportFor two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: GLP-1 FUSION PROTEINS

(57) Abstract: The present invention relates to glucagon-like-1 compounds fused to proteins that have the effect of extending the in
vivo half-life of the peptides. These fusion proteins can be used to treat non-insulin dependent diabetes mellitus as well as a variety
of other conditions.

WO 02/46227 A2

11

GLP-1 FUSION PROTEINS

The present invention relates to glucagon-like peptides including analogs and derivatives thereof fused to proteins that have the effect of extending the *in vivo* half-life of the peptides. These fusion proteins can be used to treat non-insulin dependent diabetes mellitus as well as a variety of other conditions.

Glucagon-Like Peptide 1 (GLP-1) is a 37 amino acid peptide that is secreted by the L-cells of the intestine in response to food ingestion. It has been found to stimulate insulin secretion (insulinotropic action), thereby causing glucose uptake by cells and decreased serum glucose levels [see, e.g., Mojsov, S., (1992) *Int. J. Peptide Protein Research*, 40:333-343]. However, GLP-1 is poorly active. A subsequent endogenous cleavage between the 6th and 7th position produces a more potent biologically active GLP-1(7-37)OH peptide. Numerous GLP-1 analogs and derivatives are known and are referred to herein as "GLP-1 compounds." These GLP-1 analogs include the Exendins which are peptides found in the venom of the GILA-monster. The Exendins have sequence homology to native GLP-1 and can bind the GLP-1 receptor and initiate the signal transduction cascade responsible for the numerous activities that have been attributed to GLP-1(7-37)OH.

GLP-1 compounds have a variety of physiologically significant activities. For example, GLP-1 has been shown to stimulate insulin release, lower glucagon secretion, inhibit gastric emptying, and enhance glucose utilization. [Nauck, M.A., et al. (1993) *Diabetologia* 36:741-744; Gutniak, M., et al. (1992) *New England J. of Med.* 326:1316-1322; Nauck, M.A., et al., (1993) *J. Clin. Invest.* 91:301-307].

12
15

- 2 -

GLP-1 shows the greatest promise as a treatment for non-insulin dependent diabetes mellitus (NIDDM). There are numerous oral drugs on the market to treat the insulin resistance associated with NIDDM. As the disease

5 progresses, however, patients must move to treatments that stimulate the release of insulin and eventually to treatments that involve injections of insulin. Current drugs which stimulate the release of insulin, however, can also cause hypoglycemia as can the actual administration of
10 insulin. GLP-1 activity, however, is controlled by blood glucose levels. When levels drop to a certain threshold level, GLP-1 is not active. Thus, there is no risk of hypoglycemia associated with treatment involving GLP-1.

However, the usefulness of therapy involving GLP-1
15 peptides has been limited by their fast clearance and short half-lives. For example, GLP-1(7-37) has a serum half-life of only 3 to 5 minutes. GLP-1(7-36) amide has a time action of about 50 minutes when administered subcutaneously. Even analogs and derivatives that are resistant to endogenous
20 protease cleavage, do not have half-lives long enough to avoid repeated administrations over a 24 hour period. Fast clearance of a therapeutic agent is inconvenient in cases where it is desired to maintain a high blood level of the agent over a prolonged period of time since repeated
25 administrations will then be necessary. Furthermore, a long-acting compound is particularly important for diabetic patients whose past treatment regimen has involved taking only oral medication. These patients often have an extremely difficult time transitioning to a regimen that
30 involves multiple injections of medication.

The present invention overcomes the problems associated with delivering a compound that has a short plasma half-life. The compounds of the present invention encompass GLP-1 compounds fused to another protein with a long circulating

13
H

- 3 -

half-life such as the Fc portion of an immunoglobulin or albumin.

Generally, small therapeutic peptides are difficult to manipulate because even slight changes in their structure can affect stability and/or biological activity. This has been especially true for GLP-1 compounds currently in development. For example, GLP-1(7-37)OH has a tendency to undergo a conformational change from a primarily alpha helix structure to a primarily beta sheet structure. This beta sheet form results in aggregated material that is thought to be inactive. It was, therefore, surprising that biologically active GLP-1 fusion proteins with increased half-lives could be developed. This was especially unexpected given the difficulty of working with GLP-1(7-37)OH alone and the large size of the fusion partner relative to the small GLP-1 peptide attached.

Compounds of the present invention include heterologous fusion proteins comprising a first polypeptide with a N-terminus and a C-terminus fused to a second polypeptide with a N-terminus and a C-terminus wherein the first polypeptide is a GLP-1 compound and the second polypeptide is selected from the group consisting of

- a) human albumin;
- b) human albumin analogs; and
- c) fragments of human albumin,

and wherein the C-terminus of the first polypeptide is fused to the N-terminus of the second polypeptide.

Compounds of the present invention also include a heterologous fusion protein comprising a first polypeptide with a N-terminus and a C-terminus fused to a second polypeptide with a N-terminus and a C-terminus wherein the first polypeptide is a GLP-1 compound and the second polypeptide is selected from the group consisting of

- a) human albumin;

- 4 -

17
14

- b) human albumin analogs; and
- c) fragments of human albumin,

and wherein the C-terminus of the first polypeptide is fused to the N-terminus of the second polypeptide via a peptide linker. It is preferred that the peptide linker is selected from the group consisting of:

- a) a glycine rich peptide;
- b) a peptide having the sequence [Gly-Gly-Gly-Gly-Ser]_n where n is 1, 2, 3, 4, 5 or 6; and
- 10 c) a peptide having the sequence [Gly-Gly-Gly-Gly-Ser]₃.

Additional compounds of the present invention include a heterologous fusion protein comprising a first polypeptide with a N-terminus and a C-terminus fused to a second polypeptide with a N-terminus and a C-terminus wherein the first polypeptide is a GLP-1 compound and the second polypeptide is selected from the group consisting of

- a) the Fc portion of an immunoglobulin;
 - b) an analog of the Fc portion of an immunoglobulin;
 - 20 and
 - c) fragments of the Fc portion of an immunoglobulin,
- and wherein the C-terminus of the first polypeptide is fused to the N-terminus of the second polypeptide. The GLP-1 compound may be fused to the second polypeptide via a peptide linker. It is preferable that the peptide linker is selected from the group consisting of:
- a) a glycine rich peptide;
 - b) a peptide having the sequence [Gly-Gly-Gly-Gly-Ser]_n where n is 1, 2, 3, 4, 5 or 6; and
 - 30 c) a peptide having the sequence [Gly-Gly-Gly-Gly-Ser]₃.

It is generally preferred that the GLP-1 compound that is part of the heterologous fusion protein have no more than 6 amino acids that are different from the corresponding amino acid in GLP-1(7-37)OH, GLP-1(7-36)OH, or Exendin-4.

35 It is even more preferred that the GLP-1 compound have no

18
15

- 5 -

more than 5 amino acids that differ from the corresponding amino acid in GLP-1(7-37)OH, GLP-1(7-36)OH, or Exendin-4.

It is most preferred that the GLP-1 compound have no more than 4, 3, or 2 amino acids that differ from the

5 corresponding amino acid in GLP-1(7-37)OH, GLP-1(7-36)OH, or Exendin-4. Preferably, a GLP-1 compound that is part of the heterologous fusion protein has glycine or valine at position 8.

The present invention also includes polynucleotides
10 encoding the heterologous fusion protein described herein, vectors comprising these polynucleotides and host cells transfected or transformed with the vectors described herein. Also included is a process for producing a
heterologous fusion protein comprising the steps of
15 transcribing and translating a polynucleotide described herein under conditions wherein the heterologous fusion protein is expressed in detectable amounts.

The present invention also encompasses a method for normalizing blood glucose levels in a mammal in need thereof
20 comprising the administration of a therapeutically effective amount of a heterologous fusion protein described herein.

The invention is further illustrated with reference to the following drawings:

25 Figure 1: IgG1 Fc amino acid sequence encompassing the hinge region, CH2 and CH3 domains.

Figure 2: Human serum albumin amino acid sequence.

Figure 3: A. SDS-PAGE gel and immunoblot of same gel illustrating the molecular weight of IgG1-Fc and GLP-1-Fc
30 fusion proteins (Lane 1, MW standards; Lane 2, Purified Fc; lane 3, Mock transfected media; Lane 4, Val⁸-GLP-1-Fc; Lane 5, Exendin-4-Fc) B. SDS-PAGE gel and immunoblot of same gel illustrating the molecular weight of human HSA and GLP-1-HSA
fusion proteins (Lane 1, MW standards; Lane 2, Purified HSA;
35 lane 3, Mock transfected media; Lane 4, Val⁸-GLP-1-HSA; Lane

16
19

- 6 -

5, Val⁸-GLP-1-[Gly-Gly-Gly-Gly-Ser]₃-HSA; Lane 6, Exendin-4-HSA; Lane 7, Exendin-4-[Gly-Gly-Gly-Gly-Ser]₃-HSA).

Figure 4: SDS-PAGE gel of purified Fc, albumin, and GLP-1 fusion proteins (Lane 1, MW standards; Lane 2, purified Fc; Lane 3, Val⁸-GLP-1-Fc; Lane 4, Exendin-4-Fc; Lane 5, MW standard; Lane 6, Val⁸-GLP-1-HSA; Lane 7, Exendin-4-HSA; Lane 8, Exendin-4-[Gly-Gly-Gly-Gly-Ser]₃-HSA).

Figure 5: Expression cloning vector containing the Fc regions illustrated in figure 1.

Figure 6: Expression cloning vector containing the albumin sequence illustrated in figure 2.

Figure 7: Expression cloning vector containing DNA encoding a 15 amino acid linker fused in frame and 5' of the albumin sequence illustrated in figure 2.

Figure 8: In vitro dose response activity of GLP-1 fusion proteins.

Figure 9: Pharmacokinetics of GLP-1 Fc and HSA fusion proteins.

Figure 10: Glucodynamic response to Exendin-Fc in two normal fasted dogs.

Figure 11: Insulinotropic response to Exendin-Fc in two normal fasted dogs.

Figure 12: DNA sequence encoding a human IgG1 Fc region.

Figure 13: DNA sequence encoding a human albumin protein.

The heterologous fusion proteins of the present invention comprise a GLP-1 compound fused to human albumin, a human albumin analog, a human albumin fragment, the Fc portion of an immunoglobulin, an analog of the Fc portion of an immunoglobulin, or a fragment of the Fc portion of an immunoglobulin. The C-terminus of the GLP-1 compound may be fused directly, or fused via a peptide linker, to the N-

20
17

terminus of an albumin or Fc protein. These heterologous fusion proteins are biologically active and have an increased half-life compared to native GLP-1.

It is preferred that the GLP-1 compounds that make up part of the heterologous fusion protein encompass polypeptides having from about twenty-five to about thirty-nine naturally occurring or non-naturally occurring amino acids that have sufficient homology to native GLP-1(7-37)OH such that they exhibit insulinotropic activity by binding to the GLP-1 receptor on β -cells in the pancreas. A GLP-1 compound typically comprises a polypeptide having the amino acid sequence of GLP-1(7-37)OH, an analog of GLP-1 (7-37)OH, a fragment of GLP-1(7-37)OH or a fragment of a GLP-1(7-37)OH analog. GLP-1(7-37)OH has the amino acid sequence of SEQ ID NO: 1:

7	8	9	10	11	12	13	14	15	16	17
His	Ala	Glu	Gly	Thr	Phe	Thr	Ser	Asp	Val	Ser
18	19	20	21	22	23	24	25	26	27	28
Ser	Tyr	Leu	Glu	Gly	Gln	Ala	Ala	Lys	Glu	Phe
29	30	31	32	33	34	35	36	37		
Ile	Ala	Trp	Leu	Val	Lys	Gly	Arg	Gly		

(SEQ ID NO: 1)

By custom in the art, the amino terminus of GLP-1(7-37)OH has been assigned number residue 7 and the carboxy-terminus, number 37. The other amino acids in the polypeptide are numbered consecutively, as shown in SEQ ID NO: 1. For example, position 12 is phenylalanine and position 22 is glycine.

GLP-1 compounds also encompass "GLP-1 fragments." A GLP-1 fragment is a polypeptide obtained after truncation of one or more amino acids from the N-terminus and/or C-terminus of GLP-1(7-37)OH or an analog or derivative thereof. The nomenclature used to describe GLP-1(7-37)OH is also applicable to GLP-1 fragments. For example, GLP-1(9-

21
18

- 8 -

36)OH denotes a GLP-1 fragment obtained by truncating two amino acids from the N-terminus and one amino acid from the C-terminus. The amino acids in the fragment are denoted by the same number as the corresponding amino acid in GLP-1(7-37)OH. For example, the N-terminal glutamic acid in GLP-1(9-36)OH is at position 9; position 12 is occupied by phenylalanine; and position 22 is occupied by glycine, as in GLP-1(7-37)OH. For GLP-1(7-36)OH, the glycine at position 37 of GLP-1(7-37)OH is deleted.

GLP-1 compounds also include polypeptides in which one or more amino acids have been added to the N-terminus and/ or C-terminus of GLP-1(7-37)OH, or fragments or analogs thereof. It is preferred that GLP-1 compounds of this type have up to about thirty-nine amino acids. The amino acids in the "extended" GLP-1 compound are denoted by the same number as the corresponding amino acid in GLP-1(7-37)OH. For example, the N-terminus amino acid of a GLP-1 compound obtained by adding two amino acids to the N-terminal of GLP-1(7-37)OH is at position 5; and the C-terminus amino acid of a GLP-1 compound obtained by adding one amino acid to the C-terminus of GLP-1(7-37)OH is at position 38. Thus, position 12 is occupied by phenylalanine and position 22 is occupied by glycine in both of these "extended" GLP-1 compounds, as in GLP-1(7-37)OH. Amino acids 1-6 of an extended GLP-1 compound are preferably the same as or a conservative substitution of the amino acid at the corresponding position of GLP-1(1-37)OH. Amino acids 38-45 of an extended GLP-1 compound are preferably the same as or a conservative substitution of the amino acid at the corresponding position of glucagon or Exendin-4.

GLP-1 compounds of the present invention encompass "GLP-1 analogs." A GLP-1 analog has sufficient homology to GLP-1(7-37)OH or a fragment of GLP-1(7-37)OH such that the analog has insulinotropic activity. Preferably, a GLP-1 analog has the amino acid sequence of GLP-1(7-37)OH or a

fragment thereof, modified so that from one, two, three, four or five amino acids differ from the amino acid in the corresponding position of GLP-1(7-37)OH or a fragment of GLP-1(7-37)OH. In the nomenclature used herein to

5 designate GLP-1 compounds, the substituting amino acid and its position is indicated prior to the parent structure. For example, Glu²²-GLP-1(7-37)OH designates a GLP-1 compound in which the glycine normally found at position 22 of GLP-1(7-37)OH has been replaced with glutamic acid; Val⁸-Glu²²-
10 GLP-1(7-37)OH designates a GLP-1 compound in which alanine normally found at position 8 and glycine normally found at position 22 of GLP-1(7-37)OH have been replaced with valine and glutamic acid, respectively.

GLP-1 compounds of the present invention also include
15 "GLP-1 derivatives." A GLP-1 derivative is defined as a molecule having the amino acid sequence of GLP-1 or of a GLP-1 analog, but additionally having chemical modification of one or more of its amino acid side groups, α -carbon atoms, terminal amino group, or terminal carboxylic acid
20 group. A chemical modification includes, but is not limited to, adding chemical moieties, creating new bonds, and removing chemical moieties. Modifications at amino acid side groups include, without limitation, acylation of lysine ϵ -amino groups, N-alkylation of arginine, histidine, or
25 lysine, alkylation of glutamic or aspartic carboxylic acid groups, and deamidation of glutamine or asparagine. Modifications of the terminal amino group include, without limitation, the des-amino, N-lower alkyl, N-di-lower alkyl, and N-acyl modifications. Modifications of the terminal
30 carboxy group include, without limitation, the amide, lower alkyl amide, dialkyl amide, and lower alkyl ester modifications. Lower alkyl is C₁-C₄ alkyl. Furthermore, one or more side groups, or terminal groups, may be protected by protective groups known to the ordinarily-

- 10 -

skilled protein chemist. The α -carbon of an amino acid may be mono- or dimethylated.

Any GLP-1 compound can be part of the heterologous fusion proteins of the present invention as long as the GLP-1 compound itself is able to bind and induce signaling through the GLP-1 receptor. GLP-1 receptor binding and signal transduction can be assessed using in vitro assays such as those described in EP 619,322 and U.S. Patent No. 5,120,712, respectively.

Numerous active GLP-1 fragments, analogs and derivatives are known in the art and any of these analogs and derivatives can also be part of the heterologous fusion proteins of the present invention. Some examples of novel GLP-1 analogs as well as GLP-1 analogs and derivatives known in the art are provided herein.

Some GLP-1 analogs and GLP-1 fragments known in the art include, for example, GLP-1(7-34) and GLP-1(7-35), GLP-1(7-36), Gln⁹-GLP-1(7-37), D-Gln⁹-GLP-1(7-37), Thr¹⁶-Lys¹⁸-GLP-1(7-37), and Lys¹⁸-GLP-1(7-37). GLP-1 analogs such as GLP-1(7-34) and GLP-1(7-35) are disclosed in U.S. Patent No. 5,118,666. Biologically processed forms of GLP-1 which have insulintropic properties, such as GLP-1(7-36) are also known. Other known biologically active GLP-1 compounds are disclosed in U.S. Patent No 5,977,071 to Hoffmann, et al.; U.S. Patent No. 5,545,618 to Buckley, et al., and Adelhorst, et al., *J. Biol. Chem.* 269:6275 (1994).

- 11 -

A preferred group of GLP-1 analogs is composed of GLP-1 analogs of formula I (SEQ ID NO: 2)

```

5           7   8   9   10  11  12  13  14  15  16  17
           His-Xaa-Xaa-Gly-Xaa-Phe-Thr-Xaa-Asp-Xaa-Xaa-
           18  19  20  21  22  23  24  25  26  27  28
           Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Phe-
           29  30  31  32  33  34  35  36  37  38  39
10          Ile-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-
           40  41  42  43  44  45
           Xaa-Xaa-Xaa-Xaa-Xaa-Xaa

```

Formula I (SEQ ID NO: 2)

15 wherein:

```

Xaa at position 8 is Ala, Gly, Ser, Thr, Leu, Ile, Val,
    Glu, Asp, or Lys;
Xaa at position 9 is Glu, Asp, or Lys;
Xaa at position 11 is Thr, Ala, Gly, Ser, Leu, Ile, Val,
20     Glu, Asp, or Lys;
Xaa at position 14 is Ser, Ala, Gly, Thr, Leu, Ile, Val,
    Glu, Asp, or Lys;
Xaa at position 16 is Val, Ala, Gly, Ser, Thr, Leu, Ile,
    Tyr, Glu, Asp, Trp, or Lys;
25 Xaa at position 17 is Ser, Ala, Gly, Thr, Leu, Ile, Val,
    Glu, Asp, or Lys;
Xaa at position 18 is Ser, Ala, Gly, Thr, Leu, Ile, Val,
    Glu, Asp, Trp, Tyr, or Lys;
Xaa at position 19 is Tyr, Phe, Trp, Glu, Asp, Gln, or Lys;
30 Xaa at position 20 is Leu, Ala, Gly, Ser, Thr, Ile, Val,
    Glu, Asp, Met, Trp, Tyr, or Lys;
Xaa at position 21 is Glu, Asp, or Lys;
Xaa at position 22 is Gly, Ala, Ser, Thr, Leu, Ile, Val,
    Glu, Asp, or Lys;
35 Xaa at position 23 is Gln, Asn, Arg, Glu, Asp, or Lys;

```

- 12 -

- Xaa at position 24 is Ala, Gly, Ser, Thr, Leu, Ile, Val,
Arg, Glu, Asp, or Lys;
- Xaa at position 25 is Ala, Gly, Ser, Thr, Leu, Ile, Val,
Glu, Asp, or Lys;
- 5 Xaa at position 26 is Lys, Arg, Gln, Glu, Asp, or His;
Xaa at position 27 is Leu, Glu, Asp, or Lys;
Xaa at position 30 is Ala, Gly, Ser, Thr, Leu, Ile, Val,
Glu, Asp, or Lys;
- Xaa at position 31 is Trp, Phe, Tyr, Glu, Asp, or Lys;
- 10 Xaa at position 32 is Leu, Gly, Ala, Ser, Thr, Ile, Val,
Glu, Asp, or Lys;
Xaa at position 33 is Val, Gly, Ala, Ser, Thr, Leu, Ile,
Glu, Asp, or Lys;
- Xaa at position 34 is Asn, Lys, Arg, Glu, Asp, or His;
- 15 Xaa at position 35 is Gly, Ala, Ser, Thr, Leu, Ile, Val,
Glu, Asp, or Lys;
Xaa at position 36 is Gly, Arg, Lys, Glu, Asp, or His;
Xaa at position 37 is Pro, Gly, Ala, Ser, Thr, Leu, Ile,
Val, Glu, Asp, or Lys, or is deleted;
- 20 Xaa at position 38 is Ser, Arg, Lys, Glu, Asp, or His, or is
deleted;
Xaa at position 39 is Ser, Arg, Lys, Glu, Asp, or His, or is
deleted;
- Xaa at position 40 is Gly, Asp, Glu, or Lys, or is deleted;
- 25 Xaa at position 41 is Ala, Phe, Trp, Tyr, Glu, Asp, or Lys,
or is deleted;
- Xaa at position 42 is Ser, Pro, Lys, Glu, or Asp, or is
deleted;
- Xaa at position 43 is Ser, Pro, Glu, Asp, or Lys, or is
30 deleted;
- Xaa at position 44 is Gly, Pro, Glu, Asp, or Lys, or is
deleted;
- and
- Xaa at position 45 is Ala, Ser, Val, Glu, Asp, or Lys, or is
35 deleted;

- 13 -

provided that when the amino acid at position 37, 38, 39, 40, 41, 42, 43, or 44 is deleted, then each amino acid downstream of that amino acid is also deleted.

It is preferred that the GLP-1 compound of formula I contain less than six amino acids that differ from the corresponding amino acid in GLP-1(7-37)OH or Exendin-4. It is more preferred that less than five amino acids differ from the corresponding amino acid in GLP-1(7-37)OH or Exendin-4. It is even more preferred that less than four amino acids differ from the corresponding amino acid in GLP-1(7-37)OH or Exendin-4.

GLP-1 compounds of the present invention include derivatives of formula I such as a C-1-6-ester, or amide, or C-1-6-alkylamide, or C-1-6-dialkylamide thereof. WO99/43706 describes derivatives of GLP-1 compounds of formula I and is incorporated by reference herein in its entirety. The compounds of formula I derivatized as described in WO99/43706 and underivatized are encompassed by the present invention.

Another preferred group of GLP-1 compounds is composed of GLP-1 analogs of formula II (SEQ ID NO: 3):

7	8	9	10	11	12	13	14	15	16	17
Xaa-Xaa-Xaa-Gly-Xaa-Xaa-Thr-Ser-Asp-Xaa-Ser-										
18	19	20	21	22	23	24	25	26	27	28
Xaa-Xaa-Leu-Glu-Gly-Xaa-Xaa-Ala-Xaa-Xaa-Phe-										
29	30	31	32	33	34	35	36	37		
Ile-Xaa-Xaa-Leu-Xaa-Xaa-Xaa-Xaa-R										

Formula II (SEQ ID NO: 3)

wherein:

Xaa at position 7 is: L-histidine, D-histidine, desamino-histidine, 2-amino-histidine, β -hydroxy-histidine, homohistidine, α -fluoromethyl-histidine or α -methyl-histidine;

- 14 -

Xaa at position 8 is: Gly, Ala, Val, Leu, Ile, Ser, or Thr;

Xaa at position 9 is: Thr, Ser, Arg, Lys, Trp, Phe, Tyr,

Glu, or His;

Xaa at position 11 is: Asp, Glu, Arg, Thr, Ala, Lys, or His;

5 Xaa at position 12 is: His, Trp, Phe, or Tyr;

Xaa at position 16 is: Leu, Ser, Thr, Trp, His, Phe, Asp,

Val, Tyr, Glu, or Ala;

Xaa at position 18 is: His, Pro, Asp, Glu, Arg, Ser, Ala, or

Lys;

10 Xaa at position 19 is: Gly, Asp, Glu, Gln, Asn, Lys, Arg, or
Cys;

Xaa at position 23 is: His, Asp, Lys, Glu, Gln, or Arg;

Xaa at position 24 is: Glu, Arg, Ala, or Lys;

Xaa at position 26 is: Trp, Tyr, Phe, Asp, Lys, Glu, or His;

15 Xaa at position 27 is: Ala, Glu, His, Phe, Tyr, Trp, Arg, or
Lys;

Xaa at position 30 is: Ala, Glu, Asp, Ser, or His;

Xaa at position 31 is: Asp, Glu, Ser, Thr, Arg, Trp, or Lys;

Xaa at position 33 is: Asp, Arg, Val, Lys, Ala, Gly, or Glu;

20 Xaa at position 34 is: Glu, Lys, or Asp;

Xaa at position 35 is: Thr, Ser, Lys, Arg, Trp, Tyr, Phe,
Asp, Gly, Pro, His, or Glu;

Xaa at position 36 is: Thr, Ser, Asp, Trp, Tyr, Phe, Arg,
Glu, or His;

25 R at position 37 is: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr,
Phe, His, Gly, Gly-Pro, or is deleted.

- 15 -

Another preferred group of GLP-1 compounds is composed of GLP-1 analogs of formula III (SEQ ID NO: 4):

5 7 8 9 10 11 12 13 14 15 16 17
 Xaa-Xaa-Glu-Gly-Xaa-Xaa-Thr-Ser-Asp-Xaa-Ser-
 18 19 20 21 22 23 24 25 26 27 28
 Ser-Tyr-Leu-Glu-Xaa-Xaa-Xaa-Xaa-Lys-Xaa-Phe-
 29 30 31 32 33 34 35 36 37
 Ile-Xaa-Trp-Leu-Xaa-Xaa-Xaa-Xaa-R
10 formula III (SEQ ID NO: 4)

wherein:

15 Xaa at position 7 is: L-histidine, D-histidine, desamino-histidine, 2-amino-histidine, β -hydroxy-histidine, homohistidine, α -fluoromethyl-histidine or α -methyl-histidine;

 Xaa at position 8 is: Gly, Ala, Val, Leu, Ile, Ser, or Thr;

 Xaa at position 11 is: Asp, Glu, Arg, Thr, Ala, Lys, or His;

 Xaa at position 12 is: His, Trp, Phe, or Tyr;

20 Xaa at position 16 is: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu, or Ala;

 Xaa at position 22: Gly, Asp, Glu, Gln, Asn, Lys, Arg, or Cys;

 Xaa at position 23 is: His, Asp, Lys, Glu, or Gln;

25 Xaa at position 24 is: Glu, His, Ala, or Lys;

 Xaa at position 25 is: Asp, Lys, Glu, or His;

 Xaa at position 27 is: Ala, Glu, His, Phe, Tyr, Trp, Arg, or Lys;

 Xaa at position 30 is: Ala, Glu, Asp, Ser, or His;

30 Xaa at position 33 is: Asp, Arg, Val, Lys, Ala, Gly, or Glu;

 Xaa at position 34 is: Glu, Lys, or Asp;

 Xaa at position 35 is: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, or Glu;

 Xaa at position 36 is: Arg, Glu, or His;

- 16 -

R at position 37 is: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, or is deleted.

Another preferred group of GLP-1 compounds is composed of GLP-1 analogs of formula IV (SEQ ID NO: 5):

```

      7   8   9   10  11  12  13  14  15  16  17
Xaa-Xaa-Glu-Gly-Thr-Xaa-Thr-Ser-Asp-Xaa-Ser-
18  19  20  21  22  23  24  25  26  27  28
Ser-Tyr-Leu-Glu-Xaa-Xaa-Ala-Ala-Xaa-Glu-Phe-
10  29  30  31  32  33  34  35  36  37
Ile-Xaa-Trp-Leu-Val-Lys-Xaa-Arg-R
      formula IV (SEQ ID NO: 5)

```

wherein:

15 Xaa at position 7 is: L-histidine, D-histidine, desamino-histidine, 2-amino-histidine, β -hydroxy-histidine, homohistidine, α -fluoromethyl-histidine or α -methyl-histidine;

20 Xaa at position 8 is: Gly, Ala, Val, Leu, Ile, Ser, Met, or Thr;

Xaa at position 12 is: His, Trp, Phe, or Tyr;

Xaa at position 16 is: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu, or Ala;

25 Xaa at position 22 is: Gly, Asp, Glu, Gln, Asn, Lys, Arg, or Cys;

Xaa at position 23 is: His, Asp, Lys, Glu, or Gln;

Xaa at position 26 is: Asp, Lys, Glu, or His;

Xaa at position 30 is: Ala, Glu, Asp, Ser, or His;

30 Xaa at position 35 is: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, or Glu;

R at position 37 is: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, or is deleted.

- 17 -

Another preferred group of GLP-1 compounds is composed of GLP-1 analogs of formula V (SEQ ID NO: 6):

5 7 8 9 10 11 12 13 14 15 16 17
Xaa-Xaa-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Val-Ser-
18 19 20 21 22 23 24 25 26 27 28
Ser-Tyr-Leu-Glu-Xaa-Xaa-Xaa-Ala-Lys-Glu-Phe-
29 30 31 32 33 34 35 36 37
Ile-Xaa-Trp-Leu-Val-Lys-Gly-Arg-R
10 formula V (SEQ ID NO: 6)

wherein:

Xaa at position 7 is: L-histidine, D-histidine, desamino-histidine, 2-amino-histidine, β -hydroxy-histidine, homohistidine, α -fluoromethyl-histidine or α -methyl-histidine;
15 Xaa at position 8 is: Gly, Ala, Val, Leu, Ile, Ser, or Thr;
Xaa at position 22 is: Gly, Asp, Glu, Gln, Asn, Lys, Arg, or Cys;
Xaa at position 23 is: His, Asp, Lys, Glu, or Gln;
20 Xaa at position 24 is: Ala, Glu, His, Phe, Tyr, Trp, Arg, or Lys;
Xaa at position 30 is: Ala, Glu, Asp, Ser, or His;
R at position 37 is: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, or is deleted.

25 Preferred GLP-1 compounds of formula I, II, III, IV, and V comprise GLP-1 analogs or fragments of GLP-1 analogs wherein the analogs or fragments contain an amino acid other than alanine at position 8 (position 8 analogs). It is
30 preferable that these position 8 analogs contain one or more additional changes at positions 9, 11, 12, 16, 18, 22, 23, 24, 26, 27, 30, 31, 33, 34, 35, 36, and 37 compared to the corresponding amino acid of native GLP-1(7-37)OH. It is also preferable that these analogs have 6 or fewer changes
35 compared to the corresponding amino acids in native GLP-1(7-

- 18 -

37)OH or GLP-1(7-36)OH. More preferred analogs have 5 or fewer changes compared to the corresponding amino acids in native GLP-1(7-37)OH or GLP-1(7-36)OH or have 4 or fewer changes compared to the corresponding amino acids in native GLP-1(7-37)OH or GLP-1(7-36)OH. It is even more preferable that these analogs have 3 or fewer changes compared to the corresponding amino acids in native GLP-1(7-37)OH or GLP-1(7-36)OH. It is most preferable that these analogs have 2 or fewer changes compared to the corresponding amino acids in native GLP-1(7-37)OH.

It has been found that the compounds of formula II, III, IV, and V have a reduced propensity to aggregate and generate insoluble forms. This is also important in the context of a fusion protein wherein the relatively small GLP-1 peptide must maintain an active conformation despite being fused to a much larger protein. Preferred GLP-1 compounds of formula II, III, IV, and V encompassed by the fusion proteins of the present invention comprise GLP-1 analogs or fragments of GLP-1 analogs in which glycine at position 22 and preferably alanine at position 8 have been replaced with another amino acid.

When position 22 is aspartic acid, glutamic acid, arginine or lysine, position 8 is preferably glycine, valine, leucine, isoleucine, serine, threonine or methionine and more preferably valine or glycine. When position 22 is a sulfonic acid such as cysteic acid, position 8 is preferably glycine, valine, leucine, isoleucine, serine, threonine or methionine and more preferably valine or glycine.

Other preferred GLP-1 compounds include GLP-1 analogs of formula IV (SEQ ID NO:5) wherein the analogs have the sequence of GLP-1(7-37)OH except that the amino acid at position 8 is preferably glycine, valine, leucine, isoleucine, serine, threonine, or methionine and more preferably valine or glycine and position 30 is glutamic

- 19 -

acid, aspartic acid, serine, or histidine and more preferably glutamatic acid.

Other preferred GLP-1 compounds include GLP-1 analogs of formula IV (SEQ ID NO:5) wherein the analogs have the sequence of GLP-1(7-37)OH except that the amino acid at position 8 is preferably glycine, valine, leucine, isoleucine, serine, threonine, or methionine and more preferably valine or glycine and position 37 is histidine, lysine, arginine, threonine, serine, glutamic acid, aspartic acid, tryptophan, tyrosine, phenylalanine and more preferably histidine.

Other preferred GLP-1 compounds include GLP-1 analogs of formula IV (SEQ ID NO:5) wherein the analogs have the sequence of GLP-1(7-37)OH except that the amino acid at position 8 is preferably glycine, valine, leucine, isoleucine, serine, threonine, or methionine and more preferably valine or glycine and position 22 is glutamic acid, lysine, aspartic acid, or arginine and more preferably glutamic acid or lysine and position 23 is lysine, arginine, glutamic acid, aspartic acid, and histidine and more preferably lysine or glutamic acid.

Other preferred GLP-1 compounds include GLP-1 analogs of formula V (SEQ ID NO:6) wherein the analogs have the sequence of GLP-1(7-37)OH except that the amino acid at position 8 is preferably glycine, valine, leucine, isoleucine, serine, threonine, or methionine and more preferably valine or glycine and position 22 is glutamic acid, lysine, aspartic acid, or arginine and more preferably glutamine acid or lysine and position 27 is alanine, lysine, arginine, tryptophan, tyrosine, phenylalanine, or histidine and more preferably alanine.

Other preferred GLP-1 compounds include GLP-1 analogs of formula II wherein the analogs have the sequence of GLP-

- 20 -

1(7-37)OH except that the amino acid at position 8 and one, two, or three amino acids selected from the group consisting of position 9, position 11, position 12, position 16, position 18, position 22, position 23, position 24, position 26, position 27, position 30, position 31, position 33, position 34, position 35, position 36, and position 37, differ from the amino acid at the corresponding position of native GLP-1(7-37)OH.

Other preferred GLP-1 compounds of formula II include:

10 Val⁸-GLP-1(7-37)OH, Gly⁸-GLP-1(7-37)OH, Glu²²-GLP-1(7-37)OH, Asp²²-GLP-1(7-37)OH, Arg²²-GLP-1(7-37)OH, Lys²²-GLP-1(7-37)OH, Cys²²-GLP-1(7-37)OH, Val⁸-Glu²²-GLP-1(7-37)OH, Val⁸-Asp²²-GLP-1(7-37)OH, Val⁸-Arg²²-GLP-1(7-37)OH, Val⁸-Lys²²-GLP-1(7-37)OH, Val⁸-Cys²²-GLP-1(7-37)OH, Gly⁸-Glu²²-GLP-1(7-37)OH, Gly⁸-Asp²²-GLP-1(7-37)OH, Gly⁸-Arg²²-GLP-1(7-37)OH, Gly⁸-Lys²²-GLP-1(7-37)OH, Gly⁸-Cys²²-GLP-1(7-37)OH, Glu²²-GLP-1(7-36)OH, Asp²²-GLP-1(7-36)OH, Arg²²-GLP-1(7-36)OH, Lys²²-GLP-1(7-36)OH, Cys²²-GLP-1(7-36)OH, Val⁸-Glu²²-GLP-1(7-36)OH, Val⁸-Asp²²-GLP-1(7-36)OH, Val⁸-Arg²²-GLP-1(7-36)OH, Val⁸-Lys²²-GLP-1(7-36)OH, Val⁸-Cys²²-GLP-1(7-36)OH, Gly⁸-Glu²²-GLP-1(7-36)OH, Gly⁸-Asp²²-GLP-1(7-36)OH, Gly⁸-Arg²²-GLP-1(7-36)OH, Gly⁸-Lys²²-GLP-1(7-36)OH, Gly⁸-Cys²²-GLP-1(7-36)OH, Lys²³-GLP-1(7-37)OH, Val⁸-Lys²³-GLP-1(7-37)OH, Gly⁸-Lys²³-GLP-1(7-37)OH, His²⁴-GLP-1(7-37)OH, Val⁸-His²⁴-GLP-1(7-37)OH, Gly⁸-His²⁴-GLP-1(7-37)OH, Lys²⁴-GLP-1(7-37)OH, Val⁸-Lys²⁴-GLP-1(7-37)OH, Gly⁸-Lys²³-GLP-1(7-37)OH, Glu³⁰-GLP-1(7-37)OH, Val⁸-Glu³⁰-GLP-1(7-37)OH, Gly⁸-Glu³⁰-GLP-1(7-37)OH, Asp³⁰-GLP-1(7-37)OH, Val⁸-Asp³⁰-GLP-1(7-37)OH, Gly⁸-Asp³⁰-GLP-1(7-37)OH, Gln³⁰-GLP-1(7-37)OH, Val⁸-Gln³⁰-GLP-1(7-37)OH, Gly⁸-Gln³⁰-GLP-1(7-37)OH, Tyr³⁰-

- 21 -

GLP-1(7-37)OH, Val⁸-Tyr³⁰-GLP-1(7-37)OH, Gly⁸-Tyr³⁰-GLP-1(7-37)OH, Ser³⁰-GLP-1(7-37)OH, Val⁸-Ser³⁰-GLP-1(7-37)OH, Gly⁸-Ser³⁰-GLP-1(7-37)OH, His³⁰-GLP-1(7-37)OH, Val⁸-His³⁰-GLP-1(7-37)OH, Gly⁸-His³⁰-GLP-1(7-37)OH, Glu³⁴-GLP-1(7-37)OH, Val⁸-Glu³⁴-GLP-1(7-37)OH, Gly⁸-Glu³⁴-GLP-1(7-37)OH, Ala³⁴-GLP-1(7-37)OH, Val⁸-Ala³⁴-GLP-1(7-37)OH, Gly⁸-Ala³⁴-GLP-1(7-37)OH, Gly³⁴-GLP-1(7-37)OH, Val⁸-Gly³⁴-GLP-1(7-37)OH, Gly⁸-Gly³⁴-GLP-1(7-37)OH, Ala³⁵-GLP-1(7-37)OH, Val⁸-Ala³⁵-GLP-1(7-37)OH, Gly⁸-Ala³⁵-GLP-1(7-37)OH, Lys³⁵-GLP-1(7-37)OH, Val⁸-Lys³⁵-GLP-1(7-37)OH, Gly⁸-Lys³⁵-GLP-1(7-37)OH, His³⁵-GLP-1(7-37)OH, Val⁸-His³⁵-GLP-1(7-37)OH, Gly⁸-His³⁵-GLP-1(7-37)OH, Pro³⁵-GLP-1(7-37)OH, Val⁸-Pro³⁵-GLP-1(7-37)OH, Gly⁸-Pro³⁵-GLP-1(7-37)OH, Glu³⁵-GLP-1(7-37)OH, Val⁸-Glu³⁵-GLP-1(7-37)OH, Gly⁸-Glu³⁵-GLP-1(7-37)OH, Val⁸-Ala²⁷-GLP-1(7-37)OH, Val⁸-His³⁷-GLP-1(7-37)OH, Val⁸-Glu²²-Lys²³-GLP-1(7-37)OH, Val⁸-Glu²²-Glu²³-GLP-1(7-37)OH, Val⁸-Glu²²-Ala²⁷-GLP-1(7-37)OH, Val⁸-Gly³⁴-Lys³⁵-GLP-1(7-37)OH, Val⁸-His³⁷-GLP-1(7-37)OH, Gly⁸-His³⁷-GLP-1(7-37)OH, Val⁸-Glu²²-Ala²⁷-GLP-1(7-37)OH, Gly⁸-Glu²²-Ala²⁷-GLP-1(7-37)OH, Val⁸-Lys²²-Glu²³-GLP-1(7-37)OH, and Gly⁸-Lys²²-Glu²³-GLP-1(7-37)OH.

Another preferred group of GLP-1 analogs and derivatives for use in the present invention is composed of molecules of formula VI (SEQ ID NO: 7)

25 R₁-X-Glu-Gly¹⁰-Thr-Phe-Thr-Ser-Asp¹⁵-Val-Ser-
Ser-Tyr-Leu²⁰-Y -Gly-Gln-Ala-Ala²⁵-Lys- Z -
Phe-Ile-Ala³⁰-Trp-Leu-Val-Lys-Gly³⁵-Arg-R₂
formula VI (SEQ ID NO:7)

- 22 -

wherein: R₁ is selected from the group consisting of L-histidine, D-histidine, desamino-histidine, 2-amino-histidine, β -hydroxy-histidine, homohistidine, alpha-fluoromethyl-histidine, and alpha-methyl-histidine; X is
5 selected from the group consisting of Ala, Gly, Val, Thr, Ile, and alpha-methyl-Ala; Y is selected from the group consisting of Glu, Gln, Ala, Thr, Ser, and Gly; Z is selected from the group consisting of Glu, Gln, Ala, Thr, Ser, and Gly; and R₂ is Gly-OH.

10 Another preferred group of GLP-1 compounds for use in the present invention is disclosed in WO 91/11457, and consists essentially of GLP-1(7-34), GLP-1(7-35), GLP-1(7-36), or GLP-1(7-37), or the amide form thereof, and pharmaceutically-acceptable salts thereof, having at least
15 one modification selected from the group consisting of:

(a) substitution of glycine, serine, cysteine, threonine, asparagine, glutamine, tyrosine, alanine, valine, isoleucine, leucine, methionine, phenylalanine, arginine, or D-lysine for lysine at position 26 and/or position 34; or
20 substitution of glycine, serine, cysteine, threonine, asparagine, glutamine, tyrosine, alanine, valine, isoleucine, leucine, methionine, phenylalanine, lysine, or a D-arginine for arginine at position 36;

(b) substitution of an oxidation-resistant amino
25 acid for tryptophan at position 31;

(c) substitution of at least one of: tyrosine for valine at position 16; lysine for serine at position 18; aspartic acid for glutamic acid at position 21; serine for glycine at position 22; arginine for glutamine at position
30 23; arginine for alanine at position 24; and glutamine for lysine at position 26; and

(d) substitution of at least one of: glycine, serine, or cysteine for alanine at position 8; aspartic acid, glycine, serine, cysteine, threonine, asparagine,
35 glutamine, tyrosine, alanine, valine, isoleucine, leucine,

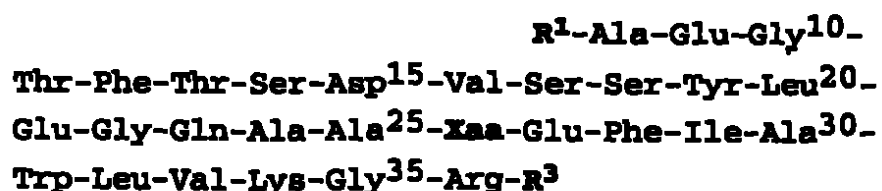
- 23 -

methionine, or phenylalanine for glutamic acid at position 9; serine, cysteine, threonine, asparagine, glutamine, tyrosine, alanine, valine, isoleucine, leucine, methionine, or phenylalanine for glycine at position 10; and glutamic acid for aspartic acid at position 15; and

(e) substitution of glycine, serine, cysteine, threonine, asparagine, glutamine, tyrosine, alanine, valine, isoleucine, leucine, methionine, or phenylalanine, or the D- or N-acylated or alkylated form of histidine for histidine at position 7; wherein, in the substitutions is (a), (b), (d), and (e), the substituted amino acids can optionally be in the D-form and the amino acids substituted at position 7 can optionally be in the N-acylated or N-alkylated form.

Because the enzyme, dipeptidyl-peptidase IV (DPP IV), may be responsible for the observed rapid in vivo inactivation of administered GLP-1, [see, e.g., Mentlein, R., et al., *Eur. J. Biochem.*, 214:829-835 (1993)], GLP-1 analogs and derivatives that are protected from the activity of DPP IV in the context of a fusion protein are preferred, and fusion proteins wherein the GLP-1 compound is Gly⁸-GLP-1(7-37)OH, Val⁸-GLP-1(7-37)OH, α-methyl-Ala⁸-GLP-1(7-37)OH, or Gly⁸-Gln²¹-GLP-1(7-37)OH are more preferred.

Another preferred group of GLP-1 compounds for use in the present invention consists of the compounds of formula VII (SEQ ID NO: 8) claimed in U.S. Patent No. 5,512,549, which is expressly incorporated herein by reference.



|
R²

formula VII (SEQ ID NO: 8)

- 24 -

wherein R^1 is selected from the group consisting of 4-imidazopropionyl, 4-imidazoacetyl, or 4-imidazo- α, α -dimethyl-acetyl; R^2 is selected from the group consisting of C_6 - C_{10} unbranched acyl, or is absent; R^3 is
5 selected from the group consisting of Gly-OH or NH_2 ; and, Xaa is Lys or Arg.

More preferred compounds of formula IV for use in the present invention are those in which Xaa is Arg and R^2 is C_6 - C_{10} unbranched acyl. Even more preferred compounds of
10 formula IV for use in the present invention are those in which Xaa is Arg, R^2 is C_6 - C_{10} unbranched acyl, and R^3 is Gly-OH. Other highly-preferred compounds of formula IV for use in the present invention are those in which Xaa is Arg, R^2 is C_6 - C_{10} unbranched acyl, R^3 is Gly-OH, and R^1 is 4-
15 imidazopropionyl. An especially preferred compound of formula IV for use in the present invention is that in which Xaa is Arg, R^2 is C_8 unbranched acyl, R^3 is Gly-OH, and R^1 is 4-imidazopropionyl.

Preferably, the GLP-1 compounds comprise GLP-1 analogs
20 wherein the backbone for such analogs or fragments contains an amino acid other than alanine at position 8 (position 8 analogs). The backbone may also include L-histidine, D-histidine, or modified forms of histidine such as desamino-histidine, 2-amino-histidine, β -hydroxy-histidine,
25 homohistidine, α -fluoromethyl-histidine, or α -methyl-histidine at position 7. It is preferable that these position 8 analogs contain one or more additional changes at positions 12, 16, 18, 19, 20, 22, 25, 27, 30, 33, and 37 compared to the corresponding amino acid of native GLP-1(7-
30 37)OH. It is more preferable that these position 8 analogs contain one or more additional changes at positions 16, 18,

- 25 -

22, 25 and 33 compared to the corresponding amino acid of native GLP-1(7-37)OH.

In a preferred embodiment, the GLP-1 analog is GLP-1(7-37)OH wherein the amino acid at position 12 is selected from the group consisting of tryptophan or tyrosine. It is more preferred that in addition to the substitution at position 12, the amino acid at position 8 is substituted with glycine, valine, leucine, isoleucine, serine, threonine, or methionine and more preferably valine or glycine. It is even more preferred that in addition to the substitutions at position 12 and 8, the amino acid at position 22 is substituted with glutamic acid.

In another preferred embodiment, the GLP-1 analog is GLP-1(7-37)OH wherein the amino acid at position 16 is selected from the group consisting of tryptophan, isoleucine, leucine, phenylalanine, or tyrosine. It is more preferred that in addition to the substitution at position 16, the amino acid at position 8 is substituted with glycine, valine, leucine, isoleucine, serine, threonine, or methionine and more preferably valine or glycine. It is even more preferred that in addition to the substitutions at position 16 and 8, the amino acid at position 22 is substituted with glutamic acid. It is also preferred that in addition to the substitutions at positions 16 and 8, the amino acid at position 30 is substituted with glutamic acid. It is also preferred that in addition to the substitutions at positions 16 and 8, the amino acid at position 37 is substituted with histidine.

In another preferred embodiment, the GLP-1 analog is GLP-1(7-37)OH wherein the amino acid at position 18 is selected from the group consisting of tryptophan, tyrosine, phenylalanine, lysine, leucine, or isoleucine, preferably tryptophan, tyrosine, and isoleucine. It is more preferred

- 26 -

that in addition to the substitution at position 18, the amino acid at position 8 is substituted with glycine, valine, leucine, isoleucine, serine, threonine, or methionine and more preferably valine or glycine. It is even more preferred that in addition to the substitutions at position 18 and 8, the amino acid at position 22 is substituted with glutamic acid. It is also preferred that in addition to the substitutions at positions 18 and 8, the amino acid at position 30 is substituted with glutamic acid. It is also preferred that in addition to the substitutions at positions 18 and 8, the amino acid at position 37 is substituted with histidine

In another preferred embodiment, the GLP-1 analog is GLP-1(7-37)OH wherein the amino acid at position 19 is selected from the group consisting of tryptophan or phenylalanine, preferably tryptophan. It is more preferred that in addition to the substitution at position 19, the amino acid at position 8 is substituted with glycine, valine, leucine, isoleucine, serine, threonine, or methionine and more preferably valine or glycine. It is even more preferred that in addition to the substitutions at position 19 and 8, the amino acid at position 22 is substituted with glutamic acid. It is also preferred that in addition to the substitutions at positions 19 and 8, the amino acid at position 30 is substituted with glutamic acid. It is also preferred that in addition to the substitutions at positions 19 and 8, the amino acid at position 37 is substituted with histidine

In another preferred embodiment, the GLP-1 analog is GLP-1(7-37)OH wherein the amino acid at position 20 is phenylalanine, tyrosine, or tryptophan. It is more preferred that in addition to the substitution at position 20, the amino acid at position 8 is substituted with

- 27 -

glycine, valine, leucine, isoleucine, serine, threonine, or methionine and more preferably valine or glycine. It is even more preferred that in addition to the substitutions at position 20 and 8, the amino acid at position 22 is

5 substituted with glutamic acid. It is also preferred that in addition to the substitutions at positions 20 and 8, the amino acid at position 30 is substituted with glutamic acid. It is also preferred that in addition to the substitutions at positions 20 and 8, the amino acid at position 37 is
10 substituted with histidine

In another preferred embodiment, the GLP-1 analog is GLP-1(7-37)OH wherein the amino acid at position 25 is selected from the group consisting of valine, isoleucine, and leucine, preferably valine. It is more preferred that
15 in addition to the substitution at position 25, the amino acid at position 8 is substituted with glycine, valine, leucine, isoleucine, serine, threonine, or methionine and more preferably valine or glycine. It is even more preferred that in addition to the substitutions at position
20 25 and 8, the amino acid at position 22 is substituted with glutamic acid. It is also preferred that in addition to the substitutions at positions 25 and 8, the amino acid at position 30 is substituted with glutamic acid. It is also preferred that in addition to the substitutions at positions
25 25 and 8, the amino acid at position 37 is substituted with histidine.

In another preferred embodiment, the GLP-1 analog is GLP-1(7-37)OH wherein the amino acid at position 27 is selected from the group consisting of isoleucine or alanine.
30 It is more preferred that in addition to the substitution at position 27, the amino acid at position 8 is substituted with glycine, valine, leucine, isoleucine, serine, threonine, or methionine and more preferably valine or

- 28 -

glycine. It is even more preferred that in addition to the substitutions at position 27 and 8, the amino acid at position 22 is substituted with glutamic acid. It is also preferred that in addition to the substitutions at positions 27 and 8, the amino acid at position 30 is substituted with glutamic acid. It is also preferred that in addition to the substitutions at positions 27 and 8, the amino acid at position 37 is substituted with histidine

In another preferred embodiment, the GLP-1 analog is GLP-1(7-37)OH wherein the amino acid at position 33 is isoleucine. It is more preferred that in addition to the substitution at position 33, the amino acid at position 8 is substituted with glycine, valine, leucine, isoleucine, serine, threonine, or methionine and more preferably valine or glycine. It is even more preferred that in addition to the substitutions at position 33 and 8, the amino acid at position 22 is substituted with glutamic acid. It is also preferred that in addition to the substitutions at positions 33 and 8, the amino acid at position 30 is substituted with glutamic acid. It is also preferred that in addition to the substitutions at positions 33 and 8, the amino acid at position 37 is substituted with histidine

The GLP-1 compounds have modifications at one or more of the following positions: 8, 12, 16, 18, 19, 20, 22, 25, 27, 30, 33, and 37. These GLP-1 compounds show increased potency compared with GLP-1(7-37)OH and comprise the amino acid sequence of formula IX (SEQ ID NO:12)

Xaa₇-Xaa₈-Glu-Gly-Thr-Xaa₁₂-Thr-Ser-Asp-Xaa₁₆-Ser-Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Gln-Ala-Xaa₂₅-Lys-Xaa₂₇-Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Lys-Gly-Arg-Xaa₃₇

Formula IX (SEQ ID NO: 12)

- 29 -

wherein:

Xaa₇ is: L-histidine, D-histidine, desamino-histidine, 2-amino-histidine, β -hydroxy-histidine,
 5 homohistidine, α -fluoromethyl-histidine, or α -methyl-histidine;

Xaa₈ is: Ala, Gly, Val, Leu, Ile, Ser, or Thr;

Xaa₁₂ is: Phe, Trp, or Tyr;

Xaa₁₆ is: Val, Trp, Ile, Leu, Phe, or Tyr;

10 Xaa₁₈ is: Ser, Trp, Tyr, Phe, Lys, Ile, Leu, Val;

Xaa₁₉ is: Tyr, Trp, or Phe;

Xaa₂₀ is: Leu, Phe, Tyr, or Trp;

Xaa₂₂ is: Gly, Glu, Asp, or Lys;

Xaa₂₅ is: Ala, Val, Ile, or Leu;

15 Xaa₂₇ is: Glu, Ile, or Ala;

Xaa₃₀ is: Ala or Glu

Xaa₃₃ is: Val, or Ile; and

Xaa₃₇ is: Gly, His, NH₂, or is absent.

20 Some preferred GLP-1 compounds of formula IX include
 GLP-1(7-37)OH, GLP-1(7-36)-NH₂, Gly⁸-GLP-1(7-37)OH, Gly⁸-GLP-
 1(7-36)NH₂, Val⁸-GLP-1(7-37)OH, Val⁸-GLP-1(7-36)NH₂, Leu⁸-
 GLP-1(7-37)OH, Leu⁸-GLP-1(7-36)NH₂, Ile⁸-GLP-1(7-37)OH, Ile⁸-
 GLP-1(7-36)NH₂, Ser⁸-GLP-1(7-37)OH, Ser⁸-GLP-1(7-36)NH₂,
 25 Thr⁸-GLP-1(7-37)OH, Thr⁸-GLP-1(7-36)NH₂, Val⁸-Tyr¹²-GLP-1(7-
 37)OH, Val⁸-Tyr¹²-GLP-1(7-36)NH₂, Val⁸-Tyr¹⁶-GLP-1(7-37)OH,
 Val⁸-Tyr¹⁶-GLP-1(7-36)NH₂, Val⁸-Glu²²-GLP-1(7-37)OH, Val⁸-
 Glu²²-GLP-1(7-36)NH₂, Gly⁸-Glu²²-GLP-1(7-37)OH, Gly⁸-Glu²²-GLP-
 1(7-36)NH₂, Val⁸-Asp²²-GLP-1(7-37)OH, Val⁸-Asp²²-GLP-1(7-
 30 36)NH₂, Gly⁸-Asp²²-GLP-1(7-37)OH, Gly⁸-Asp²²-GLP-1(7-36)NH₂,
 Val⁸-Lys²²-GLP-1(7-37)OH, Val⁸-Lys²²-GLP-1(7-36)NH₂, Gly⁸-
 Lys²²-GLP-1(7-37)OH, Gly⁸-Lys²²-GLP-1(7-36)NH₂, Leu⁸-Glu²²-GLP-
 1(7-37)OH, Leu⁸-Glu²²-GLP-1(7-36)NH₂, Ile⁸-Glu²²-GLP-1(7-

45
40

- 30 -

37)OH, Ile⁸-Glu²²-GLP-1(7-36)NH₂, Leu⁸-Asp²²-GLP-1(7-37)OH, Leu⁸-Asp²²-GLP-1(7-36)NH₂, Ile⁸-Asp²²-GLP-1(7-37)OH, Ile⁸-Asp²²-GLP-1(7-36)NH₂, Leu⁸-Lys²²-GLP-1(7-37)OH, Leu⁸-Lys²²-GLP-1(7-36)NH₂, Ile⁸-Lys²²-GLP-1(7-37)OH, Ile⁸-Lys²²-GLP-1(7-36)NH₂, Ser⁸-Glu²²-GLP-1(7-37)OH, Ser⁸-Glu²²-GLP-1(7-36)NH₂, Thr⁸-Glu²²-GLP-1(7-37)OH, Thr⁸-Glu²²-GLP-1(7-36)NH₂, Ser⁸-Asp²²-GLP-1(7-37)OH, Ser⁸-Asp²²-GLP-1(7-36)NH₂, Thr⁸-Asp²²-GLP-1(7-37)OH, Thr⁸-Asp²²-GLP-1(7-36)NH₂, Ser⁸-Lys²²-GLP-1(7-37)OH, Ser⁸-Lys²²-GLP-1(7-36)NH₂, Thr⁸-Lys²²-GLP-1(7-37)OH, Thr⁸-Lys²²-GLP-1(7-36)NH₂, Glu²²-GLP-1(7-37)OH, Glu²²-GLP-1(7-36)NH₂, Asp²²-GLP-1(7-37)OH, Asp²²-GLP-1(7-36)NH₂, Lys²²-GLP-1(7-37)OH, Lys²²-GLP-1(7-36)NH₂, Val⁸-Ala²⁷-GLP-1(7-37)OH, Val⁸-Glu²²-Ala²⁷-GLP-1(7-37)OH, Val⁸-Glu³⁰-GLP-1(7-37)OH, Val⁸-Glu³⁰-GLP-1(7-36)NH₂, Gly⁸-Glu³⁰-GLP-1(7-37)OH, Gly⁸-Glu³⁰-GLP-1(7-36)NH₂, Leu⁸-Glu³⁰-GLP-1(7-37)OH, Leu⁸-Glu³⁰-GLP-1(7-36)NH₂, Ile⁸-Glu³⁰-GLP-1(7-37)OH, Ile⁸-Glu³⁰-GLP-1(7-36)NH₂, Ser⁸-Glu³⁰-GLP-1(7-37)OH, Ser⁸-Glu³⁰-GLP-1(7-36)NH₂, Thr⁸-Glu³⁰-GLP-1(7-37)OH, Thr⁸-Glu³⁰-GLP-1(7-36)NH₂, Val⁸-His³⁷-GLP-1(7-37)OH, Val⁸-His³⁷-GLP-1(7-36)NH₂, Gly⁸-His³⁷-GLP-1(7-37)OH, Gly⁸-His³⁷-GLP-1(7-36)NH₂, Leu⁸-His³⁷-GLP-1(7-37)OH, Leu⁸-His³⁷-GLP-1(7-36)NH₂, Ile⁸-His³⁷-GLP-1(7-37)OH, Ile⁸-His³⁷-GLP-1(7-36)NH₂, Ser⁸-His³⁷-GLP-1(7-37)OH, Ser⁸-His³⁷-GLP-1(7-36)NH₂, Thr⁸-His³⁷-GLP-1(7-37)OH, Thr⁸-His³⁷-GLP-1(7-36)NH₂.

Some preferred GLP-1 compounds of formula IX having multiple substitutions include GLP-1(7-37)OH wherein position 8 is valine or glycine, position 22 is glutamic acid, position 16 is tyrosine, leucine or tryptophan, position 18 is tyrosine, tryptophan, or isoleucine, position 25 is valine and position 33 is isoleucine. Other preferred GLP-1 compounds include the following: Val⁸-Tyr¹⁶-GLP-1(7-37)OH, Val⁸-Tyr¹²-Glu²²-GLP-1(7-37)OH, Val⁸-Tyr¹⁶-Phe¹⁹-GLP-1(7-37)OH, Val⁸-Tyr¹⁶-Glu²²-GLP-1(7-37)OH, Val⁸-Trp¹⁶-Glu²²-GLP-1(7-37)OH, Val⁸-Leu¹⁶-Glu²²-GLP-1(7-37)OH, Val⁸-Ile¹⁶-

49
41

- 31 -

Glu²²-GLP-1(7-37)OH, Val⁸-Phe¹⁶-Glu²²-GLP-1(7-37)OH; Val⁸-Trp¹⁸-Glu²²-GLP-1(7-37)OH, Val⁸-Tyr¹⁸-Glu²²-GLP-1(7-37)OH, Val⁸-Phe¹⁸-Glu²²-GLP-1(7-37)OH, and Val⁸-Ile¹⁸-Glu²²-GLP-1(7-37)OH.

5 The GLP-1 compounds of the present invention also encompass Exendin compounds. Exendin-3 and Exendin-4 are biologically active peptides first isolated from Helodermatidae lizard venoms and have been shown to bind the GLP-1 receptor and stimulate cAMP-dependent H⁺ production in
10 mammalian parietal cells. Exendin-3 and Exendin-4 are both 39 amino acid peptides which are approximately 53% homologous to GLP-1. They act as potent agonists of GLP-1 activity. Notably, an N-terminally truncated derivative of Exendin, known as Exendin(9-39 amino acids), is an inhibitor
15 of Exendin-3, Exendin-4 and GLP-1.

20 An Exendin compound typically comprises a polypeptide having the amino acid sequence of Exendin-3, Exendin-4, or an analog or fragment thereof. Exendin-3 and Exendin-4 are disclosed in U.S. Patent No. 5,424,286.

Exendin-3 has the amino acid sequence of SEQ ID NO: 9:

	7	8	9	10	11	12	13	14	15	16	17
	His	Ser	Asp	Gly	Thr	Phe	Thr	Ser	Asp	Leu	Ser
	18	19	20	21	22	23	24	25	26	27	28
25	Lys	Gln	Met	Glu	Glu	Glu	Ala	Val	Arg	Leu	Phe
	29	30	31	32	33	34	35	36	37	38	39
	Ile	Glu	Trp	Leu	Lys	Asn	Gly	Gly	Pro	Ser	Ser
	40	41	42	43	44	45					
	Gly	Ala	Pro	Pro	Pro	Ser					

30 (SEQ ID NO: 9)

45
42

- 32 -

Exendin-4 has the amino acid sequence of SEQ ID NO: 10:

5 7 8 9 10 11 12 13 14 15 16 17
 His-Gly-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Leu-Ser-
18 19 20 21 22 23 24 25 26 27 28
 Lys-Gln-Met-Glu-Glu-Glu-Ala-Val-Arg-Leu-Phe-
29 30 31 32 33 34 35 36 37 38 39
 Ile-Glu-Trp-Leu-Lys-Asn-Gly-Gly-Pro-Ser-Ser-
40 41 42 43 44 45
10 Gly-Ala-Pro-Pro-Pro-Ser
 (SEQ ID NO: 10)

15 GLP-1 compounds also include Exendin fragments which
are polypeptides obtained after truncation of one or more
amino acids from the N-terminus and/or C-terminus of Exendin
or an Exendin analog. Furthermore, GLP-1 compounds include
Exendin polypeptides in which one or more amino acids have
been added to the N-terminus and/or C-terminus of Exendin r
fragments thereof. Exendin compounds of this type have up to
20 about forty-five amino acids.

GLP-1 compounds also include "Exendin analogs." An
Exendin analog has sufficient homology to Exendin-4,
Exendin-3, or a fragment thereof such that the analog has
insulinotropic activity. The activity of Exendin fragments
25 and/or analogs can be assessed using *in vitro* assays such as
those described in EP 619,322 and U.S. Patent No. 5,120,712.

Preferably, an Exendin analog has the amino acid
sequence of Exendin-4 or a fragment thereof, modified so
that from one, two, three, four or five amino acids differ
30 from the amino acid in corresponding position of Exendin-4
or the fragment of Exendin-4. In the nomenclature used
herein to designate Exendin compounds, the substituting
amino acid and its position is indicated prior to the parent

- 33 -

structure. For example, Val⁸-Exendin-4 designates an Exendin compound in which the glycine normally found at position 8 of Exendin-4 has been replaced with valine.

Another preferred group of GLP-1 compounds is composed of GLP-1/Exendin-4 analogs of formula VIII (SEQ ID NO:11).

7 8 9 10 11 12 13 14 15 16 17
Xaa-Xaa-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Xaa-Ser-
18 19 20 21 22 23 24 25 26 27 28
10 Xaa-Xaa-Xaa-Glu-Xaa-Xaa-Ala-Xaa-Xaa-Xaa-Phe-
29 30 31 32 33 34 35 36 37
Ile-Xaa-Trp-Leu-Xaa-Xaa-Gly-Xaa-R
formula VIII (SEQ ID NO: 11)

15 wherein:

Xaa at position 7 is: L-histidine, D-histidine, desamino-histidine, 2-amino-histidine, β -hydroxy-histidine, homohistidine, α -fluoromethyl-histidine or α -methyl-histidine;

20 Xaa at position 8 is: Gly, Ala, or Val;

Xaa at position 16 is: Leu or Val;

Xaa at position 18 is Lys or Ser;

Xaa at position 19 is: Gln or Tyr;

Xaa at position 20 is: Met or Leu;

25 Xaa at position 22 is: Glu or Gln;

Xaa at position 23 is: Glu, or Gln;

Xaa at position 25 is: Val or Ala;

Xaa at position 26 is: Arg or Lys;

Xaa at position 27 is Leu or Glu;

30 Xaa at position 30 is: Glu or Ala;

Xaa at position 33 is: Val or Lys;

Xaa at position 34 is: Asn or Lys;

Xaa at position 36 is: Gly or Arg; and

- 34 -

R at position 37 is: Gly, Pro, Pro-Ser-Ser-Gly-Ala-Pro-Pro-Pro-Ser, or is absent. The activity of 18 different species that fall within this genus is provided in Table 6.

Further Exendin-analogs that are useful for the present invention are described in PCT patent publications WO 99/25728 (Beeley et al.), WO 99/25727 (Beeley et al.), WO 98/05351 (Young et al.), WO 99/40788 (Young et al.), WO 99/07404 (Beeley et al), and WO 99/43708 (Knudsen et al).

The GLP-1 fusion proteins of the present invention can comprise glycosylation sites. Glycosylation is a chemical modification wherein sugar moieties are added to the protein at specific sites. Glycosylation of proteins play a role in ensuring the correct charge, confirmation, and stability of maturing protein and can target the protein to the cell surface and eventual secretion of the protein. Most importantly, glycosylation effects the *in vivo* clearance rate for many proteins. Sugars can be O-linked or N-linked. Generally, O-linked sugars are added to the hydroxyl-group oxygen of serine and threonine, while N-linked sugars are added to the amide nitrogen of asparagine. The consensus site for N-glycosylation is Asn X1 X2 wherein X1 is any amino acid except Pro and X2 is Ser or Thr.

GLP-1 compounds are generally not glycosylated *in vivo*; however, interestingly the GLP-1 fusion proteins of the present invention that comprise a GLP-1 compound with a C terminal extension fused to an Fc sequence is glycosylated at the last serine in the C terminal extension (SSGAPPPS*) and at threonine at position 11 in the N terminal region of Fc (AEPKSCDKTHT*CPPC . . .).

30

- 35 -

Heterologous Fc fusion proteins:

The GLP-1 compounds described above can be fused directly or via a peptide linker to the Fc portion of an immunoglobulin.

Immunoglobulins are molecules containing polypeptide chains held together by disulfide bonds, typically having two light chains and two heavy chains. In each chain, one domain (V) has a variable amino acid sequence depending on the antibody specificity of the molecule. The other domains (C) have a rather constant sequence common to molecules of the same class.

As used herein, the Fc portion of an immunoglobulin has the meaning commonly given to the term in the field of immunology. Specifically, this term refers to an antibody fragment which is obtained by removing the two antigen binding regions (the Fab fragments) from the antibody. One way to remove the Fab fragments is to digest the immunoglobulin with papain protease. Thus, the Fc portion is formed from approximately equal sized fragments of the constant region from both heavy chains, which associate through non-covalent interactions and disulfide bonds. The Fc portion can include the hinge regions and extend through the CH2 and CH3 domains to the C-terminus of the antibody. Representative hinge regions for human and mouse immunoglobulins can be found in Antibody Engineering, A Practical Guide, Borrebaeck, C.A.K., ed., W.H. Freeman and Co., 1992, the teachings of which are herein incorporated by reference. The Fc portion can further include one or more glycosylation sites. The amino acid sequence of a representative Fc protein containing a hinge region, CH2 and CH3 domains, and one N-glycosylation site at position 82 is shown in Figure 1.

- 36 -

There are five types of human immunoglobulin Fc regions with different effect or and pharmacokinetic properties: IgG, IgA, IgM, IgD, and IgE. IgG is the most abundant immunoglobulin in serum. IgG also has the longest half-life in serum of any immunoglobulin (23 days). Unlike other immunoglobulins, IgG is efficiently recirculated following binding to an Fc receptor. There are four IgG subclasses G1, G2, G3, and G4, each of which have different effect or functions. G1, G2, and G3 can bind C1q and fix complement while G4 cannot. Even though G3 is able to bind C1q more efficiently than G1, G1 is more effective at mediating complement-directed cell lysis. G2 fixes complement very inefficiently. The C1q binding site in IgG is located at the carboxy terminal region of the CH2 domain.

All IgG subclasses are capable of binding to Fc receptors (CD16, CD32, CD64) with G1 and G3 being more effective than G2 and G4. The Fc receptor binding region of IgG is formed by residues located in both the hinge and the carboxy terminal regions of the CH2 domain.

IgA can exist both in a monomeric and dimeric form held together by a J-chain. IgA is the second most abundant Ig in serum, but it has a half-life of only 6 days. IgA has three effect or functions. It binds to an IgA specific receptor on macrophages and eosinophils, which drives phagocytosis and degranulation, respectively. It can also fix complement via an unknown alternative pathway.

IgM is expressed as either a pentamer or a hexamer, both of which are held together by a J-chain. IgM has a serum half-life of 5 days. It binds weakly to C1q via a binding site located in its CH3 domain. IgD has a half-life of 3 days in serum. It is unclear what effect or functions are attributable to this Ig. IgE is a monomeric Ig and has a serum half-life of 2.5 days. IgE binds to two Fc receptors which drives degranulation and results in the release of proinflammatory agents.

- 37 -

Depending on the desired *in vivo* effect, the heterologous fusion proteins of the present invention may contain any of the isotypes described above or may contain mutated Fc regions wherein the complement and/or Fc receptor binding functions have been altered. Thus, the heterologous fusion proteins of the present invention may contain the entire Fc portion of an immunoglobulin, fragments of the Fc portion of an immunoglobulin, or analogs thereof fused to a GLP-1 compound.

10 The fusion proteins of the present invention can consist of single chain proteins or as multi-chain polypeptides. Two or more Fc fusion proteins can be produced such that they interact through disulfide bonds that naturally form between Fc regions. These multimers can
15 be homogeneous with respect to the GLP-1 compound or they may contain different GLP-1 compounds fused at the N-terminus of the Fc portion of the fusion protein.

Regardless of the final structure of the fusion protein, the Fc or Fc-like region must serve to prolong the
20 *in vivo* plasma half-life of the GLP-1 compound fused at the N-terminus. Furthermore, the fused GLP-1 compound must retain some biological activity. An increase in half-life can be demonstrated using the method described in Example 7 wherein the half-life of the fusion protein is compared to
25 the half-life of the GLP-1 compound alone. Biological activity can be determined by *in vitro* and *in vivo* methods known in the art. Representative biological assays are described in Examples 6, 8, and 9.

Since the Fc region of IgG produced by proteolysis has
30 the same *in vivo* half-life as the intact IgG molecule and Fab fragments are rapidly degraded, it is believed that the relevant sequence for prolonging half-life reside in the CH2 and/or CH3 domains. Further, it has been shown in the literature that the catabolic rates of IgG variants that do
35 not bind the high-affinity Fc receptor or Clq are

- 38 -

indistinguishable from the rate of clearance of the parent wild-type antibody, indicating that the catabolic site is distinct from the sites involved in Fc receptor or Clq binding. [Wawrzynczak et al., (1992) *Molecular Immunology* 29:221]. Site-directed mutagenesis studies using a murine IgG1 Fc region suggested that the site of the IgG1 Fc region that controls the catabolic rate is located at the CH2-CH3 domain interface.

Based on these studies, Fc regions can be modified at the catabolic site to optimize the half-life of the fusion proteins. It is preferable that the Fc region used for the heterologous fusion proteins of the present invention be derived from an IgG1 or an IgG4 Fc region. It is even more preferable that the Fc region be IgG4 or derived from IgG4. Preferably the IgG Fc region contains both the CH2 and CH3 regions including the hinge region.

Heterologous albumin fusion proteins:

The GLP-1 compounds described above can be fused directly or via a peptide linker to albumin or an analog, fragment, or derivative thereof.

Generally the albumin proteins making up part of the fusion proteins of the present invention can be derived from albumin cloned from any species. However, human albumin and fragments and analogs thereof are preferred to reduce the risk of the fusion protein being immunogenic in humans. Human serum albumin (HSA) consists of a single non-glycosylated polypeptide chain of 585 amino acids with a formula molecular weight of 66,500. The amino acid sequence of human HSA is shown in figure 2. [See Meloun, et al. (1975) *FEBS Letters* 58:136; Behrens, et al. (1975) *Fed. Proc.* 34:591; Lawn, et al. (1981) *Nucleic Acids Research* 9:6102-6114; Minghetti, et al. (1986) *J. Biol. Chem.* 261:6747]. A variety of polymorphic variants as well as analogs and fragments of albumin have been described. [See

- 39 -

Weitkamp, et al., (1973) Ann. Hum. Genet. 37:219]. For example, in EP 322,094, the inventors disclose various shorter forms of HSA. Some of these fragments include HSA(1-373), HSA(1-388), HSA(1-389), HSA(1-369), and HSA(1-419) and fragments between 1-369 and 1-419. EP 399,666 discloses albumin fragments that include HSA(1-177) and HSA(1-200) and fragments between HSA(1-177) and HSA(1-200).

It is understood that the heterologous fusion proteins of the present invention include GLP-1 compounds that are coupled to any albumin protein including fragments, analogs, and derivatives wherein such fusion protein is biologically active and has a longer plasma half-life than the GLP-1 compound alone. Thus, the albumin portion of the fusion protein need not necessarily have a plasma half-life equal to that of native human albumin. Fragments, analogs, and derivatives are known or can be generated that have longer half-lives or have half-lives intermediate to that of native human albumin and the GLP-1 compound of interest.

The heterologous fusion proteins of the present invention encompass proteins having conservative amino acid substitutions in the GLP-1 compound and/or the Fc or albumin portion of the fusion protein. A "conservative substitution" is the replacement of an amino acid with another amino acid that has the same net electronic charge and approximately the same size and shape. Amino acids with aliphatic or substituted aliphatic amino acid side chains have approximately the same size when the total number carbon and heteroatoms in their side chains differs by no more than about four. They have approximately the same shape when the number of branches in their side chains differs by no more than one. Amino acids with phenyl or substituted phenyl groups in their side chains are considered to have about the same size and shape. Except as otherwise specifically provided herein, conservative

- 40 -

substitutions are preferably made with naturally occurring amino acids.

However, the term "amino acid" is used herein in its broadest sense, and includes naturally occurring amino acids as well as non-naturally occurring amino acids, including amino acid analogs and derivatives. The latter includes molecules containing an amino acid moiety. One skilled in the art will recognize, in view of this broad definition, that reference herein to an amino acid includes, for example, naturally occurring proteogenic L-amino acids; D-amino acids; chemically modified amino acids such as amino acid analogs and derivatives; naturally occurring non-proteogenic amino acids such as norleucine, β -alanine, ornithine, GABA, etc.; and chemically synthesized compounds having properties known in the art to be characteristic of amino acids. As used herein, the term "proteogenic" indicates that the amino acid can be incorporated into a peptide, polypeptide, or protein in a cell through a metabolic pathway.

The incorporation of non-natural amino acids, including synthetic non-native amino acids, substituted amino acids, or one or more D-amino acids into the heterologous fusion proteins of the present invention can be advantageous in a number of different ways. D-amino acid-containing peptides, etc., exhibit increased stability *in vitro* or *in vivo* compared to L-amino acid-containing counterparts. Thus, the construction of peptides, etc., incorporating D-amino acids can be particularly useful when greater intracellular stability is desired or required. More specifically, D-peptides, etc., are resistant to endogenous peptidases and proteases, thereby providing improved bioavailability of the molecule, and prolonged lifetimes *in vivo* when such properties are desirable. Additionally, D-peptides, etc., cannot be processed efficiently for major histocompatibility complex class II-restricted presentation to T helper cells,

- 41 -

and are therefore, less likely to induce humoral immune responses in the whole organism.

In addition to structure/function analyses of the various polypeptides encompassed by the present invention, there are numerous factors that can be considered when selecting amino acids for substitution. One factor that can be considered in making such changes is the hydropathic index of amino acids. The importance of the hydropathic amino acid index in conferring interactive biological function on a protein has been discussed by Kyte and Doolittle (1982, *J. Mol. Biol.*, 157: 105-132). It is accepted that the relative hydropathic character of amino acids contributes to the secondary structure of the resultant protein. This, in turn, affects the interaction of the protein with molecules such as enzymes, substrates, receptors, ligands, DNA, antibodies, antigens, etc. Based on its hydrophobicity and charge characteristics, each amino acid has been assigned a hydropathic index as follows:

isoleucine (+4.5); valine (+4.2); leucine (+3.8);
phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2);
glutamate/glutamine/aspartate/asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

As is known in the art, certain amino acids in a peptide, polypeptide, or protein can be substituted for other amino acids having a similar hydropathic index or score and produce a resultant peptide, etc., having similar or even improved biological activity. In making such changes, it is preferable that amino acids having hydropathic indices within ± 2 are substituted for one another. More preferred substitutions are those wherein the amino acids have hydropathic indices within ± 1 . Most

- 42 -

preferred substitutions are those wherein the amino acids have hydropathic indices within ± 0.5 .

Like amino acids can also be substituted on the basis of hydrophilicity. U.S. Patent No. 4,554,101 discloses that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein. The following hydrophilicity values have been assigned to amino acids: arginine/lysine (+3.0); aspartate/glutamate (+3.0 $\pm$ 1); serine (+0.3); asparagine/glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5 $\pm$ 1); alanine/histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine/isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); and tryptophan (-3.4). Thus, one amino acid in a peptide, polypeptide, or protein can be substituted by another amino acid having a similar hydrophilicity score and still produce a resultant peptide, etc., having similar biological activity, i.e., still retaining correct biological function. In making such changes, amino acids having hydropathic indices within ± 2 are preferably substituted for one another, those within ± 1 are more preferred, and those within ± 0.5 are most preferred.

As outlined above, amino acid substitutions in the fusion proteins of the present invention can be based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, etc. Furthermore, substitutions can be made based on secondary structure propensity. For example, a helical amino acid can be replaced with an amino acid that would preserve the helical structure. Exemplary substitutions that take various of the foregoing characteristics into consideration in order to produce conservative amino acid changes resulting in silent changes within the present peptides, etc., can be selected from their members of the class to which the naturally

- 43 -

occurring amino acid belongs. Amino acids can be divided into the following four groups: (1) acidic amino acids; (2) basic amino acids; (3) neutral polar amino acids; and (4) neutral non-polar amino acids.

5

General methods for making the heterologous fusion proteins of the present invention.

Although the heterologous fusion proteins of the present invention can be made by a variety of different methods, recombinant methods are preferred. For purposes of the present invention, as disclosed and claimed herein, the following general molecular biology terms and abbreviations are defined below. The terms and abbreviations used in this document have their normal meanings unless otherwise designated. For example, "°C" refers to degrees Celsius; "mmol" refers to millimole or millimoles; "mg" refers to milligrams; "µg" refers to micrograms; "ml or mL" refers to milliliters; and "µl or µL" refers to microliters. Amino acids abbreviations are as set forth in 37 C.F.R. § 1.822 (b) (2) (1994).

"Base pair" or "bp" as used herein refers to DNA or RNA. The abbreviations A, C, G, and T correspond to the 5'-monophosphate forms of the deoxyribonucleosides (deoxy)adenosine, (deoxy)cytidine, (deoxy)guanosine, and thymidine, respectively, when they occur in DNA molecules. The abbreviations U, C, G, and A correspond to the 5'-monophosphate forms of the ribonucleosides uridine, cytidine, guanosine, and adenosine, respectively when they occur in RNA molecules. In double stranded DNA, base pair may refer to a partnership of A with T or C with G. In a DNA/RNA, heteroduplex base pair may refer to a partnership of A with U or C with G. (See the definition of "complementary", infra.)

"Digestion" or "Restriction" of DNA refers to the catalytic cleavage of the DNA with a restriction enzyme that

- 44 -

acts only at certain sequences in the DNA ("sequence-specific endonucleases"). The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors, and other requirements were used as

5 would be known to one of ordinary skill in the art.

Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer or can be readily found in the literature.

"Ligation" refers to the process of forming
10 phosphodiester bonds between two double stranded nucleic acid fragments. Unless otherwise provided, ligation may be accomplished using known buffers and conditions with a DNA ligase, such as T4 DNA ligase.

"Plasmid" refers to an extrachromosomal (usually) self-
15 replicating genetic element. Plasmids are generally designated by a lower case "p" followed by letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available
20 plasmids in accordance with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

"Recombinant DNA cloning vector" as used herein refers
25 to any autonomously replicating agent, including, but not limited to, plasmids and phages, comprising a DNA molecule to which one or more additional DNA segments can or have been added.

"Recombinant DNA expression vector" as used herein
30 refers to any recombinant DNA cloning vector in which a promoter to control transcription of the inserted DNA has been incorporated.

"Transcription" refers to the process whereby
information contained in a nucleotide sequence of DNA is
35 transferred to a complementary RNA sequence.

- 45 -

"Transfection" refers to the uptake of an expression vector by a host cell whether or not any coding sequences are, in fact, expressed. Numerous methods of transfection are known to the ordinarily skilled artisan, for example, calcium phosphate co-precipitation, liposome transfection, and electroporation. Successful transfection is generally recognized when any indication of the operation of this vector occurs within the host cell.

"Transformation" refers to the introduction of DNA into an organism so that the DNA is replicable, either as an extrachromosomal element or by chromosomal integration. Methods of transforming bacterial and eukaryotic hosts are well known in the art, many of which methods, such as nuclear injection, protoplast fusion or by calcium treatment using calcium chloride are summarized in J. Sambrook, et al., Molecular Cloning: A Laboratory Manual, (1989). Generally, when introducing DNA into Yeast the term transformation is used as opposed to the term transfection.

"Translation" as used herein refers to the process whereby the genetic information of messenger RNA (mRNA) is used to specify and direct the synthesis of a polypeptide chain.

"Vector" refers to a nucleic acid compound used for the transfection and/or transformation of cells in gene manipulation bearing polynucleotide sequences corresponding to appropriate protein molecules which, when combined with appropriate control sequences, confers specific properties on the host cell to be transfected and/or transformed. Plasmids, viruses, and bacteriophage are suitable vectors. Artificial vectors are constructed by cutting and joining DNA molecules from different sources using restriction enzymes and ligases. The term "vector" as used herein includes Recombinant DNA cloning vectors and Recombinant DNA expression vectors.

- 46 -

"Complementary" or "Complementarity", as used herein, refers to pairs of bases (purines and pyrimidines) that associate through hydrogen bonding in a double stranded nucleic acid. The following base pairs are complementary:
5 guanine and cytosine; adenine and thymine; and adenine and uracil.

"Hybridization" as used herein refers to a process in which a strand of nucleic acid joins with a complementary strand through base pairing. The conditions employed in the
10 hybridization of two non-identical, but very similar, complementary nucleic acids varies with the degree of complementarity of the two strands and the length of the strands. Such techniques and conditions are well known to practitioners in this field.

15 "Isolated amino acid sequence" refers to any amino acid sequence, however, constructed or synthesized, which is locationally distinct from the naturally occurring sequence.

"Isolated DNA compound" refers to any DNA sequence, however constructed or synthesized, which is locationally
20 distinct from its natural location in genomic DNA.

"Isolated nucleic acid compound" refers to any RNA or DNA sequence, however constructed or synthesized, which is locationally distinct from its natural location.

"Primer" refers to a nucleic acid fragment which
25 functions as an initiating substrate for enzymatic or synthetic elongation.

"Promoter" refers to a DNA sequence which directs transcription of DNA to RNA.

"Probe" refers to a nucleic acid compound or a
30 fragment, thereof, which hybridizes with another nucleic acid compound.

"Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe
35 length, washing temperature, and salt concentration. In

- 47 -

general, longer probes require higher temperatures for proper annealing, while short probes need lower temperatures.

Hybridization generally depends on the ability of denatured DNA to re-anneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature that can be used. As a result, it follows that higher relative temperatures would tend to make the reactions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., *Current Protocols in Molecular Biology*, Wiley Interscience Publishers, 1995.

"Stringent conditions" or "high stringency conditions", as defined herein, may be identified by those that (1) employ low ionic strength and high temperature for washing, for example, 15 mM sodium chloride/1.5 mM sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride/75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5X SSC (750 mM sodium chloride, 75 mM sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5X Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C with washes at 42°C in 0.2X SSC (30 mM sodium chloride/3 mM sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1X SSC containing EDTA at 55°C.

"Moderately stringent conditions" may be identified as described by Sambrook et al. [*Molecular Cloning: A Laboratory Manual*, New York: Cold Spring Harbor Press, (1989)], and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength, and %SDS) less

- 48 -

stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5X SSC (750 mM sodium chloride, 75 mM sodium citrate), 50 mM sodium phosphate at pH 7.6, 5X Denhardt's solution, 10% dextran sulfate, and 20 mg/mL denatured sheared salmon sperm DNA, followed by washing the filters in 1X SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc., as necessary to accommodate factors such as probe length and the like.

"PCR" refers to the widely-known polymerase chain reaction employing a thermally-stable DNA polymerase.

"Leader sequence" refers to a sequence of amino acids which can be enzymatically or chemically removed to produce the desired polypeptide of interest.

"Secretion signal sequence" refers to a sequence of amino acids generally present at the N-terminal region of a larger polypeptide functioning to initiate association of that polypeptide with the cell membrane compartments like endoplasmic reticulum and secretion of that polypeptide through the plasma membrane.

Construction of DNA encoding the heterologous fusion proteins of the present invention:

Wild-type albumin and Immunoglobulin proteins can be obtained from a variety of sources. For example, these proteins can be obtained from a cDNA library prepared from tissue or cells which express the mRNA of interest at a detectable level. Libraries can be screened with probes designed using the published DNA or protein sequence for the particular protein of interest. For example, immunoglobulin light or heavy chain constant regions are described in Adams, et al. (1980) Biochemistry 19:2711-2719; Goughet, et al. (1980) Biochemistry 19:2702-2710; Dolby, et al. (1980) Proc. Natl. Acad. Sci. USA 77:6027-6031; Rice et al. (1982)

- 49 -

Proc. Natl. Acad. Sci. USA 79:7862-7862; Falkner, et al. (1982) Nature 298:286-288; and Morrison, et al. (1984) Ann. Rev. Immunol. 2:239-256. Some references disclosing albumin protein and DNA sequences include Meloun, et al. (1975) FEBS Letters 58:136; Behrens, et al. (1975) Fed. Proc. 34:591; Lawn, et al. (1981) Nucleic Acids Research 9:6102-6114; and Minghetti, et al. (1986) J. Biol. Chem. 261:6747

Screening a cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, NY (1989). An alternative means to isolate a gene encoding an albumin or immunoglobulin protein is to use PCR methodology [Sambrook et al., supra; Dieffenbach et al., *PCR Primer: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, NY (1995)]. PCR primers can be designed based on published sequences.

Generally the full-length wild-type sequences cloned from a particular species can serve as a template to create analogs, fragments, and derivatives that retain the ability to confer a longer plasma half-life on the GLP-1 compound that is part of the fusion protein. It is preferred that the Fc and albumin portions of the heterologous fusion proteins of the present invention be derived from the native human sequence in order to reduce the risk of potential immunogenicity of the fusion protein in humans.

In particular, it is preferred that the immunoglobulin portion of a fusion protein encompassed by the present invention contain only an Fc fragment of the immunoglobulin. Depending on whether particular effect or functions are desired and the structural characteristics of the fusion protein, an Fc fragment may contain the hinge region along with the CH2 and CH3 domains or some other combination thereof. These Fc fragments can be generated using PCR techniques with primers designed to hybridize to sequences

- 50 -

corresponding to the desired ends of the fragment.

Similarly, if fragments of albumin are desired, PCR primers can be designed which are complementary to internal albumin sequences. PCR primers can also be designed to create
5 restriction enzyme sites to facilitate cloning into expression vectors.

DNA encoding the GLP-1 compounds of the present invention can be made by a variety of different methods including cloning methods like those described above as well
10 as chemically synthesized DNA. Chemical synthesis may be attractive given the short length of the encoded peptide. The amino acid sequence for GLP-1 has been published as well as the sequence of the preproglucagon gene. [Lopez, et al. (1983) Proc. Natl. Acad. Sci., USA 80:5485-5489; Bell, et
15 al. (1983) Nature, 302:716-718; Heinrich, G., et al. (1984) Endocrinol, 115:2176-2181; Ghigliione, M., et al. 1984) Diabetologia 27:599-600]. Thus, primers can be designed to PCR native GLP-1 compounds and fragments thereof.

The gene encoding a fusion protein can then be
20 constructed by ligating DNA encoding a GLP-1 compound in-frame to DNA encoding an albumin or Fc protein. The gene encoding the GLP-1 compound and the gene encoding the albumin or Fc protein can also be joined in-frame via DNA encoding a linker peptide.

25 The *in vivo* function and stability of the heterologous fusion proteins of the present invention can be optimized by adding small peptide linkers to prevent potentially unwanted domain interactions. Although these linkers can potentially be any length and consist of any combination of amino acids,
30 it is preferred that the length be no longer than necessary to prevent unwanted domain interactions and/or optimize biological activity and/or stability. Generally, the linkers should not contain amino acids with extremely bulky side chains or amino acids likely to introduce significant
35 secondary structure. It is preferred that the linker be

- 51 -

serine-glycine rich and be less than 30 amino acids in length. It is more preferred that the linker be no more than 20 amino acids in length. It is even more preferred that the linker be no more than 15 amino acids in length. A preferred linker contains repeats of the sequence Gly-Gly-Gly-Gly-Ser. It is preferred that there be between 2 and 6 repeats of this sequence. It is even more preferred that there be between 3 and 4 repeats of this sequence.

The DNA encoding wild-type GLP-1, albumin, and Fc polypeptides and fragments thereof can be mutated either before ligation or in the context of a cDNA encoding an entire fusion protein. A variety of mutagenesis techniques are well known in the art. For example, a mutagenic PCR method utilizes strand overlap extension to create specific base mutations for the purposes of changing a specific amino acid sequence in the corresponding protein. This PCR mutagenesis requires the use of four primers, two in the forward orientation (primers A and C) and two in the reverse orientation (primers B and D). A mutated gene is amplified from the wild-type template in two different stages. The first reaction amplifies the gene in halves by performing an A to B reaction and a separate C to D reaction wherein the B and C primers target the area of the gene to be mutated. When aligning these primers with the target area, they contain mismatches for the bases that are targeted to be changed. Once the A to B and C to D reactions are complete, the reaction products are isolated and mixed for use as the template for the A to D reaction. This reaction then yields the full, mutated product.

Once a gene encoding an entire fusion protein is produced it can be cloned into an appropriate expression vector. Specific strategies that can be employed to make the GLP-1 fusion proteins of the present invention are described in example 1.

- 52 -

General methods to recombinantly express the heterologous fusion proteins of the present invention:

Host cells are transfected or transformed with expression or cloning vectors described herein for heterologous fusion protein production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in *Mammalian Cell Biotechnology: A Practical Approach*, M. Butler, ed. (IRL Press, 1991) and Sambrook, et al., supra. Methods of transfection are known to the ordinarily skilled artisan, for example, CaPO₄ and electroporation. General aspects of mammalian cell host system transformations have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of van Solingen et al., *J Bact.* 130(2): 946-7 (1977) and Hsiao et al., *Proc. Natl. Acad. Sci. USA* 76(8): 3829-33 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, r polycations, e.g., polybrene or polyomithine, may also be used. For various techniques for transforming mammalian cells, see Keown, et al., *Methods in Enzymology* 185: 527-37 (1990) and Mansour, et al., *Nature* 336(6197): 348-52 (1988).

Suitable host cells for cloning or expressing the nucleic acid (e.g., DNA) in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 3 1.446); *E. coli* X1 776 (ATCC 3 1.537); *E. coli* strain W3 110

- 53 -

(ATCC 27.325) and K5 772 (ATCC 53.635). Other suitable prokaryotic host cells include Enterobacteriaceae such as *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescans*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (e.g., *B. licheniformis* 4 1 P disclosed in DD266,7 10, published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3 110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype *ronA*; *E. coli* W3 110 strain 9E4, which has the complete genotype *ton4 ptr3*; *E. coli* W3110 strain 27C7 (ATCC 55,244), which has the complete genotype *tonA, ptr3 phoA* E15 (*argF-lac*) I69 *degP ompT /can'*; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vivo* methods of cloning, e.g., PCR or other nucleic acid polymerase reactions, are suitable.

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for fusion protein vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* [Beach and Nurse, *Nature* 290: 140-3 (1981); EP 139,383 published 2 May 1995]; *Muyveromyces* hosts [U.S. Patent No. 4,943,529; Flear, et al., *Bio/Technology* 9(10): 968-75 (1991)] such as, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574) [de Louvencourt et al., J.

- 54 -

Bacteriol. 154(2): 737-42 (1983)]; *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickerhamii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilae* (ATCC 36,906) [Van den Berg et al., *Bio/Technology* 8(2): 135-9 (1990)]; *K.*

5 *thermotolerans*, and *K. marxianus*; *Yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070) [Sreekrishna et al., *J. Basic Microbiol.* 28(4): 265-78 (1988)]; *Candida*; *Trichoderma reesei* (EP 244,234); *Neurospora crassa* [Case, et al., *Proc. Natl. Acad. Sci. USA* 76(10): 5259-63 (1979)]; *Schwanniomyces* such as

10 *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, e.g., *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* [Ballance et al., *Biochem. Biophys. Res. Comm.* 112(1): 284-9 (1983)];

15 Tilburn, et al., *Gene* 26(2-3): 205-21 (1983); Yelton, et al., *Proc. Natl. Acad. Sci. USA* 81(5): 1470-4 (1984)] and *A. niger* [Kelly and Hynes, *EMBO J.* 4(2): 475-9 (1985)]. Methylophilic yeasts are selected from the genera consisting of *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis*, and

20 *Rhodotorula*. A list of specific species that are exemplary of this class of yeast may be found in C. Antony, *The Biochemistry of Methylophilic Yeasts* 269 (1982).

Suitable host cells for the expression of the fusion proteins of the present invention are derived from

25 multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sp, *Spodoptera* high5 as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney

30 CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line [293 or 293 cells subcloned for growth in suspension culture, Graham, et al., *J. Gen Virol.*, 36(1): 59-74 (1977)]; Chinese hamster ovary cells/-DHFR [CHO, Urlaub and Chasin, *Proc. Natl. Acad. Sci. USA*, 77(7): 4216-20

35 (1980)]; mouse sertoli cells [TM4, Mather, *Biol. Reprod.*

- 55 -

23(1):243-52 (1980)]; human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

- 5 The fusion proteins of the present invention may be recombinantly produced directly, or as a protein having a signal sequence or other additional sequences which create a specific cleavage site at the N-terminus of the mature fusion protein. In general, the signal sequence may be a component
- 10 of the vector, or it may be a part of the fusion protein-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders.
- 15 For yeast secretion the signal sequence may be, e.g., the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* cc-factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179), or
- 20 the signal described in WO 90/13646. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species as well as viral secretory leaders.
- 25 Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-
- 30 negative bacteria, the 2u plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.
- Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical
- 35 selection genes encode proteins that (a) confer resistance to

- 56 -

antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement autotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*.

An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the fusion protein-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described [Urlaub and Chasin, *Proc. Natl. Acad. Sci. USA*, 77(7): 4216-20 (1980)]. A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7 [Stinchcomb, et al., *Nature* 282(5734): 39-43 (1979); Kingsman, et al., *Gene* 7(2): 141-52 (1979); Tachumper, et al., *Gene* 10(2): 157-66 (1980)]. The *trp1* gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEPC1 [Jones, *Genetics* 85: 23-33 (1977)].

Expression and cloning vectors usually contain a promoter operably linked to the fusion protein-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β -lactamase and lactose promoter systems [Chang, et al., *Nature* 275(5681): 617-24 (1978); Goeddel, et al., *Nature* 281(5732): 544-8 (1979)], alkaline phosphatase, a tryptophan (up) promoter system [Goeddel, *Nucleic Acids Res.* 8(18): 4057-74 (1980); EP 36,776 published 30 September 1981], and hybrid promoters such as the *tat* promoter [deBoer, et al., *Proc. Natl. Acad. Sci. USA* 80(1): 21-5 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the fusion protein.

70
67

- 57 -

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman, et al., *J. Biol. Chem.* 255(24): 12073-80 (1980)] or other glycolytic enzymes [Hess et al., *J. Adv. Enzyme Reg.* 7: 149 (1968); Holland, *Biochemistry* 17(23): 4900-7 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657. Transcription of fusion protein-encoding mRNA from vectors in mammalian host cells may be controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus, adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a polynucleotide encoding a fusion protein by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin,

71
68

- 58 -

elastase, albumin, a-ketoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the fusion protein coding sequence but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and occasionally 3' untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding the fusion protein.

Various forms of a fusion protein may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g., Triton-X 100) or by enzymatic cleavage. Cells employed in expression of a fusion protein can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

Purification of the heterologous fusion proteins of the present invention:

Once the heterologous fusion proteins of the present invention are expressed in the appropriate host cell, the analogs can be isolated and purified. The following procedures are exemplary of suitable purification procedures: fractionation on carboxymethyl cellulose; gel filtration such

- 59 -

as Sephadex G-75; anion exchange resin such as DEAE or Mono-Q; cation exchange such as CM or Mono-S; protein A sepharose to remove contaminants such as IgG; metal chelating columns to bind epitope-tagged forms of the polypeptide; reversed-phase HPLC; chromatofocusing; silica gel; ethanol precipitation; and ammonium sulfate precipitation.

Various methods of protein purification may be employed and such methods are known in the art and described, for example, in Deutscher, *Methods in Enzymology* 182: 83-9 (1990) and Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag, NY (1982). The purification step(s) selected will depend on the nature of the production process used and the particular fusion protein produced. For example, fusion proteins comprising an Fc fragment can be effectively purified using a Protein A or Protein G affinity matrix. Low or high pH buffers can be used to elute the fusion protein from the affinity matrix. Mild elution conditions will aid in preventing irreversible denaturation of the fusion protein. Imidazole-containing buffers can also be used. Example 3 describes some successful purification protocols for the fusion proteins of the present invention.

Characterization of the heterologous fusion proteins of the present invention:

Numerous methods exist to characterize the fusion proteins of the present invention. Some of these methods include: SDS-PAGE coupled with protein staining methods or immunoblotting using anti-IgG or anti-HSA antibodies. Other methods include matrix assisted laser desorption/ionization-mass spectrometry (MALDI-MS), liquid chromatography/mass spectrometry, isoelectric focusing, analytical anion exchange, chromatofocussing, and circular dichroism to name a few. A representative number of heterologous fusion proteins were characterized using SDS-PAGE coupled with immunoblotting as

77
70

- 60 -

well as mass spectrometry (See examples 4 and 5 and figures 3 and 4).

For example table 3 (see example 5) illustrates the calculated molecular mass for a representative number of fusion proteins as well as the mass as determined by mass spectrometry. In addition, Figures 3 and 4 illustrate molecular weights of a representative number of fusion proteins as determined by SDS PAGE. All heterologous fusion proteins tested were expressed and secreted transiently. In addition, the IgK signal sequence was cleaved to yield proteins with the correct N-terminus.

Further, table 3 illustrates that in some instances the mass determined by mass spectrometry is greater than expected. This is the result of glycosylation of the Fc portion and the C terminal extension. Enzymatic digestion of the fusion proteins followed by reversed-phase HPLC and mass spectrometry can identify peptide fractions that contain sugar moieties. These fractions can then be N-terminal amino acid sequenced to identify the potential glycosylation site. For example, characterization of Exendin-4-Fc (SEQ ID NO: 29) shows that the serine at position 39 and threonine at position 50 are O-linked glycosylated and the asparagine at position 122 is N-linked glycosylated.

A representative number of GLP-1 fusion proteins were also tested for activity. Numerous methods exist to detect GLP-1 activity *in vitro* and *in vivo* (see examples 6, 7, 8, and 9). Table 4 (example 6) illustrates GLP-1 receptor activity associated with several GLP-1 fusions. The numbers are relative to the activity associated with Val⁸-GLP-1(7-37)OH. All fusion proteins tested had GLP-1 receptor activity. A low level of *in vitro* activity is not necessarily indicative of a weak effect *in vivo*. Because of the substantial increase in the half-life of these fusion proteins, weak *in vitro* activity is not generally a predictor of weak *in vivo* activity. Figure 7 and example 7 illustrate the prolonged half-life associated

74
71

- 61 -

with the fusion proteins of the present invention. For example, Val⁸-GLP-1-Fc had a half-life of approximately 45 hours in monkeys, Val⁸-GLP-1-HSA had a half-life of about 87 hours in monkeys, Gly⁸-Glu²²-GLP-1-CEx-Linker-IgG1 had a half-life after IV administration of approximately 55 hours in dogs, and Gly⁸-Glu²²-GLP-1-CEx-Linker-IgG1 had a half-life after SC administration of approximately 38 hours in dogs.

Compositions of the invention:

10 Physical stability is also an essential feature for therapeutic protein formulations. GLP-1 compounds have been especially difficult to manufacture and formulate due to structural changes that occur during processing. For example, some GLP-1 compounds have a general tendency to aggregate. In
15 addition, it has been shown that some GLP-1 compounds convert from a soluble and active α -helix form to an insoluble and potentially inactive β -sheet form. The fusion of GLP-1 compounds to large proteins such as the Fc region of an IgG or albumin not only acts to increase the half-life of the GLP-1
20 compound, but also contributes to the physical and conformational stability of the GLP-1 compound. For example, Val⁸-GLP-1-Linker-HSA in PBS is stable at 37°C out to about 30 days.

25 The heterologous fusion proteins of the present invention may be formulated with one or more excipients. The active fusion proteins of the present invention may be combined with a pharmaceutically acceptable buffer, and the pH adjusted to provide acceptable stability, and a pH acceptable for administration such as parenteral administration.

30 Optionally, one or more pharmaceutically-acceptable anti-microbial agents may be added. Meta-cresol and phenol are preferred pharmaceutically-acceptable microbial agents. One or more pharmaceutically-acceptable salts may be added to adjust the ionic strength or tonicity. One or more excipients
35 may be added to further adjust the is tonicity of the

formulation. Glycerin is an example of an isotonicity-adjusting excipient. Pharmaceutically acceptable means suitable for administration to a human or other animal and thus, does not contain toxic elements or undesirable
5 contaminants and does not interfere with the activity of the active compounds therein.

A pharmaceutically-acceptable salt form of the heterologous fusion proteins of the present invention may be used in the present invention. Acids commonly employed to
10 form acid addition salts are inorganic acids such as hydrochloric acid, hydrobromic acid, hydriodic acid, sulfuric acid, phosphoric acid, and the like, and organic acids such as p-toluenesulfonic acid, methanesulfonic acid, oxalic acid, p-bromophenyl-sulfonic acid, carbonic acid, succinic acid,
15 citric acid, benzoic acid, acetic acid, and the like. Preferred acid addition salts are those formed with mineral acids such as hydrochloric acid and hydrobromic acid.

Base addition salts include those derived from inorganic bases, such as ammonium or alkali or alkaline earth metal
20 hydroxides, carbonates, bicarbonates, and the like. Such bases useful in preparing the salts of this invention thus include sodium hydroxide, potassium hydroxide, ammonium hydroxide, potassium carbonate, and the like.

25 Administration of Compositions:

Administration may be via any route known to be effective by the physician of ordinary skill. Peripheral, parenteral is one such method. Parenteral administration is commonly understood in the medical literature as the injection of a
30 dosage form into the body by a sterile syringe or some other mechanical device such as an infusion pump. Peripheral parenteral routes can include intravenous, intramuscular, subcutaneous, and intraperitoneal routes of administration.

The heterologous fusion proteins of the present invention
35 may also be amenable to administration by oral, rectal, nasal,

76
73

- 63 -

or lower respiratory routes, which are non-parenteral routes. Of these non-parenteral routes, the lower respiratory route and the oral route are preferred.

The fusion proteins of the present invention can be used to treat a wide variety of diseases and conditions. The fusion proteins of the present invention primarily exert their biological effects by acting at a receptor referred to as the "GLP-1 receptor." Subjects with diseases and/or conditions that respond favorably to GLP-1 receptor stimulation or to the administration of GLP-1 compounds can therefore be treated with the GLP-1 fusion proteins of the present invention. These subjects are said to "be in need of treatment with GLP-1 compounds" or "in need of GLP-1 receptor stimulation". Included are subjects with non-insulin dependent diabetes, insulin dependent diabetes, stroke (see WO 00/16797), myocardial infarction (see WO 98/08531), obesity (see WO 98/19698), catabolic changes after surgery (see U.S. Patent No. 6,006,753), functional dyspepsia and irritable bowel syndrome (see WO 99/64060). Also included are subjects requiring prophylactic treatment with a GLP-1 compound, e.g., subjects at risk for developing non-insulin dependent diabetes (see WO 00/07617). Subjects with impaired glucose tolerance or impaired fasting glucose, subjects whose body weight is about 25% above normal body weight for the subject's height and body build, subjects with a partial pancreatectomy, subjects having one or more parents with non-insulin dependent diabetes, subjects who have had gestational diabetes and subjects who have had acute or chronic pancreatitis are at risk for developing non-insulin dependent diabetes.

An "effective amount" of a GLP-1 compound is the quantity which results in a desired therapeutic and/or prophylactic effect without causing unacceptable side-effects when administered to a subject in need of GLP-1 receptor stimulation. A "desired therapeutic effect" includes one or more of the following: 1) an amelioration of

77
74

- 64 -

the symptom(s) associated with the disease or condition; 2) a delay in the onset of symptoms associated with the disease or condition; 3) increased longevity compared with the absence of the treatment; and 4) greater quality of life compared with the absence of the treatment. For example, an "effective amount" of a GLP-1 compound for the treatment of diabetes is the quantity that would result in greater control of blood glucose concentration than in the absence of treatment, thereby resulting in a delay in the onset of diabetic complications such as retinopathy, neuropathy or kidney disease. An "effective amount" of a GLP-1 compound for the prevention of diabetes is the quantity that would delay, compared with the absence of treatment, the onset of elevated blood glucose levels that require treatment with anti-hypoglycaemic drugs such as sulfonyl ureas, thiazolidinediones, insulin and/or bisguanidines.

The dose of fusion protein effective to normalize a patient's blood glucose will depend on a number of factors, among which are included, without limitation, the subject's sex, weight and age, the severity of inability to regulate blood glucose, the route of administration and bioavailability, the pharmacokinetic profile of the fusion protein, the potency, and the formulation.

The present invention comprises GLP-1 compounds that have improved biochemical and biophysical properties by virtue of being fused to an albumin protein, an albumin fragment, an albumin analog, a Fc protein, a Fc fragment, or a Fc analog. These heterologous proteins can be successfully expressed in host cells, retain signaling activities associated with activation of the GLP-1 receptor, and have prolonged half-lives.

The following examples are presented to further describe the present invention. The scope of the present invention is not to be construed as merely consisting of the following examples. Those skilled in the art will recognize

78
75

- 65 -

that the particular reagents, equipment, and procedures described are merely illustrative and are not intended to limit the present invention in any manner.

5 Example 1: Construction of DNA
 encoding heterologous fusion proteins

Example 1a Construction of DNA encoding Val⁸-GLP-1(7-37)-

10 Fc: A Fc portion of human IgG1 was isolated from a
cDNA library and contains the full hinge region and the CH2
and CH3 domains. A fragment containing 696 base pairs of
this Fc portion of human IgG1 was subcloned into the *NheI*
and *Eco47III* sites of mammalian expression vector pJB02 to
create pJB02/Fc (see Figure 5). DNA encoding the Igκ
secretion signal sequence fused to Val⁸-GLP-1(7-37) was
15 generated by *in vitro* hybridization of four overlapping and
complementary oligonucleotides:

5'- CTAGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
CCAGGTTCCACTGGTGACCAGTG - 3' [SEQ ID NO:12]

20

5'- GAGGGCACCTTCACCTCCGACGTGTCCTCCTATCTGGAGGGCCAGGCCGCCAAGGA
GTTTCATCGCCTGGCTGGTGAAGGGAAGAGGC - 3' [SEQ ID NO:13]

25

5'- TGAAGGTGCCCTCCACGTGGTCACCAGTGGAACCTGGAACCCAGAGCAGCAGTA
CCCATAGCAGGAGTGTGTCTGTCTCCATGGTGG - 3' [SEQ ID
NO:14]

30

5'- GCCTCTTCCCTTCACCAGCCAGGCGATGAACTCCTTGGCGGCCTGGCCCTCCAGA
TAGGAGGACACGTCGGAGG - 3' [SEQ ID NO:15]

35

The hybridization reaction was carried out using
equivalent amounts of each oligonucleotide (1 pm/μl final
concentration for each oligo). The mixture of
oligonucleotides was heated for 5 min at 100°C in ligation
buffer (50 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 10 mM DTT, 1mM

79
76

- 66 -

ATP, 25 µg/ml bovine serum albumin) and then cooled over at least 2 hours to 30°C.

The resulting hybridization product was ligated for 2 hours at room temperature or overnight at 16°C to the

5 pJB02/Fc vector backbone which had been digested with *NheI* and *Eco47III*. The ligation products were used to transform competent XL-1 Blue cells (Stratagene). Recombinant plasmids were screened for the presence of peptide coding inserts by digesting clones with *NcoI* (encoding the Kozak sequence and first Met of the signal peptide) and sequenced. 10 The resulting expression plasmid used for transfection assays was denoted pJB02-V8-GLP-1-Fc (Figure 5).

Example 1b Construction of DNA encoding Val⁸-GLP-1(7-37)-

15 HSA: The plasmid HSA/pcDNA3.1GS was purchased from Invitrogen (Catalog # H-M12523M-pcDNA3.1/GS) and used as a template to isolate the cDNA encoding human serum albumin (HSA). The HSA cDNA was prepared using PCR wherein the DNA encoding the leader sequence as well as the six amino acid 20 pro-peptide was removed from the 5' end. In addition, stop codons were added directly at the 3' end of the HSA coding sequence. Finally, restriction enzyme sites were engineered at the 5' and 3' end to facilitate cloning. The HSA DNA sequence present in the original vector purchased from 25 Invitrogen contained a single base change in the 3' region of the gene (position 667) compared to the native human sequence. This change would result in a codon for Asn instead of Asp. Thus, using the strand overlapping PCR mutagenesis method discussed above, the codon was changed to 30 code for Asp at this position. The resulting HSA encoding DNA was cloned into the *NheI* and *HindIII* sites of pJB02 to create pJB02-HSA (Figure 6).

The IgK leader sequence fused to the Val⁸-GLP-1(7-37) sequence was generated as discussed in Example 1a. This DNA

- 67 -

was ligated into the *NheI* and *FspI* sites of pJB02-HSA to create pJB02- Val⁸-GLP-1-HSA.

Example 1c Construction of DNA encoding Val⁸-GLP-1(7-37)-linker-HSA:

The vector pJB02-HSA was prepared as discussed in Example 1b. DNA encoding the linker sequence [GGGGS], was ligated in frame to the 5' end of the HSA encoding DNA to create pJB02-linker-HSA. (Figure 7), DNA encoding the Igk leader sequence and fused to the Val⁸-GLP-1(7-37) sequence and the 5' part of the linker sequence was generated as discussed in Example 1a. This DNA was ligated into the *NheI* and *BspEI* sites of pJB02 to create pJB02- Val⁸-GLP-1-linker-HSA.

15

Example 1d Construction of DNA encoding Exendin-4-Fc:

The plasmid pJB02/Fc was prepared as described in Example 1a. DNA encoding the Igk signal sequence fused to Exendin-4 was generated by *in vitro* hybridization of the following overlapping and complementary oligonucleotides:

20

5' - CTAGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTG
GGTTCAGGTTCCACCGGTCAC - 3' [SEQ ID NO:16]

25

5' - GGAGAGGGAACCTTCACCAGCGACCTGAGCAAGCAGATGGAGGAGGAGGCCGT
GAGACTG - 3' [SEQ ID NO:17]

5' - TTCATCGAGTGGCTGAAGAACGGAGGACCAAGCAGCGGAGCCCCTCCTCCT
AGC - 3' [SEQ ID NO:18]

30

5' - GAACCTGGAACCCAGAGCAGCAGTACCCATAGCAGGAGTGTGTCTGTCTCCA
TGGTGG - 3' [SEQ ID NO:19]

35

5' - CTCCTCCTCCATCTGCTTGCTCAGGTCGCTGGTGAAGGTTCCCTCTCCGTGA
CCGGTG - 3' [SEQ ID NO:20]

- 68 -

5' - GCTAGGAGGAGGGGCTCCGCTGCTTGGTCCTCCGTTCTTCAGCCACTCGAT
GAACAGTCTCACGGC - 3' [SEQ ID NO:21]

The hybridization reaction was carried out as described
5 in Example 1a. The hybridized product was ligated to the
pJB02 vector which had been digested with *NheI* and *Eco47III*
as described in Example 1a to create pJB02-Exendin-4-Fc.

Example 1e Construction of DNA encoding Exendin-4-HSA:

10 The plasmid pJB02-HSA was prepared as described in
Example 1b. DNA encoding the Igk signal sequence fused to
Exendin-4 was generated by *in vitro* hybridization of the
same overlapping and complementary oligonucleotides
described in Example 1d. Hybridization reactions were also
15 carried out as described above. DNA was cloned into unique
NheI and *FspI* sites in pJB02-HSA to create pJB02-Exendin-4-
HSA.

Example 1f Construction of DNA encoding Exendin-4-linker-
20 HSA:

The plasmid pJB02-linker-HSA was constructed as
described in Example 1c. DNA encoding the Igk signal
sequence fused to Exendin-4 and the 5' part of the linker
sequence was generated as in Example 1d. This DNA was
25 cloned into unique *NheI* and *BspEI* sites in pJB02-linker-HSA
to create pJB02-Exendin-4-linker-HSA.

Example 1g Construction of DNA encoding Val⁸-GLP-1/C-Ex-Fc:

The plasmid pJB02-Exendin-4-Fc was prepared as
30 described in Example 1d. The Exendin-4 encoding DNA was
excised from the vector with *AgeI* and *Eco47III*. The Val⁸-
GLP-1/C-Ex encoding DNA was generated by *in vitro*
hybridization of the following overlapping and complementary
oligonucleotides:

35

- 69 -

5' -CCGGTCACGTGGAGGGGCACCTTCACCTCCGACGTGTCCTCCTATCTGGA
GGGCCAGGCCGCCA - 3' [SEQ ID NO:22]

5' - AGGAATTCATCGCCTGGCTGGTGAAGGGCCGGGGCAGCAGCGG
5 AGCCCCTCCTCCTAGC - 3' [SEQ ID NO:23]

5' - CTCCAGATAGGAGGACACGTCGGAGGTGAAGGTGCCCTCCAC
GTGA - 3' [SEQ ID NO:24]

10 5' - GCTAGGAGGAGGGGCTCCGCTGCTGCCCCGGCCCTTCACCAGCCAGGCGA
TGAATTCCTTGGCGGCCTGGCC - 3' [SEQ ID NO:25]

15 The hybridization reaction was carried out as described
in Example 1a. The hybridized product was ligated in place
of Exendin-4 in the pJB02-Exendin-4-Fc expression vector to
create pJB02-Val⁸-GLP-1/C-Ex-Fc.

Example 1h Construction of DNA encoding Val⁸-Glu²²-GLP-1-Fc:

20 The plasmid pJB02-Exendin-4-Fc was prepared as
described in Example 1d. The Exendin-4 encoding DNA was
excised from the vector with *AgeI* and *Eco47III*. The Val⁸-
Glu²²-GLP-1 encoding DNA was generated by *in vitro*
hybridization of the following overlapping and complementary
oligonucleotides:

25 5' -CCGGTCACGTGGAGGGGCACCTTCACCTCCGACGTGTCCTCCTATCTCGA
GGAGCAGGCCGCCA - 3' [SEQ ID NO:26]

5' - AGGAGTTCATCGCCTGGCTGGTGAAGGGCCGGGGC - 3' [SEQ ID NO:27]

30 5' - GCCCCGGCCCTTCACCAGCCAGGCGATGAACTCCTTGGCGGCC
TGCTC - 3' [SEQ ID NO:28]

5' - CTCGAGATAGGAGGACACGTCGGAGGTGAAGGTGCCCT
35 CCACGTGA - 3' [SEQ ID NO:29]

- 70 -

The hybridization reaction was carried out as described in Example 1a. The hybridized product was ligated in place of Exendin-4 in the pJB02-Exendin-4-Fc expression vector to create pJB02-Val⁸-Glu²²-GLP-1-Fc.

Example 1i Construction of DNA encoding Val⁸-Glu²²GLP-1/C-Ex-Fc:

The plasmid pJB02-Exendin-4-Fc was prepared as described in Example 1d. The Exendin-4 encoding DNA was excised from the vector with AgeI and Eco47III. The Val⁸-Glu²²GLP-1/C-Ex encoding DNA was generated by in vitro hybridization of the following overlapping and complementary oligonucleotides:

5' - CCGGTCACGTGGAGGGCACCTTCACCTCCGACGTGTCCTCCTATCTCGA
GGAGCAGGCCGCCA - 3' [SEQ ID NO:30]

5' - AGGAATTCATCGCCTGGCTGGTGAAGGGCCGGGCAGCAGCGGA
GCCCCCTCCTCCTAGC - 3' [SEQ ID NO:31]

5' - CTCGAGATAGGAGGACACGTCGGAGGTGAAGGTGCCC
TCCACGTGA - 3' [SEQ ID NO:32]

5' - GCTAGGAGGAGGGGCTCCGCTGCTGCCCCGGGCCCTTCACCAGCCAGGCCGA
TGAATTCCTTGGCGGCCTGCTC - 3' [SEQ ID NO:33]

The hybridization reaction was carried out as described in Example 1a. The hybridized product was ligated in place of Exendin-4 in the pJB02-Exendin-4-Fc expression vector to create pJB02-Val⁸-Glu²²-GLP-1/C-Ex-Fc.

Example 1j Construction of DNA encoding Gly⁸-GLP-1-Fc:

The plasmid pJB02-Exendin-4-Fc was prepared as described in Example 1d. The Exendin-4 encoding DNA was

- 71 -

excised from the vector with *AgeI* and *Eco47III*. The Gly⁸-GLP-1 encoding DNA was generated by *in vitro* hybridization of the following overlapping and complementary oligonucleotides:

5

5' - CCGGTCACGGCGAGGGGCACCTTCACTAGTGACGTGTCCTCCTATCTGGA
GGGCCAGGCCGCCA - 3' [SEQ ID NO:34]

5' - AGGAGTTCATCGCCTGGCTGGTGAAGGGCCGGGGC - 3' [SEQ ID NO:35]

10

5' - CTCCAGATAGGAGGACACGTCACTAGTGAAGGTGCCCTC
GCCGTGA - 3' [SEQ ID NO:36]

15

5' - GCCCCGGCCCTTCACCAGCCAGGCGATGAACTCCTTGGCGGC
CTGGCC - 3' [SEQ ID NO:37]

20

The hybridization reaction was carried out as described in Example 1a. The hybridized product was ligated in place of Exendin-4 in the pJB02-Exendin-4-Fc expression vector to create pJB02-Gly⁸-GLP-1-Fc.

Example 2: Expression of heterologous fusion proteins

25

Expression of the fusion proteins encoded by the DNA constructs of Example 1 was carried out by transiently transfecting HEK 293EBNA cells (both adherent and suspension). Cells were counted and seeded 24 hours prior to transfection. The transfection cocktail was prepared by mixing FuGeneTM6 transfection reagent (Roche Molecular Biochemicals, catalog # 1814443) with OptiMEM (Gibco/BRL) and incubating at room temperature for 5 min at which point DNA was added and the cocktail was incubated for an additional 15 min. Immediately before transfection, fresh growth media was added to the plate. Tables 1 and 2 provide further transfection details.

35

- 72 -

Table 1: Reagents used in transient transfections of 293EBNA cells.

Tissue culture vessel	Number of cells seeded	DNA (μ g)	FuGene (μ l)	OptiMEM media (ml)	Vol. of growth medium (ml)
35 mm	5×10^5	1.5	9	0.102	2
100 mm	2×10^6	12	73	0.820	10
700 cm ² (RB)	2×10^7	65	400	4.0	100

5 Table 2: Media composition

Growth and transfection medium	Harvesting medium
DMEM F12 3:1	Hybritech base
5 % FBS	1 mM Ca ²⁺
20 mM HEPES	20 mM HEPES
2 mM L-glutamine	1 μ g/ml Nuselin (human insulin)
50 μ g/ml geneticin (G418 NEO)	1 μ g/ml human transferrin
50 μ g/ml tobromycin	50 μ g/ml tobromycin

For small-scale transfections (35mm - 10mm vessels), cells were rinsed with PBS and switched to harvesting media 24 hours post- transfection and media was collected and replaced every 24 hours for several days. In the case of large-scale transfections (700 cm² roller bottles), the roller bottles were rinsed with PBS 48 hours post-transfection and changed to harvesting media. Media was collected and changed every 24 hours for at least 10 consecutive days. Routinely, only 10 harvests were used for subsequent protein purification.

- 73 -

Example 3: Purification of heterologous fusion proteinsExample 3a purification of Val⁸-GLP-1-Fc

Approximately 4.5 liters of conditioned medium (fusion protein expression level approximately 20 µg/ml) from large-scale transfections was filtered using a CUNO filter system and concentrated to 250 ml using a ProFlux tangential flow filtration system with a 10 K filter membrane. Val⁸-GLP-1-Fc was captured with a 5 ml HiTrap protein A column in 1x PBS, pH 7.4 at a flow rate of 2 ml/min and eluted with 50 mM citric acid pH 3.3. Fractions (1 ml) were collected in tubes containing 4 ml 1x PBS and 100µl 1 M Tris pH 8.

Fractions containing the fusion protein, as determined by SDS-PAGE and reverse phase-HPLC on Zorbax C8, were pooled and applied to a Superdex 75 60/60 column in 1x PBS pH 7.4 at a flow rate of 10 ml/min. Positive fractions (20 mls/tube) were collected and pooled. Pooled fractions were then subjected to C4 Reverse Phase Chromatography in 0.1 %TFA water at a flow rate of 3 ml/min. Val⁸-GLP-1-Fc was eluted using a gradient from 5% B (0.1% TFA in acetonitrile) to 100% B in 70 min. Eluant fractions (3 mls/tube) were collected. Acetonitrile was removed by vacuum drying and 1 ml of H₂O was added. The purified sample (approximately 32 mls) was dialyzed twice against 4 liters of 1x PBS pH7.4.

The dialyzed sample was then filtered using a MILLEX-GV 0.22 µm Filter Unit and concentration was determined using absorption at 280 nm.

Example 3b purification of Val⁸-GLP-1-HSA or Val⁸-GLP-1-Linker-HSA

Approximately 6.5 liters of conditioned medium (fusion protein expression level approximately 10µg/ml) was filtered using a CUNO filter system and concentrated to 380 mls using a ProFlux tangential flow filtration system with a 10 K filter membrane.

- 74 -

The fusion protein was captured using a 50 ml Fast Flow Q column (Pharmacia) in 20 mM Tris pH 7.4 at a flow rate of 5ml/min. Protein was eluted using a gradient: from 0% to 50% 20mM Tris pH 7.4, 1M NaCl in 10 CV, then to 100%B in 2 CV.

5 Fractions containing the fusion protein were pooled and subjected to C4 Reverse Phase Chromatography in 0.1% TFA water at a flow rate of 5 ml/min. The fusion protein was eluted using a gradient from 20% B (0.1% TFA in acetonitrile) to 90% B in 120min. Fractions (3.5 ml/tube)
10 were collected. Acetonitrile was removed by vacuum drying.

Approximately 9 mls of pooled sample was diluted with 1x PBS pH 7.4 to 40ml and dialyzed against 4 liters of 1x PBS pH 7.4 overnight. The sample was filtered and concentration was determined by absorption at 280nm.

15

Example 3c purification of Exendin-4-Fc:

Approximately 4 liters of conditioned medium (fusion protein expression level approximately 8 µg/ml) was filtered using a CUNO filter system and concentrated to 250 mls using
20 a ProFlux tangential flow filtration system with a 30K filter membrane.

Exendin-4-Fc was captured with a 5 ml HiTrap protein A column in 1x PBS, pH 7.4 at a flow rate of 2 ml/min and eluted with 50 mM citric acid pH 3.3. Fractions containing
25 the fusion protein were pooled, filtered, and dialyzed against 4 liters of 1 x PBS over night. The dialyzed sample was then applied to a Superdex 75 60/60 column in 1x PBS pH7.4, 0.5M NaCl at a flow rate of 10 ml/min. Fractions (20 ml/tube) containing the fusion protein were collected,
30 pooled, and concentrated to about 1 mg/ml. Concentrated samples were then filtered using a MILLEX-GV 0.22 um Filter Unit.

- 75 -

Example 3d purification of Exendin-4-HSA and Exendin-4-linker-HSA:

Approximately 1.1 liters of conditioned medium (fusion protein expression level approximately 6 μ g/ml) was filtered using a CUNO filter system and concentrated to 175 mls using a ProFlux tangential flow filtration system with a 30K filter membrane.

The fusion protein was captured using a 5 ml HiTrap Q-sepharose column (Pharmacia) in 20 mM Tris pH 7.4 at a flow rate of 2 ml/min. Protein was eluted using a gradient from 0% to 50% 20mM Tris pH 7.4, 1M NaCl in 12 CV and then to 100%B in 4 CV.

Fractions containing the fusion protein were pooled and subjected to C4 Reverse Phase Chromatography in 0.1% TFA water at a flow rate of 5 ml/min. The fusion protein was eluted using a gradient from 10% B (0.1% TFA in acetonitrile) to 100% B in 70 min. Fractions (10 ml/tube) containing the fusion protein were collected. Acetonitrile was removed using a vacuum dryer.

Approximately 8 mls of pooled sample was dialyzed against 4 liters of 1x PBS pH 7.4 overnight. The sample was filtered and concentration was determined by absorption at 280nm. The dialyzed sample was then applied to a Superdex 200 26/60 column in 1x PBS pH 7.4, 0.5 M NaCl at a flow rate of 2 ml/min. Fractions (3 ml/tube) containing the fusion protein were collected, pooled, concentrated, and filtered.

Example 4: Characterization of fusion proteins by SDS PAGE:

SDS-PAGE followed by immunoblotting was used to analyze both purified fusion protein as well as conditioned medium from cells transfected with various fusion protein expression vectors. SDS-PAGE was performed on a Novex Powerease 500 system using Novex 16% Tris-Glycine Precast gels (EC6498), running buffer (10x, LC2675) and sample

- 76 -

buffer (L2676). Samples were reduced with 50 mM DTT and heated 3-5 min at 95°C prior to loading.

After running the SDS-PAGE gel, water and transfer buffer (1X Tris-Glycine Seprabuff (Owl Scientific Cat. No. ER26-S) with 20% methanol) were used to rinse SDS from the gels. A Novex transfer apparatus was used with PVDF (BioRad, Cat. No. 162-0174) and nitrocellulose membranes (BioRad, Cat. No. 1703965 or 1703932). Transfer was carried out at room temperature for 90 min at 30-35 V. Membranes were blocked in 1X PBS with 0.1% Tween-20 (Sigma, Cat. No. P-7949) and 5% Milk (BioRad, Cat. No. 170-6404) for 1-12 hours at 4°C. Antibodies are diluted into 1X PBS +5% Milk and the blots are incubated in these solutions for 1-2 h at 4°C. Between incubations, the blots are washed four times for 5 min each with 1X PBS and 0.2% Tween-20 at room temperature. PBS was made from either GIBCO 10X PBS (Cat No. 70011), to give a final composition of 1 mM monobasic potassium phosphate, 3 mM dibasic sodium phosphate, 153 mM sodium chloride, pH 7.4, or PBS pouches from Sigma (Cat. No. 1000-3), to give 120 mM NaCl, 2.7 mM KCl and 10 mM phosphate, pH 7.4 at 25°C.

The primary antibodies were either a polyclonal goat anti-IgG1 or rabbit anti-HSA. The secondary antibody was either an anti-goat IgG HRP or an anti-rabbit IgG HRP. The secondary antibody was diluted 1:5000. An ECL system (Amersham Pharmacia Biotech, Cat. No. RN2108 and Cat. No. RPN1674H) was used for developing blots.

Figure 3A compares purified Fc protein to conditioned media from pJB02-Val⁸-GLP-1-Fc and pJB02-Exendin-4-Fc transfected cells. The decrease in mobility is consistent with the increased size due to the GLP-1 portion of the fusion protein. Figure 3B similarly compares purified HSA with conditioned media from cells transfected with pJB02-Val⁸-GLP-2-HSA, pJB02-Val⁸-GLP-1-Linker-HSA, pJB02-Exendin-

- 77 -

4-HSA, or pJB02-Exendin-4-Linker-HSA. Figure 4 identifies purified fusion protein preparations.

Example 5: Characterization of fusion
proteins using mass spectrometry:

5

All experiments were performed on a Micromass TofSpec-2E mass spectrometer equipped with Time Lag Focusing electronics, a Reflectron (used in analysis of the 0-8000 Da peptide range), a Linear detector (used during high mass / good signal analysis), and Post Acceleration Detector (or P.A.D., used for high mass / extremely low signal analysis). The effective path length of the instrument in Linear mode is 1.2 meters, in Reflectron mode it is 2.3 meters. Two dual micro-channel plate detectors are fitted for linear and reflectron mode detection. The laser used is a Laser Science Inc. VSL-337i nitrogen laser operating at 337 nm at 5 laser shots per second. All data were acquired using a 2 GHz, 8 bit internal digitizer and up to 50 laser shots were averaged per spectrum.

10

15

20

25

The instrument was operated in linear mode for the analysis of the GLP-1 fusion proteins in question. The linear detector is a device that detects ions that travel down the flight tube of the MALDI-ToF-MS instrument. It measures the ion abundance over time and sends a signal to the digitizer for conversion. The digitizer is an analog-to-digital converter that allows the signal from the mass spectrometer to be transferred to the computer, where it is reconstructed into a usable m/z spectrum.

30

35

A recrystallized saturated sinapinic acid solution (diluted in 50/50 Acn / H₂O and 0.1% TFA) was utilized as the ionization matrix. Sinapinic acid is a proper matrix for proteins above 10 kDa. Mass appropriate reference proteins were used for internal and external calibration files in order to obtain accurate mass determinations for the samples analyzed. Samples were all analyzed using a 1:2 sample to

- 78 -

matrix dilution. The instrument was initially set up under the following linear detector conditions:

Source Voltage: 20.0 keV Pulse Voltage: 3.0 keV
 Extraction Voltage: 20.0 keV Laser Coarse: 50
 Focus Voltage: 16.0 keV Laser Fine: 50
 Linear detector: 3.7 keV
 P.A.D.: (off line)

These settings were modified (as needed) to give the best signal/noise ratio and highest resolution. Table 3 provides a characterization of different GLP-1 fusion proteins.

Table 3

Fusion Protein	Expected Mass (KDa)	Mass determined by Mass Spec (kDa)
Val ⁸ -GLP-1-IgG1	59.08	61.94
Val ⁸ -Glu ²² -GLP-1-IgG1	59.23	63.61
Gly ⁸ -GLP-1-IgG1	59.00	62.93
Val ⁸ -GLP-1-CEX-IgG1	60.45	65.1-65.6
Val ⁸ -Glu ²² -GLP-1-CEX-IgG1	60.69	65.86
Exendin-4-IgG1	60.69	65.86
Val ⁸ -GLP-1-Linker-HSA	70.70	69.89, 70.74
Exendin-4-HSA	70.56	70.62
Exendin-4-Linker-HSA	71.56	71.62

CEX refers to a C-terminal extension and comprises the sequence of Ser-Ser-Gly-Ala-Pro-Pro-Pro-Ser.

Linker is Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Ser.

- 79 -

Example 6: Activity of heterologous fusion proteins:

The ability of the fusion proteins of the present invention to activate the GLP-1 receptor was assessed using *in vitro* assays such as those described in EP 619,322 to Gelfand, et al., and U.S. Patent No. 5,120,712, respectively. The activity of these compounds relative to the activity of Val⁸-GLP-1(7-37)OH is reported in Table 4. Figure 8 represents *in vitro* dose response curves for Val⁸-GLP-1 and Exendin-4 fusion proteins. In addition, Table 5a and 5b provide the *in vitro* activity of a large group of GLP-1 analogs that can be fused to an Fc or an albumin protein to make biologically active fusion proteins. These activities are compared to GLP-1(7-37)OH.

Table 4: In vitro activity of GLP-1 fusion proteins

Fusion Protein	In Vitro Activity (% of Val ⁸ -GLP-1)
Val ⁸ -GLP-1-IgG1	1
Exendin-4-IgG1	240
Val ⁸ -GLP-1-Linker-HSA	0.2
Exendin-4-HSA	20
Exendin-4-Linker-HSA	90
Exendin-4	500
Val ⁸ -Glu ²² -GLP-1-IgG1	3.7
Gly ⁸ -GLP-1-IgG1	3.3
Val ⁸ -GLP-1-CEX-IgG1	3.3
Val ⁸ -Glu ²² -GLP-1-CEX-IgG1	29
Gly ⁸ -Glu ²² -GLP-1-C2-IgG1	75
Gly ⁸ -Glu ²² -GLP-1-CEX-Linker-IgG1	150
Exendin-4-C2-IgG1	250
Exendin-4-Linker-IgG1	330
Gly ⁸ -Glu ²² -GLP-1-CEX-Linker-HSA	4
Gly ⁸ -Glu ²² -GLP-1-CEX-Linker-IgG4	80

- 80 -

CEx refers to a C-terminal extension and comprises the sequence of Ser-Ser-Gly-Ala-Pro-Pro-Pro-Ser. Linker is Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Ser

5 C2 is Ser-Ser-Gly-Ala-Ser-Ser-Gly-Ala.

The amino acid sequences of the fusion proteins described in Tables 3 and 4 are represented in SEQ ID NO: 13 to SEQ ID NO: 31.

Val⁸-GLP-1-Human serum albumin amino acid sequence is

10 represented by SEQ ID NO: 13.

1 HVEGTFTSDV SSYLEGQAAK EFIAMLVKGR GDAHKSEVAV RFDLGEENF KALVLIAPAQ
61 YLQOCPPFDH VKLVNEVTEF AKTCVADESA ENCDKSLHTL PGDKLCTVAT LRETYGENAD
121 CCAKQEPERN ECFLQHKDDN PNLPRLVPE VDVMTAFHD WEETFLKKYL YELARRHPYF
181 YAPELLFFAK RYKAAFTCEC QAADKAACIL PKLDELDEG KASSAKQRLK CASLQKFGKR
15 241 AFKAWAVARL SQRFPKAEFA EVSKLVTDLT KVHTECCHGD LLECADDRAD LAKYICENQD
301 SISKLEKCEC EKPLLEKSHC LAEVENDENP ADLPSLAADF VESKDVCKNY AEAKDVFLGM
361 FLYEYARRHP DYSVVLLRL ARTYETTLEK CCAAADPHEC YAKVFDEFKP LVEEPQNLIK
421 QNCELFEQLG EYKQWALLV RYTKKVPQVS TPTLVEVSRN LGKVGSKCK HPEAKRMPCA
481 EDYLSVVLNQ LCVLHEKTPV SDRVTKCTE SLVNRPCFS ALEVDETYVP KEFNAETFTF
20 541 HADICTLSEK ERQIKQTAL VELVGHKPKA TKEQLKAVMD DFAAFVERCC KADDKETCFA
601 EGGKKLVAAS QAALGL [SEQ ID NO: 13]

Val⁸-GLP-1-Linker-Human serum albumin amino acid sequence is represented by SEQ ID NO: 14.

25 1 HVEGTFTSDV SSYLEGQAAK EFIAMLVKGR GGGGGSGGGG SGGGGSDAHK SEVAHRFKDL
61 GEENFKALVL IAPAYLQOC PFEDHVKLVN EVTEFAKTCV ADESAENCDK SLHTLPGDKL
121 CTVATLRETY GEMADCCAKQ EPERNECFLO HKDDNPNLPR LVRPEVDVMC TAFHDNEETF
181 LKKYLYELAR RHPYFYAPEL LFFAKRYKAA FTECCQAADK AACLLPKLDE LRDEGKASSA
241 KQRLKASLQ KGERAFKAW AVARLSQRFP KAFAEVSKL VTDLTKVHTE CCHGDILLECA
30 301 DDRADLAKYI CENQDSISSK LKECCEKPLL EKSHCIAEVE NDENPADLPS LAADFVESKD
361 VCKNYAEAKD VFLGMFLY EYARRHPDYSVV LLLRLAKTYE TTLEKCCAAA DPHECYAKVF
421 DEFKPLVEEP QNLIKQNC ELFEQLGEYKFO NALLVRYTKK VPQVSTPTLV EVSRNLGKVG
481 SKCKHPEAK RMPCAEDYLS VVLNQLCVLH EKTVPVSDRVT KCCTESLVNR RPCFSALEVD
541 ETVVPKEFMA ETFTPHADIC TLSEKERQIK QQTALVELVK HKPKATKEQL KAVMDFAAF
35 601 VERCKKADDK ETCFAEGKK LVAASQAALG L [SEQ ID NO: 14]

- 81 -

Gly⁸-Glu²²-GLP-1-CEx-Linker-Human serum albumin amino acid sequence is represented by SEQ ID NO: 15.

```

1 HEGTFTSDV SSYLEEQAQK EFLANLVRGR GSSGAPPPSG GGGSGGGGS GGGGSDAHKS
61 EVAHRFKDLG EENFKALVLI AFAQYLQQCP FEDHVKLVNE VTEFAKTCVA DESAENCCKS
5 121 LHTLFGDKLC TVATLRETYG EMADCCAKQE PERNECFLOH KDDNPNLPRL VRPEVDVMCT
181 AFHDNEETFL KKLYEELARR HPYFYAPELL FFAKRYKAAF TECCQAADKA ACILPKLDEL
241 RDEGKASSAK QRLKASLQK FGERAFKAMA VARLSQRFPK AEFAEVSKLV TDLTKVHTEC
301 CHGDLLECAD DRADLAKYIC ENQDSISSKL KECCEKPLLE KSHCIAEVEN DEMPADLPSL
361 AADFVESKDV CKNYAEAKDV FLGMFLYEYA RRHPDYSVVL LLRLAKTYET TLEKCCAAAD
10 421 PHECYAKVPD EFKPLVEEPQ NLIKQNCLEF EQLGEYKQFN ALLVRYTKKV PQVSTPTLVE
481 VSRNLGKVGK KCKKHPEAKR MPCAEDYLSV VLNQLCVLHE KTPVSDRVTK CCTESLVNRR
541 PCPSALEVDE TYVPKEFNAE TTFPHADICT LSEKERQIKK QTALVELVKH KPKATKEQLK
601 AVMDDFAAVF EKCKADDKE TCFABEGKKL VAASQAALGL [SEQ ID NO: 15]

```

15 Exendin-4-Human serum albumin amino acid sequence is represented by SEQ ID NO: 16.

```

1 HEGTFTSDL SKQMEEEAVR LFIWLNKNGG PSSGAPPPSD AHKSEVAHRF KDLGEENFKA
61 LVLIAFAQYL QQCFFEDHVK LVNEVTEFAK TCVADESAEN CDKSLHTLFG DKLCTVATLR
121 ETYGENADCC AKQEPERNEC FLQHKDDNPN LPRLVRPEVD VMCTAFHDNE ETFLKKLYE
20 181 IARRHPYFYA PELLFFAKRY KAAFTECCQA ADKAACLLPK LDELRLDEGKA SSKAQRLLKA
241 SLQKFGERAF KAMAVARLSQ RFPKAEFAEV SKLVTDLTKV HTECCHGDLL ECADDRADLA
301 KYICENQDSI SSKLKECCEK PILKSHCIA EVENDENPAD LPSLAADFVE SKDVCKNYAE
361 AKDVFLGMFL YEYARRHPDY SVVLLLRLAK TYETTLKCC AAADPHECYA KVFDEFKPLV
421 EEPQNLIKQ NCLEFQELGEY KQFNALLVRY TKKVPQVSTP TLVEVSRNLG KVGSKCKHPE
25 481 EAKRMPCAED YLSVVLNQLC VLHEKTPVED RVTKCCTESL VNRRPCPSAL EVDETYVPKE
541 FNAETTFPHA DICTLSEKER QIKKQTALVE LVKHKPKATK EQLKAVMDDF AAFVEKCKKA
601 DDKETCFABE GKILVAASQA ALGL [SEQ ID NO: 16]

```

30 Exendin-4-Linker-Human serum albumin amino acid sequence is represented by SEQ ID NO: 17.

```

1 HEGTFTSDL SKQMEEEAVR LFIWLNKNGG PSSGAPPPSG GGGSGGGGS GGGGSDAHKS
61 EVAHRFKDLG EENFKALVLI AFAQYLQQCP FEDHVKLVNE VTEFAKTCVA DESAENCCKS
121 LHTLFGDKLC TVATLRETYG EMADCCAKQE PERNECFLOH KDDNPNLPRL VRPEVDVMCT
181 AFHDNEETFL KKLYEELARR HPYFYAPELL FFAKRYKAAF TECCQAADKA ACILPKLDEL
35 241 RDEGKASSAK QRLKASLQK FGERAFKAMA VARLSQRFPK AEFAEVSKLV TDLTKVHTEC
301 CHGDLLECAD DRADLAKYIC ENQDSISSKL KECCEKPLLE KSHCIAEVEN DEMPADLPSL
361 AADFVESKDV CKNYAEAKDV FLGMFLYEYA RRHPDYSVVL LLRLAKTYET TLEKCCAAAD
421 PHECYAKVPD EFKPLVEEPQ NLIKQNCLEF EQLGEYKQFN ALLVRYTKKV PQVSTPTLVE

```

- 82 -

481 VSRNLGKVGK KCKKHPEAKR MPCEADYLSV VLNQLCVLHE KTFVSDRVTK CCTESLVNRR
 541 PCFSALEVDK TYVPKEFNAE TPTFHADICT LSEKERQIKK QTALVELVKH KPKATKEQLK
 601 AVMDDFAAAFV EKCKKADDEK TCFAEEGKKL VAASQAALGL [SEQ ID NO: 17]

5 Val⁸-GLP-1-IgG1 amino acid sequences represented by SEQ ID NO: 18.

1 HVEGTFTSDV SSYLEGQAAK EFIANLVKGR GAEPKSCDKT HTCPPCPAPE LLGGPSVFLF
 61 PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VDMAKTKPRE EQYNSTYRVV
 121 SVLTVLHQDW LNKKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SREMTKNQV
 10 181 SLTCLVKGFPY PSDIAVEWES NGQPENNYKT TPPVLDSGGS FFLYSKLTVD KSRNQOGNVF
 241 SCSVMHEALH NHYTQKSLSL SPGK [SEQ ID NO: 18]

Val⁸-GLP-1-Cex-IgG1 amino acid sequence is represented by SEQ ID NO: 19.

15 1 HVEGTFTSDV SSYLEGQAAK EFIANLVKGR GSSGAPPPSA EPKSCDKTHT CPPCPAPELL
 61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
 121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR
 181 EEMTKNQVSL TCVKGFPYPS DIAVEWESNG QPENNYKTTT FVLDSGGSFF LYSKLTVDKS
 241 RQOQGNVFSV SVMHEALHNN YTQKSLSLSP GK [SEQ ID NO: 19]

20

Val⁸-Glu²² GLP-1-IgG1 amino acid sequence is represented by SEQ ID NO: 20.

1 HVEGTFTSDV SSYLEGQAAK EFIANLVKGR GAEPKSCDKT HTCPPCPAPE LLGGPSVFLF
 61 PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VDMAKTKPRE EQYNSTYRVV
 25 121 SVLTVLHQDW LNKKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SREMTKNQV
 181 SLTCLVKGFPY PSDIAVEWES NGQPENNYKT TPPVLDSGGS FFLYSKLTVD KSRNQOGNVF
 241 SCSVMHEALH NHYTQKSLSL SPGK [SEQ ID NO: 20]

Val⁸-Glu²² GLP-1-CEX-IgG1 amino acid sequence is represented by SEQ ID NO: 21.

30 1 HVEGTFTSDV SSYLEGQAAK EFIANLVKGR GSSGAPPPSA EPKSCDKTHT CPPCPAPELL
 61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
 121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR
 181 EEMTKNQVSL TCVKGFPYPS DIAVEWESNG QPENNYKTTT FVLDSGGSFF LYSKLTVDKS
 35 241 RQOQGNVFSV SVMHEALHNN YTQKSLSLSP GK [SEQ ID NO: 21]

- 83 -

Gly⁸-Glu²² GLP-1-C2-IgG1 amino acid sequence is represented by SEQ ID NO: 22.

1 HEGGTFTSDV SSYLEEQAAK EFLAWLVKGR GSSGASSGAA EPKSCDKTHT CPPCPAPELL
61 GGPSVFLPPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
5 121 YNSTYRVVSV LTVLHQDWLN GKEYCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR
181 EEMTKNQVSL TCLVKGFYPS DIAVENESNG QPENNYKTFP FVLDSGDSFF LYSKLTVDKS
241 RWQOGNVFSC SVMHEALHNH YTQKSLSLSP GK [SEQ ID NO: 22]

Gly⁸-Glu²² GLP-1-CEx-Linker-IgG1 amino acid sequence is represented by SEQ ID NO: 23.

1 HEGGTFTSDV SSYLEEQAAK EFLAWLVKGR GSSGAPPPSG GGGSGGGGSG GGGSAEPKSC
61 DKTHTCPPCP APELLGGPSV FLFPPKPKDT LMISRTPEVT CVVVDVSHED PEVKFNWYVD
121 GVEVHNATK PREEQYNSTY KVSVLTVLH QDWLNGKEYK CKVSNKALPA PIEKTISKAK
181 GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTFPFVLDS
15 241 DGSFFLYSKL TVDKSRWQOG NVFSCSVME ALHNHYTQKS LSLSPGK [SEQ ID NO: 23]

Gly⁸-Glu²² GLP-1-CEx-Linker-IgG4 amino acid sequence is represented by SEQ ID NO: 24.

1 HEGGTFTSDV SSYLEEQAAK EFLAWLVKGR GSSGAPPPSG GGGSGGGGSG GGGSAESKYG
20 61 PPCPSCPAPE FLGGPSVFLF PPKPKDTLMI SRTPEVTCVV VDVSDQEDPEV QFNWYVDGVE
121 VHNATKPRE EQFNSTYRVV SVLTVLHQDW LNKKEYCKV SNKGLPSSIE KTISKAKGQP
181 REPQVYTLFP SQEEMTKNQV SLTCLVKGFY PSDIAVENES NGQPENNYKT TPFVLDSGDS
241 PFLYSRLTVD KSRWQEGNVF SCSVMHEALH NHYTQKSLSL SLGK [SEQ ID NO: 24]

Gly⁸-Glu²² GLP-1-CEx-2Linker-IgG1 amino acid sequence is represented by SEQ ID NO: 25.

1 HEGGTFTSDV SSYLEEQAAK EFLAWLVKGR GSSGAPPPSG GGGSGGGGSG GGGSGGGGSG
61 GGGSGGGGSA EPKSCDKTHT CPPCPAPELL GGPSVFLPPP KPKDTLMISR TPEVTCVVVD
121 VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ YNSTYRVVSV LTVLHQDWLN GKEYCKVSN
30 181 KALPAPIEKT ISKAKGQPRE PQVYTLPPSR EEMTKNQVSL TCLVKGFYPS DIAVENESNG
241 QPENNYKTFP FVLDSGDSFF LYSKLTVDKS RWQOGNVFSC SVMHEALHNH YTQKSLSLSP
301 GK [SEQ ID NO: 25]

Gly⁸-Glu²² GLP-1-2Linker-IgG1 amino acid sequence is represented by SEQ ID NO: 26.

- 84 -

1 HEGTFTSDV SSYLEEQAAK EFIANLVKGR GGGGGSGGGG SGGGGSGGGG SGGGGSGGGG
 61 SAEPKSCDKT HTCPPCPAPE LLGGPSVFLF PPKFKDTLMI SRTPEVTCVV VDVSHEDPEV
 121 KFNWYVDGVE VHNATKPRE EQYNSTYRVV SVLTVLHQDW LNGKEYKCKV SNKALPAPIE
 181 KTISKAKGQP REPQVYTLPP SREMTKNQV SLTCLVKGFY PSDIAVENES NGQPENNYKT
 5 241 TFPVLDSDGS FFLYSKLTVD KSRWQQGNV SCSVMHEALH NHYTQKSLSL SPGK
 [SEQ ID NO: 26]

Gly⁸-Glu²² GLP-1-2CEX-IgG1 amino acid sequence is represented
 by SEQ ID NO: 27.

10 1 HEGTFTSDV SSYLEEQAAK EFIANLVKGR GSSGAPPPSS SGAPPPSAEP KSCDKTHTCP
 61 PCPAPELLGG PSVFLFPPKP KDTLMISRTPEVTCVVVDVS HEDPEVKFNW YVDGVEVHNA
 121 KTKPREEQYN STYRVVSVLT VLHQDWLNGK EYKCKVSNKA LPAPIEKTIS KAKGQPREPQ
 181 VYTLPPSRRE MTKNQVSLTC LVKGFYPSDI AVENESNGQP ENNYKTTFPV LDSDGSFFLY
 241 SKLTVDKSRW QQGNVFSCSV MHEALHNHYT QKSLSLSPGK [SEQ ID NO: 27]

15

Gly⁸-Glu²²-Val²⁵-Ile³³ GLP-1-CEX-Linker-IgG1 amino acid
 sequence is represented by SEQ ID NO: 28.

20 1 HEGTFTSDV SSYLEEQAVK EFIANLIKGR GSSGAPPPSG GGGSGGGGSG GGGSAEPKSC
 61 DKHTTCPPCP APELLGGPSV FLPPPKPKDT LMISRTPEVT CVVVDVSHED PEVKFNWYVD
 121 GVEVHNATK PREEQYNSTY RVVSVLTVLH QDWLNGKEYK CKVSNKALPA PIKTISKAK
 181 GQPREPQVYT LPPSRREMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTFPVLDG
 241 DGSFFLYSKL TVDKSRWQQG NVFSCSVME ALHNHYTQKS LSLSPGK [SEQ ID NO: 28]

25 Exendin-4-IgG1 amino acid sequence is represented by SEQ ID
 NO: 29.

30 1 HEGTFTSDL SKQMEEEAVR LFIEMLNKGG PSSGAPPPSA EPKSCDKTHT CPFCPAPELL
 61 GGPVFLPPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NMYVDGVEVH NATKPREEQ
 121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR
 181 EEMTKNQVSL TCLVKGFYPS DIAVENESNG QPENNYKTTT FVLDSDGSFF LYSKLTVDKS
 241 RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GK [SEQ ID NO: 29]

Exendin-4-C2-IgG1 amino acid sequence is represented by SEQ ID NO: 30.

1 HGEGETFTSDL SKQMEKEAVR LFIENLKNGG PSSGASSGAA EPKSCDKTHT CPPCPAPELL
5 61 GGPSVFLPPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR
181 EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTFP PVLDS DGSFF LYSKLTVDKS
241 RWQQGNVFSC SVMHEALHMH YTKSLSLSP GK [SEQ ID NO: 30]

10 Exendin-4-Linker-IgG1 amino acid sequence is represent by SEQ ID NO: 31.

1 HGEGETFTSDL SKQMEERAVR LFIENLKNGG PSSGAPPPSG GGGSGGGGSG GGGSAEPKSC
61 DKTHTCPPCP APELLGGPSV FLFPPKPKDT LMISRTPEVT CVVVDVSHED PEVKFNWYVD
121 GVEVHNATK PREEQYNSTY RVVSVLTVLH QDMLNGKEYK CKVENKALPA PIEKTISKAK
15 181 GQPREPQVYT LPPSRREMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTPPVLDS
241 DGSFFLYSKL TVDKSRWQQG NVFSCSVME ALHNHYTQKS LSLSPGK [SEQ ID NO: 31]

Table 5a: In vitro GLP-1 analog activity

GLP-1 Receptor	
<u>GLP-1 Compound</u>	<u>Activation</u>
20 GLP-1(7-37)OH	1.0
Val ⁸ -GLP-1(7-37)OH	0.47 (n = 6)
25 Gly ⁸ -His ¹¹ -GLP-1(7-37)OH	0.282
Val ⁸ -Ala ¹¹ -GLP-1(7-37)OH	0.021
30 Val ⁸ -Lys ¹¹ -GLP-1(7-37)OH	0.001
Val ⁸ -Tyr ¹² -GLP-1(7-37)OH	0.81
Val ⁸ -Glu ¹⁶ -GLP-1(7-37)OH	0.047

- 86 -

	Val ⁸ -Ala ¹⁶ -GLP-1(7-37)OH	0.112
	Val ⁸ -Tyr ¹⁶ -GLP-1(7-37)OH	1.175
5	Val ⁸ -Lys ²⁰ -GLP-1(7-37)OH	0.33
	Gln ²² -GLP-1(7-37)OH	0.42
10	Val ⁸ -Ala ²² -GLP-1(7-37)OH	0.56
	Val ⁸ -Ser ²² -GLP-1(7-37)OH	0.50
	Val ⁸ -Asp ²² -GLP-1(7-37)OH	0.40
15	Val ⁸ -Glu ²² -GLP-1(7-37)OH	1.29
	Val ⁸ -Lys ²² -GLP-1(7-37)OH	0.58
20	Val ⁸ -Pro ²² -GLP-1(7-37)OH	0.01
	Val ⁸ -His ²² -GLP-1(7-37)OH	0.14
	Val ⁸ -Lys ²² -GLP-1(7-36)NH ₂	0.53
25	Val ⁸ -Glu ²² -GLP-1(7-36)NH ₂	1.0
	Gly ⁸ -Glu ²² -GLP-1(7-37)OH	1.07
30	Val ⁸ -Lys ²³ -GLP-1(7-37)OH	0.18
	Val ⁸ -His ²⁴ -GLP-1(7-37)OH	0.007

100
97

- 87 -

	Val ⁸ -Lys ²⁴ -GLP-1(7-37)OH	0.02
	Val ⁸ -His ²⁶ -GLP-1(7-37)OH	1.6
5	Val ⁸ -Glu ²⁶ -GLP-1(7-37)OH	1.5
	Val ⁸ -His ²⁷ -GLP-1(7-37)OH	0.37
	Val ⁸ -Ala ²⁷ -GLP-1(7-37)OH	0.47
10	Gly ⁸ -Glu ³⁰ -GLP-1(7-37)OH	0.29
	Val ⁸ -Glu ³⁰ -GLP-1(7-37)OH	0.29
15	Val ⁸ -Asp ³⁰ -GLP-1(7-37)OH	0.15
	Val ⁸ -Ser ³⁰ -GLP-1(7-37)OH	0.19
	Val ⁸ -His ³⁰ -GLP-1(7-37)OH	0.19
20	Val ⁸ -Glu ³³ -GLP-1(7-37)OH	0.039
	Val ⁸ -Ala ³³ -GLP-1(7-37)OH	0.1
25	Val ⁸ -Gly ³³ -GLP-1(7-37)OH	0.01
	Val ⁸ -Glu ³⁴ -GLP-1(7-37)OH	0.17
	Val ⁸ -Pro ³⁵ -GLP-1(7-37)OH	0.094
30	Val ⁸ -His ³⁵ -GLP-1(7-37)OH	0.41
	Val ⁸ -Glu ³⁵ -GLP-1(7-37)OH	0.15

126
98

- 88 -

	Val ⁸ -Glu ³⁶ -GLP-1(7-37)OH	0.11
	Val ⁸ -His ³⁶ -GLP-1(7-37)OH	0.22
5	Val ⁸ -His ³⁷ -GLP-1(7-37)OH	0.33
	Val ⁸ -Leu ¹⁶ -Glu ²⁶ -GLP-1(7-37)OH	0.23
10	Val ⁸ -Lys ²² -Glu ³⁰ -GLP-1(7-37)OH	0.37
	Val ⁸ -Lys ²² -Glu ²³ -GLP-1(7-37)OH	0.35
	Val ⁸ -Glu ²² -Ala ²⁷ -GLP-1(7-37)OH	1.02
15	Val ⁸ -Glu ²² -Lys ²³ -GLP-1(7-37)OH	1.43
	Val ⁸ -Lys ³³ -Val ³⁴ -GLP-1(7-37)OH	0.08
20	Val ⁸ -Lys ³³ -Asn ³⁴ -GLP-1(7-37)OH	0.09
	Val ⁸ -Gly ³⁴ -Lys ³⁵ -GLP-1(7-37)OH	0.34
	Val ⁸ -Gly ³⁶ -Pro ³⁷ -GLP-1(7-37)NH ₂	0.53

25

Table 5b. *In vitro* GLP-1 analog activity

GLP-1 Compound	GLP-1 Receptor Activation
GLP-1(7-37)OH	1.0
Val ⁸ -GLP-1(7-37)OH	0.47
Gly ⁸ -GLP-1(7-37)OH	0.80
Val ⁸ -Tyr ¹² -GLP-1(7-37)OH	0.80
Val ⁸ -Tyr ¹² -GLP-1(7-36)NH ₂	0.52
Val ⁸ -Trp ¹² -GLP-1(7-37)OH	0.52
Val ⁸ -Leu ¹⁶ -GLP-1(7-37)OH	0.52
Val ⁸ -Val ¹⁶ -GLP-1(7-37)OH	0.52
Val ⁸ -Tyr ¹⁶ -GLP-1(7-37)OH	1.18
Gly ⁸ -Glu ²² -GLP-1(7-37)OH	1.03
Val ⁸ -Leu ²⁵ -GLP-1(7-37)OH	0.24
Val ⁸ -Tyr ¹² -Tyr ¹⁶ -GLP-1(7-37)OH	0.70
Val ⁸ -Trp ¹² -Glu ²² -GLP-1(7-37)OH	0.80
Val ⁸ -Tyr ¹² -Glu ²² -GLP-1(7-37)OH	1.27
Val ⁸ -Tyr ¹⁶ -Phe ¹⁹ -GLP-1(7-37)OH	1.32
Val ⁸ -Tyr ¹⁶ -Glu ²² -GLP-1(7-37)OH	1.69, 1.79
Val ⁸ -Trp ¹⁶ -Glu ²² -GLP-1(7-37)OH	2.30, 2.16
Val ⁸ -Leu ¹⁶ -Glu ²² -GLP-1(7-37)OH	2.02
Val ⁸ -Ile ¹⁸ -Glu ²² -GLP-1(7-37)OH	1.55
Val ⁸ -Phe ¹⁶ -Glu ²² -GLP-1(7-37)OH	1.08
Val ⁸ -Trp ¹⁸ -Glu ²² -GLP-1(7-37)OH	1.50, 3.10
Val ⁸ -Tyr ¹⁸ -Glu ²² -GLP-1(7-37)OH	2.40, 2.77
Val ⁸ -Phe ¹⁸ -Glu ²² -GLP-1(7-37)OH	0.94
Val ⁸ -Ile ¹⁸ -Glu ²² -GLP-1(7-37)OH	1.88
Val ⁸ -Lys ¹⁸ -Glu ²² -GLP-1(7-37)OH	1.18
Val ⁸ -Trp ¹⁹ -Glu ²² -GLP-1(7-37)OH	1.50
Val ⁸ -Phe ¹⁹ -Glu ²² -GLP-1(7-37)OH	0.70
Val ⁸ -Phe ²⁰ -Glu ²² -GLP-1(7-37)OH	1.27

Val ⁸ -Glu ²² -Leu ²⁵ -GLP-1(7-37)OH	1.32
Val ⁸ -Glu ²² -Ile ²⁵ -GLP-1(7-37)OH	1.46
Val ⁸ -Glu ²² -Val ²⁵ -GLP-1(7-37)OH	2.21, 1.36
Val ⁸ -Glu ²² -Ile ²⁷ -GLP-1(7-37)OH	0.94
Val ⁸ -Glu ²² -Ala ²⁷ -GLP-1(7-37)OH	1.03
Val ⁸ -Glu ²² -Ile ³³ -GLP-1(7-37)OH	2.21, 1.79, 1.60
Val ⁸ -Asp ⁹ -Ile ¹¹ -Tyr ¹⁶ -Glu ²² -GLP-1(7-37)OH	2.02
Val ⁸ -Tyr ¹⁶ -Trp ¹⁹ -Glu ²² -GLP-1(7-37)OH	1.64
Val ⁸ -Trp ¹⁶ -Glu ²² -Val ²⁵ -Ile ³³ -GLP-1(7-37)OH	2.35
Val ⁸ -Trp ¹⁶ -Glu ²² -Ile ³³ -GLP-1(7-37)OH	1.93
Val ⁸ -Glu ²² -Val ²⁵ -Ile ³³ -GLP-1(7-37)OH	2.30, 2.73, 3.15
Val ⁸ -Trp ¹⁶ -Glu ²² -Val ²⁵ -GLP-1(7-37)OH	2.07
Val ⁸ -Cys ¹⁶ -Lys ²⁶ -GLP-1(7-37)OH	1.97
Val ⁸ -Cys ¹⁶ -Lys ²⁶ -Arg ³⁴ -GLP-1(7-37)OH	2.4, 1.9

Table 6: In vitro activity of GLP/Exendin analogs

Peptide Sequence	In Vitro Activity (% of Val ⁸ -GLP-1(7-37)OH)
HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGP-NH ₂	6.21
HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPP PS-NH ₂	6.75, 3.25
HVEGTFTSDLSKQMEEEAVRLFIAWLVKGRG	2.86
HVEGTFTSDVSSYLEEEAVRLFIAWLVKGRG	1.47
HVEGTFTSDLSKQMEGQAAKEFIAWLVKGRG	0.11

127
101

- 91 -

HVEGTFTSDVSKQMEGQAAKEFLAWLVKGRG	0.04
HGEGTFTSDLKQMEGQAAKEFLIWLKNGGP-NH2	1.44
HGEGTFTSDLKQMEEEAAKEFLIWLKNGGP-NH2	2.80
HGEGTFTSDVSSYLEEEAVRLFIEWLKNGGP-NH2	5.40
HGEGTFTSDLSSYLEEEAVRLFIEWLKNGGP-NH2	5.07
HAEGTFTSDVSSYLEGQAAKEFLAWLVKGRPSSGAPPPS-NH2	3.30
HAEGTFTSDVSKQLEEEAAKEFLAWLVKGRG	2.15
HVEGTFTSDVSSYLEGQAAKEFLIWLKNGGP-NH2	2.36
HGEGTFTSDLKQMEEEAVRLFIAWLVKGRG	3.25
HVEGTFTSDVSSYLEEEAAKEFLAWLVKGRG	1.00
HVEGTFTSDVSSYLEGQAAKEFLAWLVKGRG	0.20
HVEGTFTSDVSSYLEGQAAKEFLAWLVKGRG	1.00
HAEGTFTSDVSSYLEGQAAKEFLAWLVKGRG	2.12

Example 7: In vivo pharmacokinetics of Val⁸-GLP-1-IgG1 and Val⁸-GLP-1-HSA:

A pharmacokinetic study of Val⁸-GLP-1-IgG1 and Val⁸-GLP-1-HSA was performed in cynomolgus monkeys. Monkeys were dosed at 5.6 nmoles/kg with either purified Val⁸-GLP-1-IgG1 or Val⁸-GLP-1-HSA. The compounds were administered as an intravenous bolus administration. Blood was collected pre-dose and at 0.083, 0.25, 0.5, 1, 4, 8, 12, 24, 48, 72, 96, 120, 144, 168, and 216 hours post-dose into tubes containing EDTA. Plasma concentrations of immunoreactive Val⁸-GLP-1 were determined using a radioimmunoassay that utilizes a polyclonal antiserum whose primary specificity is for the N-terminal (7-16) region of Val⁸-GLP-1(7-37).

Figure 9 depicts the plasma concentration of Val⁸-GLP-1-Fc and Val⁸-GLP-1-Linker-HSA following a single intravenous dose to two cynomolgus monkeys. The Fc fusion protein had a half-life of approximately 45 hours and the albumin fusion had a half-life of approximately 87 hours.

- 92 -

Example 8: In vivo pharmacodynamics of Exendin-4-IgG1:

Two chronically cannulated normal male beagle dogs were studied after an overnight fast. Arterial and venous vascular access ports were accessed, and a catheter was inserted percutaneously into a cephalic vein and secured. Animals were placed in cages, and their catheters were attached to a swivel/tether system. A solution containing the fusion protein Exendin-4-IgG1 (11.8 μ M) was injected intravenously (1.0 nmol/kg) through the cephalic vein catheter. The catheter was then cleared with 10 ml of saline. Two hours later, a hyperglycemic (150 mg/dl) clamp was initiated and continued for three hours. Arterial blood samples were drawn throughout this 5-hour period for determination of plasma concentrations of the fusion protein, glucose, and insulin.

The results of this study were compared to those from a similar, previous study in which both of the animals had received a bolus of saline, s.c., and three hours later were studied using a 3-hour hyperglycemic (150 mg/dl) clamp.

In both sets of studies, plasma glucose concentrations were determined using a Beckman glucose analyzer. Plasma insulin concentrations were determined by employees of Linco Research, Inc. using an RIA kit developed in their laboratories. The data is illustrated in Figures 10 and 11.

Example 9: In vivo pharmacokinetics of
Gly⁸-Glu²²-GLP-1-CEx-Linker-IgG1:

Two groups of three normal male beagle dogs received 0.1 mg/kg of Gly⁸-Glu²²-GLP-1-CEx-Linker-IgG1 by subcutaneous (SC) or intravenous (IV) administration. Plasma concentrations of Gly⁸-Glu²²-GLP-1-CEx-Linker-IgG1 immunoreactivity were determined by radioimmunoassay in samples collected from 30 minutes predose to 216 hours postdose for both the IV and SC groups. These concentrations were subsequently used to determine the

106
103

- 93 -

reported pharmacokinetic parameters. The mean elimination half-life of IV administered Gly⁸-Glu²²-GLP-1-CEx-Linker-IgG1 was approximately 55 hours and the total body clearance was 1.5 mL/h/kg. The mean elimination half-life of SC administered Gly⁸-Glu²²-GLP-1-CEx-Linker-IgG1 was approximately 38 hours.

107
104

- 94 -

WE CLAIM:

1. A heterologous fusion protein comprising a first
5 polypeptide with a N-terminus and a C-terminus fused to a
second polypeptide with a N-terminus and a C-terminus
wherein the first polypeptide is a GLP-1 compound and the
second polypeptide is selected from the group consisting of
a) human albumin;
10 b) human albumin analogs; and
c) fragments of human albumin,
and wherein the C-terminus of the first polypeptide is fused
to the N-terminus of the second polypeptide.
- 15 2. A heterologous fusion protein comprising a first
polypeptide with a N-terminus and a C-terminus fused to a
second polypeptide with a N-terminus and a C-terminus
wherein the first polypeptide is a GLP-1 compound and the
second polypeptide is selected from the group consisting of
20 a) human albumin;
b) human albumin analogs; and
c) fragments of human albumin,
and wherein the C-terminus of the first polypeptide is fused
to the N-terminus of the second polypeptide via a peptide
25 linker.
3. The heterologous fusion protein of the Claim 2 wherein
the peptide linker is selected from the group consisting of:
a) a glycine rich peptide;
30 b) a peptide having the sequence [Gly-Gly-Gly-Gly-Ser]_n
where n is 1, 2, 3, 4, 5 or 6; and
c) a peptide having the sequence [Gly-Gly-Gly-Gly-
Ser]₃.

128
105

- 95 -

4. The heterologous fusion protein of Claims 1, 2, or 3 wherein the GLP-1 compound comprises the sequence of formula 1 [SEQ ID NO: 2]

5 7 8 9 10 11 12 13 14 15 16 17
 His-Xaa-Xaa-Gly-Xaa-Phe-Thr-Xaa-Asp-Xaa-Xaa-
 18 19 20 21 22 23 24 25 26 27 28
 Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Phe-
 29 30 31 32 33 34 35 36 37 38 39
10 Ile-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-
 40 41 42 43 44 45
 Xaa-Xaa-Xaa-Xaa-Xaa-Xaa

Formula I (SEQ ID NO: 2)

15 wherein:
 Xaa at position 8 is Ala, Gly, Ser, Thr, Leu, Ile, Val,
 Glu, Asp, or Lys;
 Xaa at position 9 is Glu, Asp, or Lys;
 Xaa at position 11 is Thr, Ala, Gly, Ser, Leu, Ile,
20 Val, Glu, Asp, or Lys;
 Xaa at position 14 is Ser, Ala, Gly, Thr, Leu, Ile,
 Val, Glu, Asp, or Lys;
 Xaa at position 16 is Val, Ala, Gly, Ser, Thr, Leu,
 Ile, Tyr, Glu, Asp, Trp, or Lys;
25 Xaa at position 17 is Ser, Ala, Gly, Thr, Leu, Ile,
 Val, Glu, Asp, or Lys;
 Xaa at position 18 is Ser, Ala, Gly, Thr, Leu, Ile,
 Val, Glu, Asp, Trp, Tyr, or Lys;
 Xaa at position 19 is Tyr, Phe, Trp, Glu, Asp, Gln, or
30 Lys;
 Xaa at position 20 is Leu, Ala, Gly, Ser, Thr, Ile,
 Val, Glu, Asp, Met, Trp, Tyr, or Lys;
 Xaa at position 21 is Glu, Asp, or Lys;
 Xaa at position 22 is Gly, Ala, Ser, Thr, Leu, Ile,
35 Val, Glu, Asp, or Lys;

129
106

- 96 -

- Xaa at position 23 is Gln, Asn, Arg, Glu, Asp, or Lys;
Xaa at position 24 is Ala, Gly, Ser, Thr, Leu, Ile,
Val, Arg, Glu, Asp, or Lys;
Xaa at position 25 is Ala, Gly, Ser, Thr, Leu, Ile,
5 Val, Glu, Asp, or Lys;
Xaa at position 26 is Lys, Arg, Gln, Glu, Asp, or His;
Xaa at position 27 is Leu, Glu, Asp, or Lys;
Xaa at position 30 is Ala, Gly, Ser, Thr, Leu, Ile,
Val, Glu, Asp, or Lys;
10 Xaa at position 31 is Trp, Phe, Tyr, Glu, Asp, or Lys;
Xaa at position 32 is Leu, Gly, Ala, Ser, Thr, Ile,
Val, Glu, Asp, or Lys;
Xaa at position 33 is Val, Gly, Ala, Ser, Thr, Leu,
Ile, Glu, Asp, or Lys;
15 Xaa at position 34 is Asn, Lys, Arg, Glu, Asp, or His;
Xaa at position 35 is Gly, Ala, Ser, Thr, Leu, Ile,
Val, Glu, Asp, or Lys;
Xaa at position 36 is Gly, Arg, Lys, Glu, Asp, or His;
Xaa at position 37 is Pro, Gly, Ala, Ser, Thr, Leu,
20 Ile, Val, Glu, Asp, or Lys, or is deleted;
Xaa at position 38 is Ser, Arg, Lys, Glu, Asp, or His,
or is deleted;
Xaa at position 39 is Ser, Arg, Lys, Glu, Asp, or His,
or is deleted;
25 Xaa at position 40 is Gly, Asp, Glu, or Lys, or is
deleted;
Xaa at position 41 is Ala, Phe, Trp, Tyr, Glu, Asp, or
Lys, or is deleted;
Xaa at position 42 is Ser, Pro, Lys, Glu, or Asp, or is
30 deleted;
Xaa at position 43 is Ser, Pro, Glu, Asp, or Lys, or is
deleted;
Xaa at position 44 is Gly, Pro, Glu, Asp, or Lys, or is
deleted; and

112
107

- 97 -

Xaa at position 45 is Ala, Ser, Val, Glu, Asp, or Lys,
r is

deleted;

provided that when the amino acid at position 37, 38, 39,
5 40, 41, 42, 43, or 44 is deleted, then each amino acid
downstream of that amino acid is also deleted.

5. The heterologous fusion protein of Claims 1, 2, or 3.
wherein the GLP-1 compound comprises the sequence of formula

10 II (SEQ ID NO: 3):

7 8 9 10 11 12 13 14 15 16 17
Xaa-Xaa-Xaa-Gly-Xaa-Xaa-Thr-Ser-Asp-Xaa-Ser-
18 19 20 21 22 23 24 25 26 27 28
15 Xaa-Tyr-Leu-Glu-Xaa-Xaa-Xaa-Ala-Xaa-Xaa-Phe-
29 30 31 32 33 34 35 36 37
Ile-Xaa-Xaa-Leu-Xaa-Xaa-Xaa-Xaa-Xaa
Formula II (SEQ ID NO: 3)

20 wherein:

Xaa at position 7 is: L-histidine, D-histidine,
desamino-histidine, 2-amino-histidine, β -hydroxy-
histidine, homohistidine, α -fluoromethyl-histidine or
 α -methyl-histidine;

25 Xaa at position 8 is: Gly, Ala, Val, Leu, Ile, Ser, or
Thr;

Xaa at position 9 is: Thr, Ser, Arg, Lys, Trp, Phe,
Tyr, Glu, or His;

Xaa at position 11 is: Asp, Glu, Arg, Thr, Ala, Lys, or His;

30 Xaa at position 12 is: His, Trp, Phe, or Tyr;

Xaa at position 16 is: Leu, Ser, Thr, Trp, His, Phe, Asp,
Val, Tyr, Glu, or Ala;

Xaa at position 18 is: His, Pro, Asp, Glu, Arg, Ser, Ala, or
Lys;

35 Xaa at position 19 is: Gly, Asp, Glu, Gln, Asn, Lys, Arg, or

11
108

- 98 -

Cys;

Xaa at position 23 is: His, Asp, Lys, Glu, Gln, or Arg;

Xaa at position 24 is: Glu, Arg, Ala, or Lys;

Xaa at position 26 is: Trp, Tyr, Phe, Asp, Lys, Glu, or His;

5 Xaa at position 27 is: Ala, Glu, His, Phe, Tyr, Trp, Arg, or Lys;

Xaa at position 30 is: Ala, Glu, Asp, Ser, or His;

Xaa at position 31 is: Asp, Glu, Ser, Thr, Arg, Trp, or Lys;

Xaa at position 33 is: Asp, Arg, Val, Lys, Ala, Gly, or Glu;

10 Xaa at position 34 is: Glu, Lys, or Asp;

Xaa at position 35 is: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, or Glu;

Xaa at position 36 is: Thr, Ser, Asp, Trp, Tyr, Phe, Arg, Glu, or His;

15 Xaa at position 37 is: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, or is deleted.

6. The heterologous fusion protein of Claims 1, 2, or 3 wherein the GLP-1 compound comprises the sequence of formula

20 III (SEQ ID NO: 4):

	7	8	9	10	11	12	13	14	15	16	17
	Xaa	Xaa	Glu	Gly	Xaa	Xaa	Thr	Ser	Asp	Xaa	Ser
	18	19	20	21	22	23	24	25	26	27	28
	Ser	Tyr	Leu	Glu	Xaa	Xaa	Xaa	Ala	Xaa	Xaa	Phe
25	29	30	31	32	33	34	35	36	37		
	Ile	Ala	Xaa	Leu	Xaa	Xaa	Xaa	Xaa	Xaa		
	formula III (SEQ ID NO: 4)										

wherein:

30 Xaa at position 7 is: L-histidine, D-histidine, desamino-histidine, 2-amino-histidine, β -hydroxy-histidine, homohistidine, α -fluoromethyl-histidine or α -methyl-histidine;

Xaa at position 8 is: Gly, Ala, Val, Leu, Ile, Ser, or Thr;

35 Xaa at position 11 is: Asp, Glu, Arg, Thr, Ala, Lys, or His;

- 99 -

Xaa at position 12 is: His, Trp, Phe, or Tyr;

Xaa at position 16 is: Leu, Ser, Thr, Trp, His, Phe, Asp,
Val, Glu, or Ala;

Xaa at position 22: Gly, Asp, Glu, Gln, Asn, Lys, Arg, or
Cys;

Xaa at position 23 is: His, Asp, Lys, Glu, or Gln;

Xaa at position 24 is: Glu, His, Ala, or Lys;

Xaa at position 25 is: Asp, Lys, Glu, or His;

Xaa at position 27 is: Ala, Glu, His, Phe, Tyr, Trp, Arg,
or Lys;

Xaa at position 30 is: Ala, Glu, Asp, Ser, or His;

Xaa at position 33 is: Asp, Arg, Val, Lys, Ala, Gly, or
Glu;

Xaa at position 34 is: Glu, Lys, or Asp;

Xaa at position 35 is: Thr, Ser, Lys, Arg, Trp, Tyr, Phe,
Asp, Gly, Pro, His, or Glu;

Xaa at position 36 is: Arg, Glu, or His;

Xaa at position 37 is: Lys, Arg, Thr, Ser, Glu, Asp, Trp,
Tyr, Phe, His, Gly, Gly-Pro, or is deleted.

7. The heterologous fusion protein of Claim 1, 2, or 3
wherein the GLP-1 compound comprises the sequence of formula
IV (SEQ ID NO: 5):

7	8	9	10	11	12	13	14	15	16	17
Xaa-Xaa-Glu-Gly-Thr-Xaa-Thr-Ser-Asp-Xaa-Ser-										
18	19	20	21	22	23	24	25	26	27	28
Ser-Tyr-Leu-Glu-Xaa-Xaa-Ala-Ala-Xaa-Glu-Phe-										
29	30	31	32	33	34	35	36	37		
Ile-Xaa-Trp-Leu-Val-Lys-Xaa-Arg-Xaa										

formula IV (SEQ ID NO: 5)

wherein:

Xaa at position 7 is: L-histidine, D-histidine, desamino-
histidine, 2-amino-histidine, β -hydroxy-histidine,

117
110

- 100 -

homohistidine, α -fluoromethyl-histidine or α -methyl-histidine;

Xaa at position 8 is: Gly, Ala, Val, Leu, Ile, Ser, or Thr;

Xaa at position 12 is: His, Trp, Phe, or Tyr;

5 Xaa at position 16 is: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu, or Ala;

Xaa at position 22 is: Gly, Asp, Glu, Gln, Asn, Lys, Arg, or Cys;

Xaa at position 23 is: His, Asp, Lys, Glu, or Gln;

10 Xaa at position 26 is: Asp, Lys, Glu, or His;

Xaa at position 30 is: Ala, Glu, Asp, Ser, or His;

Xaa at position 35 is: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, or Glu;

15 Xaa at position 37 is: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, $-\text{NH}_2$, Gly, Gly-Pro, or Gly-Pro- NH_2 , or is deleted.

8. The heterologous fusion protein of Claims 1, 2, or 3 wherein the GLP-1 compound comprises the sequence of formula

20 V (SEQ ID NO: 6)

7	8	9	10	11	12	13	14	15	16	17
Xaa-Xaa-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Val-Ser-										
18	19	20	21	22	23	24	25	26	27	28
Ser-Tyr-Leu-Glu-Xaa-Xaa-Ala-Ala-Lys-Xaa-Phe-										
29	30	31	32	33	34	35	36	37		
Ile-Xaa-Trp-Leu-Val-Lys-Gly-Arg-Xaa										
formula V (SEQ ID NO: 6)										

30 wherein:

Xaa at position 7 is: L-histidine, D-histidine, desamino-histidine, 2-amino-histidine, β -hydroxy-histidine, homohistidine, α -fluoromethyl-histidine or α -methyl-histidine;

35 Xaa at position 8 is: Gly, Ala, Val, Leu, Ile, Ser, or Thr;

- 101 -

Xaa at position 22 is: Gly, Asp, Glu, Gln, Asn, Lys, Arg, or Cys;

Xaa at position 23 is: His, Asp, Lys, Glu, or Gln;

Xaa at position 24 is: Ala, Glu, His, Phe, Tyr, Trp, Arg, or Lys

Xaa at position 30 is: Ala, Glu, Asp, Ser, or His;

Xaa at position 37 is: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, or is deleted.

9. The heterologous fusion protein of Claims 1, 2, or 3 wherein the GLP-1 compound comprises the sequence of formula VIII (SEQ ID NO:11).

	7	8	9	10	11	12	13	14	15	16	17
15	Xaa-Xaa-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Xaa-Ser-										
	18	19	20	21	22	23	24	25	26	27	28
	Xaa-Xaa-Xaa-Glu-Xaa-Xaa-Ala-Xaa-Xaa-Xaa-Phe-										
	29	30	31	32	33	34	35	36	37		
	Ile-Xaa-Trp-Leu-Xaa-Xaa-Gly-Xaa-Xaa										
20	formula VIII (SEQ ID NO: 11)										

wherein:

Xaa at position 7 is: L-histidine, D-histidine, desamino-histidine, 2-amino-histidine, β -hydroxy-histidine, homohistidine, α -fluoromethyl-histidine or α -methyl-histidine;

Xaa at position 8 is: Gly, Ala, or Val;

Xaa at position 16 is: Leu or Val;

Xaa at position 18 is Lys or Ser;

Xaa at position 19 is: Gln or Tyr;

Xaa at position 20 is: Met or Leu;

Xaa at position 22 is: Glu or Gln;

Xaa at position 23 is: Glu, or Gln;

Xaa at position 25 is: Val or Ala;

Xaa at position 26 is: Arg or Lys;

115.
112

- 102 -

Xaa at position 27 is Leu or Glu;

Xaa at position 30 is: Glu or Ala;

Xaa at position 33 is: Val or Lys;

Xaa at position 34 is: Asn or Lys;

5 Xaa at position 36 is: Gly or Arg; and

Xaa at position 37 is: Gly, Pro, Pro-Ser-Ser-Gly-Ala-Pro-Pro-Pro-Ser, or is absent.

10. The heterologous fusion protein of claims 1 through 9
10 wherein the GLP-1 compound has no more than 6 amino acids
that are different from the corresponding amino acid in GLP-
1(7-37)OH, GLP-1(7-36)OH, or Exendin-4.

11. The heterologous fusion protein of Claim 10 wherein the
15 GLP-1 compound has no more than 5 amino acids that differ
from the corresponding amino acid in GLP-1(7-37)OH, GLP-1(7-
36)OH, or Exendin-4.

12. The heterologous fusion protein of Claim 11 wherein the
20 GLP-1 compound has no more than 4 amino acids that differ
from the corresponding amino acid in GLP-1(7-37)OH, GLP-1(7-
36)OH, or Exendin-4.

13. The heterologous fusion protein of Claim 12 wherein the
25 GLP-1 compound has no more than 3 amino acids that differ
from the corresponding amino acid in GLP-1(7-37)OH, GLP-1(7-
36)OH, or Exendin-4.

14. The heterologous fusion protein of Claim 13 wherein the
30 GLP-1 compound has no more than 2 amino acids that differ
from the corresponding amino acid in GLP-1(7-37)OH, GLP-1(7-
36)OH, or Exendin-4.

15. The heterologous fusion protein of any one of Claims 1
35 through 14 wherein Xaa at position 8 is glycine or valine.

113

- 103 -

16. The heterologous fusion protein of Claim 8 wherein the GLP-1 compound has no more than 2 amino acids that differ from the corresponding amino acid in GLP-1(7-37)OH or GLP-1(7-36)OH and Xaa at position 8 is glycine or valine and Xaa at position 30 is alanine, glutamic acid, aspartic acid, serine, or histidine.
17. The heterologous fusion protein of Claim 16 wherein Xaa at position 30 is Glutamic acid.
18. The heterologous fusion protein of Claim 8 wherein the GLP-1 compound has no more than 2 amino acids that differ from the corresponding amino acid in GLP-1(7-37)OH or GLP-1(7-36)OH and Xaa at position 8 is glycine or valine and Xaa at position 37 is histidine, phenylalanine, tyrosine, tryptophan, aspartic acid, glutamic acid, serine, threonine, arginine, or lysine.
19. The heterologous fusion protein of Claim 18 wherein Xaa at position 37 is histidine.
20. The heterologous fusion protein of Claim 8 wherein the GLP-1 compound has no more than 2 amino acids that differ from the corresponding amino acid in GLP-1(7-37)OH or GLP-1(7-36)OH and Xaa at position 8 is glycine, valine, leucine, isoleucine, serine, threonine, or methionine and Xaa at position 22 is aspartic acid, glutamic acid, lysine, arginine, asparagine, glutamine or histidine.
21. The heterologous fusion protein of Claim 20 wherein Xaa at position 22 is lysine or glutamic acid.
22. The heterologous fusion protein of Claim 8 wherein the GLP-1 compound is Val⁸-GLP-1(7-37).

117
114

- 104 -

23. The heterologous fusion protein of Claim 8 wherein the GLP-1 compound is Gly⁸-GLP-1(7-37).

5 24. The heterologous fusion protein of any one of Claims 1 through 23 wherein the second polypeptide is human albumin.

25. The heterologous fusion protein of Claim 24 wherein the second polypeptide has the sequence of SEQ ID NO: 34.

10

26. The heterologous fusion protein of any one of Claims 1 through 23 wherein the second polypeptide is an N-terminal fragment of albumin.

15 27. A heterologous fusion protein comprising a first polypeptide with a N-terminus and a C-terminus fused to a second polypeptide with a N-terminus and a C-terminus wherein the first polypeptide is a GLP-1 compound and the second polypeptide is selected from the group consisting of
20 a) the Fc portion of an immunoglobulin;
b) an analog of the Fc portion of an immunoglobulin;
and
c) fragments of the Fc portion of an immunoglobulin,
and wherein the C-terminus of the first polypeptide is fused
25 to the N-terminus of the second polypeptide.

28. A heterologous fusion protein comprising a first polypeptide with a N-terminus and a C-terminus fused to a second polypeptide with a N-terminus and a C-terminus
30 wherein the first polypeptide is a GLP-1 compound and the second polypeptide is selected from the group consisting of
a) the Fc portion of an immunoglobulin;
b) an analog of the Fc portion of an immunoglobulin;
and
35

- 105 -

c) fragments of the Fc portion of an immunoglobulin,
and wherein the C-terminus of the first polypeptide is
fused to the N-terminus of the second polypeptide via a
peptide linker.

29. The heterologous fusion protein of the Claim 28 wherein
the peptide linker is selected from the group consisting of:

- a) a glycine rich peptide;
- 10 b) a peptide having the sequence [Gly-Gly-Gly-Gly-Ser]_n
where n is 1, 2, 3, 4, 5, or 6; and
- c) a peptide having the sequence [Gly-Gly-Gly-Gly-Ser]₃.

30. The heterologous fusion protein of any one of Claims
15 27, 28, or 29 wherein the GLP-1 compound comprises the
sequence of formula 1 [SEQ ID NO: 2].

7 8 9 10 11 12 13 14 15 16 17
His-Xaa-Xaa-Gly-Xaa-Phe-Thr-Xaa-Asp-Xaa-Xaa-
20 18 19 20 21 22 23 24 25 26 27 28
Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Phe-
29 30 31 32 33 34 35 36 37 38 39
Ile-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-
40 41 42 43 44 45
25 Xaa-Xaa-Xaa-Xaa-Xaa-Xaa

Formula I (SEQ ID NO: 2)

wherein:

- Xaa at position 8 is Ala, Gly, Ser, Thr, Leu, Ile, Val,
30 Glu, Asp, or Lys;
- Xaa at position 9 is Glu, Asp, or Lys;
- Xaa at position 11 is Thr, Ala, Gly, Ser, Leu, Ile, Val,
Glu, Asp, or Lys;
- Xaa at position 14 is Ser, Ala, Gly, Thr, Leu, Ile, Val,
35 Glu, Asp, or Lys;

- 106 -

- Xaa at position 16 is Val, Ala, Gly, Ser, Thr, Leu, Ile, Tyr, Glu, Asp, Trp, or Lys;
- Xaa at position 17 is Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, or Lys;
- 5 Xaa at position 18 is Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, Trp, Tyr, or Lys;
- Xaa at position 19 is Tyr, Phe, Trp, Glu, Asp, Gln, or Lys;
- Xaa at position 20 is Leu, Ala, Gly, Ser, Thr, Ile, Val, Glu, Asp, Met, Trp, Tyr, or Lys;
- 10 Xaa at position 21 is Glu, Asp, or Lys;
- Xaa at position 22 is Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, or Lys;
- Xaa at position 23 is Gln, Asn, Arg, Glu, Asp, or Lys;
- Xaa at position 24 is Ala, Gly, Ser, Thr, Leu, Ile, Val, Arg, Glu, Asp, or Lys;
- 15 Xaa at position 25 is Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp, or Lys;
- Xaa at position 26 is Lys, Arg, Gln, Glu, Asp, or His;
- Xaa at position 27 is Leu, Glu, Asp, or Lys;
- 20 Xaa at position 30 is Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp, or Lys;
- Xaa at position 31 is Trp, Phe, Tyr, Glu, Asp, or Lys;
- Xaa at position 32 is Leu, Gly, Ala, Ser, Thr, Ile, Val, Glu, Asp, or Lys;
- 25 Xaa at position 33 is Val, Gly, Ala, Ser, Thr, Leu, Ile, Glu, Asp, or Lys;
- Xaa at position 34 is Asn, Lys, Arg, Glu, Asp, or His;
- Xaa at position 35 is Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, or Lys;
- 30 Xaa at position 36 is Gly, Arg, Lys, Glu, Asp, or His;
- Xaa at position 37 is Pro, Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, or Lys, or is deleted;
- Xaa at position 38 is Ser, Arg, Lys, Glu, Asp, or His, or is deleted;
- 35 Xaa at position 39 is Ser, Arg, Lys, Glu, Asp, or His, or

- 107 -

is deleted;

Xaa at position 40 is Gly, Asp, Glu, or Lys, or is deleted;

Xaa at position 41 is Ala, Phe, Trp, Tyr, Glu, Asp, or Lys,
or is deleted;

5 Xaa at position 42 is Ser, Pro, Lys, Glu, or Asp, or is
deleted;

Xaa at position 43 is Ser, Pro, Glu, Asp, or Lys, or is
deleted;

10 Xaa at position 44 is Gly, Pro, Glu, Asp, or Lys, or is
deleted;

and

Xaa at position 45 is Ala, Ser, Val, Glu, Asp, or Lys, or is
deleted;

15 provided that when the amino acid at position 37, 38, 39,
40, 41, 42, 43, or 44 is deleted, then each amino acid
downstream of that amino acid is also deleted.

31. The heterologous fusion protein of Claims 27, 28, or 29
wherein the GLP-1 compound comprises the sequence of formula
20 II (SEQ ID NO: 3):

7 8 9 10 11 12 13 14 15 16 17
Xaa-Xaa-Xaa-Gly-Xaa-Xaa-Thr-Ser-Asp-Xaa-Ser-
18 19 20 21 22 23 24 25 26 27 28
25 Xaa-Tyr-Leu-Glu-Xaa-Xaa-Xaa-Ala-Xaa-Xaa-Phe-
29 30 31 32 33 34 35 36 37
Ile-Xaa-Xaa-Leu-Xaa-Xaa-Xaa-Xaa-Xaa

Formula II (SEQ ID NO: 3)

30 wherein:

Xaa at position 7 is: L-histidine, D-histidine, desamino-
histidine, 2-amino-histidine, β -hydroxy-histidine,
homohistidine, α -fluoromethyl-histidine or α -methyl-
histidine;

35 Xaa at position 8 is: Gly, Ala, Val, Leu, Ile, Ser, or Thr;

Xaa at position 9 is: Thr, Ser, Arg, Lys, Trp, Phe, Tyr,
Glu, or His;

Xaa at position 11 is: Asp, Glu, Arg, Thr, Ala, Lys, or His;

Xaa at position 12 is: His, Trp, Phe, or Tyr;

5 Xaa at position 16 is: Leu, Ser, Thr, Trp, His, Phe, Asp,
Val, Tyr, Glu, or Ala;

Xaa at position 18 is: His, Pro, Asp, Glu, Arg, Ser, Ala, or Lys;

10 Xaa at position 19 is: Gly, Asp, Glu, Gln, Asn, Lys, Arg, or Cys;

Xaa at position 23 is: His, Asp, Lys, Glu, Gln, or Arg;

Xaa at position 24 is: Glu, Arg, Ala, or Lys;

Xaa at position 26 is: Trp, Tyr, Phe, Asp, Lys, Glu, or His;

Xaa at position 27 is: Ala, Glu, His, Phe, Tyr, Trp, Arg, or
15 Lys;

Xaa at position 30 is: Ala, Glu, Asp, Ser, or His;

Xaa at position 31 is: Asp, Glu, Ser, Thr, Arg, Trp, or Lys;

Xaa at position 33 is: Asp, Arg, Val, Lys, Ala, Gly, or Glu;

Xaa at position 34 is: Glu, Lys, or Asp;

20 Xaa at position 35 is: Thr, Ser, Lys, Arg, Trp, Tyr, Phe,
Asp, Gly, Pro, His, or Glu;

Xaa at position 36 is: Thr, Ser, Asp, Trp, Tyr, Phe, Arg,
Glu, or His;

25 Xaa at position 37 is: Lys, Arg, Thr, Ser, Glu, Asp, Trp,
Tyr, Phe, His, Gly, Gly-Pro, or is deleted.

32. The heterologous fusion protein of Claims 27, 28, or 29 wherein the GLP-1 compound comprises the sequence of formula III (SEQ ID NO: 4):

```

30      7   8   9   10  11  12  13  14  15  16  17
      Xaa-Xaa-Glu-Gly-Xaa-Xaa-Thr-Ser-Asp-Xaa-Ser-
      18  19  20  21  22  23  24  25  26  27  28
      Ser-Tyr-Leu-Glu-Xaa-Xaa-Xaa-Ala-Xaa-Xaa-Phe-
      29  30  31  32  33  34  35  36  37
35      Ile-Ala-Xaa-Leu-Xaa-Xaa-Xaa-Xaa-R

```

- 109 -

formula III (SEQ ID NO: 4)

wherein:

- 5 Xaa at position 7 is: L-histidine, D-histidine, desamino-histidine, 2-amino-histidine, β -hydroxy-histidine, homohistidine, α -fluoromethyl-histidine or α -methyl-histidine;
- Xaa at position 8 is: Gly, Ala, Val, Leu, Ile, Ser, or Thr;
- 10 Xaa at position 11 is: Asp, Glu, Arg, Thr, Ala, Lys, or His;
- Xaa at position 12 is: His, Trp, Phe, or Tyr;
- Xaa at position 16 is: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu, or Ala;
- Xaa at position 22: Gly, Asp, Glu, Gln, Asn, Lys, Arg, or Cys;
- 15 Xaa at position 23 is: His, Asp, Lys, Glu, or Gln;
- Xaa at position 24 is: Glu, His, Ala, or Lys;
- Xaa at position 25 is: Asp, Lys, Glu, or His;
- Xaa at position 27 is: Ala, Glu, His, Phe, Tyr, Trp, Arg, or Lys;
- 20 Xaa at position 30 is: Ala, Glu, Asp, Ser, or His;
- Xaa at position 33 is: Asp, Arg, Val, Lys, Ala, Gly, or Glu;
- Xaa at position 34 is: Glu, Lys, or Asp;
- Xaa at position 35 is: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, or Glu;
- 25 Xaa at position 36 is: Arg, Glu, or His;
- Xaa at position 37 is: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, or is deleted.
- 30 33. The heterologous fusion protein of Claim 27, 28, or 29 wherein the GLP-1 compound comprises the sequence of formula IV (SEQ ID NO: 5):
- | | | | | | | | | | | | |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| | Xaa | Xaa | Glu | Gly | Thr | Xaa | Thr | Ser | Asp | Xaa | Ser |
| 35 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 |

- 110 -

Ser-Tyr-Leu-Glu-Xaa-Xaa-Ala-Ala-Xaa-Glu-Phe-
29 30 31 32 33 34 35 36 37
Ile-Xaa-Trp-Leu-Val-Lys-Xaa-Arg-Xaa
formula IV (SEQ ID NO: 5)

5

wherein:

Xaa at position 7 is: L-histidine, D-histidine, desamino-
histidine, 2-amino-histidine, β -hydroxy-histidine,
homohistidine, α -fluoromethyl-histidine or α -methyl-
histidine;

10

Xaa at position 8 is: Gly, Ala, Val, Leu, Ile, Ser, or Thr;

Xaa at position 12 is: His, Trp, Phe, or Tyr;

Xaa at position 16 is: Leu, Ser, Thr, Trp, His, Phe, Asp,
Val, Glu, or Ala;

15

Xaa at position 22 is: Gly, Asp, Glu, Gln, Asn, Lys, Arg, or
Cys;

Xaa at position 23 is: His, Asp, Lys, Glu, or Gln;

Xaa at position 26 is: Asp, Lys, Glu, or His;

Xaa at position 30 is: Ala, Glu, Asp, Ser, or His;

20

Xaa at position 35 is: Thr, Ser, Lys, Arg, Trp, Tyr, Phe,
Asp, Gly, Pro, His, or Glu;

Xaa at position 37 is: Lys, Arg, Thr, Ser, Glu, Asp, Trp,
Tyr, Phe, His, Gly, Gly-Pro, or is deleted.

25

34. The heterologous fusion protein of Claims 27, 28, or 29
wherein the GLP-1 compound comprises the sequence of formula
V (SEQ ID NO: 6)

30

7 8 9 10 11 12 13 14 15 16 17
Xaa-Xaa-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Val-Ser-
18 19 20 21 22 23 24 25 26 27 28
Ser-Tyr-Leu-Glu-Xaa-Xaa-Ala-Ala-Lys-Xaa-Phe-
29 30 31 32 33 34 35 36 37
Ile-Xaa-Trp-Leu-Val-Lys-Gly-Arg-Xaa
formula V (SEQ ID NO: 6)

35

- 111 -

wherein:

Xaa at position 7 is: L-histidine, D-histidine, desamino-histidine, 2-amino-histidine, β -hydroxy-histidine,
5 homohistidine, α -fluoromethyl-histidine or α -methyl-histidine;

Xaa at position 8 is: Gly, Ala, Val, Leu, Ile, Ser, or Thr;

Xaa at position 22 is: Gly, Asp, Glu, Gln, Asn, Lys, Arg, r
Cys;

10 Xaa at position 23 is: His, Asp, Lys, Glu, or Gln;

Xaa at position 24 is: Ala, Glu, His, Phe, Tyr, Trp, Arg,
or Lys

Xaa at position 30 is: Ala, Glu, Asp, Ser, or His;

Xaa at position 37 is: Lys, Arg, Thr, Ser, Glu, Asp, Trp,
15 Tyr, Phe, His, Gly, Gly-Pro, or is deleted.

35. The heterologous fusion protein of Claims 27, 28, or 29
wherein the GLP-1 compound comprises the sequence of formula
VIII (SEQ ID NO:11).

20

7	8	9	10	11	12	13	14	15	16	17
Xaa-Xaa-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Xaa-Ser-										
18	19	20	21	22	23	24	25	26	27	28
Xaa-Xaa-Xaa-Glu-Xaa-Xaa-Ala-Xaa-Xaa-Xaa-Phe-										
29	30	31	32	33	34	35	36	37		
Ile-Xaa-Trp-Leu-Xaa-Xaa-Gly-Xaa-R										
formula VIII (SEQ ID NO: 11)										

25

wherein:

30 Xaa at position 7 is: L-histidine, D-histidine, desamino-histidine, 2-amino-histidine, β -hydroxy-histidine, homohistidine, α -fluoromethyl-histidine or α -methyl-histidine;

Xaa at position 8 is: Gly, Ala, or Val;

35 Xaa at position 16 is: Leu or Val;

j22

- 112 -

Xaa at position 18 is Lys or Ser;

Xaa at position 19 is: Gln or Tyr;

Xaa at position 20 is: Met or Leu;

Xaa at position 22 is: Glu or Gln;

5 Xaa at position 23 is: Glu, or Gln;

Xaa at position 25 is: Val or Ala;

Xaa at position 26 is: Arg or Lys;

Xaa at position 27 is Leu or Glu;

Xaa at position 30 is: Glu or Ala;

10 Xaa at position 33 is: Val or Lys;

Xaa at position 34 is: Asn or Lys;

Xaa at position 36 is: Gly or Arg; and

Xaa at position 37 is: Gly, Pro, Pro-Ser-Ser-Gly-Ala-Pro-Pro-Ser, or is absent.

15

36. The heterologous fusion protein of claims 27 through 35 wherein the GLP-1 compound has no more than 6 amino acids that are different from the corresponding amino acid in GLP-1(7-37)OH, GLP-1(7-36)OH, or Exendin-4.

20

37. The heterologous fusion protein of Claim 36 wherein the GLP-1 compound has no more than 5 amino acids that differ from the corresponding amino acid in GLP-1(7-37)OH, GLP-1(7-36)OH, or Exendin-4.

25

38. The heterologous fusion protein of Claim 37 wherein the GLP-1 compound has no more than 4 amino acids that differ from the corresponding amino acid in GLP-1(7-37)OH, GLP-1(7-36)OH, or Exendin-4.

30

39. The heterologous fusion protein of Claim 38 wherein the GLP-1 compound has no more than 3 amino acids that differ from the corresponding amino acid in GLP-1(7-37)OH, GLP-1(7-36)OH, or Exendin-4.

35

- 113 -

40. The heterologous fusion protein of Claim 39 wherein the GLP-1 compound has no more than 2 amino acids that differ from the corresponding amino acid in GLP-1(7-37)OH, GLP-1(7-36)OH, or Exendin-4.

5

41. The heterologous fusion protein of any one of Claims 27 through 40 wherein Xaa at position 8 is glycine or valine.

42. The heterologous fusion protein of Claim 34 wherein the GLP-1 compound has no more than 2 amino acids that differ from the corresponding amino acid in GLP-1(7-37)OH or GLP-1(7-36)OH and Xaa at position 8 is glycine or valine and Xaa at position 30 is alanine, glutamic acid, aspartic acid, serine, or histidine.

15

43. The heterologous fusion protein of Claim 42 wherein Xaa at position 30 is Glutamic acid.

44. The heterologous fusion protein of Claim 34 wherein the GLP-1 compound has no more than 2 amino acids that differ from the corresponding amino acid in GLP-1(7-37)OH or GLP-1(7-36)OH and Xaa at position 8 is glycine or valine and Xaa at position 37 is histidine, phenylalanine, tyrosine, tryptophan, aspartic acid, glutamic acid, serine, threonine, arginine, or lysine.

25

45. The heterologous fusion protein of Claim 44 wherein Xaa at position 37 is histidine.

124

- 114 -

46. The heterologous fusion protein of Claim 34 wherein the GLP-1 compound has no more than 2 amino acids that differ from the corresponding amino acid in GLP-1(7-37)OH or GLP-1(7-36)OH and Xaa at position 8 is glycine, valine, leucine, isoleucine, serine, threonine, or methionine and Xaa at position 22 is aspartic acid, glutamic acid, lysine, arginine, asparagine, glutamine or histidine.
47. The heterologous fusion protein of Claim 46 wherein Xaa at position 22 is lysine or glutamic acid.
48. The heterologous fusion protein of Claim 34 wherein the GLP-1 compound is Val-8-GLP-1(7-37).
49. The heterologous fusion protein of Claim 34 wherein the GLP-1 compound is Gly-8-GLP-1(7-37).
50. The heterologous fusion protein of any one of Claims 27 through 49 wherein the second polypeptide is the Fc portion of an Ig selected from the group consisting of: IgG1, IgG2, IgG3, IgG4, IgE, IgA, IgD, or IgM.
51. The heterologous fusion protein of any one of Claims 27 through 50 wherein the second polypeptide is the Fc portion of an Ig selected from the group consisting of: IgG1, IgG2, IgG3, and IgG4.
52. The heterologous fusion protein of Claim 51 wherein the second polypeptide is the Fc portion of an IgG1 immunoglobulin.
53. The heterologous fusion protein of Claim 51 wherein the second polypeptide is the Fc portion of an IgG4 immunoglobulin.

- 115 -

54. The heterologous fusion protein of Claims 50 through 53 wherein the IgG is human.

5 55. The heterologous fusion protein of any one of Claims 27 through 54 wherein the Fc portion comprises the hinge, CH2, and CH3 domains.

10 56. The heterologous fusion protein of Claim 52 wherein the second polypeptide has the sequence of SEQ ID NO: 32.

57. A polynucleotide encoding a heterologous fusion protein of any one of Claims 1 through 56.

15 58. A vector comprising the polynucleotide of Claim 57.

59. A host cell comprising the vector of Claim 58.

20 60. A host cell expressing at least one heterologous fusion protein of any one of Claims 1 through 56.

61. The host cell of Claim 60 wherein said host cell is a CHO cell.

25 62. A process for producing a heterologous fusion protein comprising the steps of transcribing and translating a polynucleotide of Claim 57 under conditions wherein the heterologous fusion protein is expressed in detectable amounts.

30 63. A method for normalizing blood glucose levels in a mammal in need thereof comprising the administration of a therapeutically effective amount of the heterologous fusion protein of any one of Claims 1 through 56.

35

- 116 -

64. A method of treating a patient with non-insulin dependent diabetes mellitus comprising the administration of a therapeutically effective amount of the heterologous fusion protein of any one of Claims 1 through 56.

5

65. A method of treating obesity comprising the administration of a therapeutically effective amount of the heterologous fusion protein of any one of Claims 1 through 56.

10

66. The use of a heterologous fusion protein as claimed in any one of Claims 1 through 56 for the manufacture of a medicament for the treatment of patients with non-insulin dependent diabetes mellitus.

15

67. The use of a heterologous fusion protein as claimed in any one of Claims 1 through 56 for the manufacture of a medicament for the treatment of patients with obesity.

20

68. A pharmaceutical formulation adapted for the treatment of patients with non-insulin dependent diabetes comprising a heterologous fusion protein of any one of Claims 1 through 56.

25

69. The heterologous fusion protein of Claims 1, 2, or 3 wherein the GLP-1 compound comprises the sequence of formula IX [SEQ ID NO: 12]

30

Xaa₇-Xaa₈-Glu-Gly-Thr-Xaa₁₂-Thr-Ser-Asp-Xaa₁₆-Ser-
Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Gln-Ala-Xaa₂₅-Lys-Xaa₂₇-
Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Lys-Gly-Arg-Xaa₃₇

Formula IX (SEQ ID NO: 12)

wherein:

35

Xaa₇ is: L-histidine, D-histidine, desamin -histidine, 2-

- 117 -

amino-histidine, β -hydroxy-histidine,
homohistidine, α -fluoromethyl-histidine, or α -
methyl-histidine;

- Xaa₉ is: Ala, Gly, Val, Leu, Ile, Ser, or Thr;
- 5 Xaa₁₁ is: Phe, Trp, or Tyr;
- Xaa₁₆ is: Val, Trp, Ile, Leu, Phe, or Tyr;
- Xaa₁₈ is: Ser, Trp, Tyr, Phe, Lys, Ile, Leu, Val;
- Xaa₁₉ is: Tyr, Trp, or Phe;
- Xaa₂₀ is: Leu, Phe, Tyr, or Trp;
- 10 Xaa₂₂ is: Gly, Glu, Asp, or Lys;
- Xaa₂₅ is: Ala, Val, Ile, or Leu;
- Xaa₂₇ is: Glu, Ile, or Ala;
- Xaa₃₀ is: Ala or Glu
- Xaa₃₃ is: Val, or Ile; and
- 15 Xaa₃₇ is: Gly, His, NH₂, or is absent.
70. The heterologous fusion protein of Claim 69 wherein the
second polypeptide is human albumin.
- 20 71. The heterologous fusion protein of Claim 70 wherein the
second polypeptide has the sequence of SEQ ID NO: 34.
72. The heterologous fusion protein of Claim 69 wherein the
second polypeptide is an N-terminal fragment of albumin.
- 25 73. The heterologous fusion protein of any one of Claims
27, 28, or 29 wherein the GLP-1 compound comprises the
sequence of formula IX [SEQ ID NO: 12]
- 30 Xaa₇-Xaa₈-Glu-Gly-Thr-Xaa₁₂-Thr-Ser-Asp-Xaa₁₆-Ser-
Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Gln-Ala-Xaa₂₅-Lys-Xaa₂₇-
Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Lys-Gly-Arg-Xaa₃₇
Formula IX (SEQ ID NO: 12)

- 118 -

wherein:

Xaa₇ is: L-histidine, D-histidine, desamino-histidine, 2-amino-histidine, β -hydroxy-histidine,

5 homohistidine, α -fluoromethyl-histidine, or α -methyl-histidine;

Xaa₈ is: Ala, Gly, Val, Leu, Ile, Ser, or Thr;

Xaa₁₂ is: Phe, Trp, or Tyr;

Xaa₁₆ is: Val, Trp, Ile, Leu, Phe, or Tyr;

10 Xaa₁₈ is: Ser, Trp, Tyr, Phe, Lys, Ile, Leu, Val;

Xaa₁₉ is: Tyr, Trp, or Phe;

Xaa₂₀ is: Leu, Phe, Tyr, or Trp;

Xaa₂₂ is: Gly, Glu, Asp, or Lys;

Xaa₂₅ is: Ala, Val, Ile, or Leu;

15 Xaa₂₇ is: Glu, Ile, or Ala;

Xaa₃₀ is: Ala or Glu

Xaa₃₃ is: Val, or Ile; and

Xaa_n is: Gly, His, NH₂, or is absent.

20 74. The heterologous fusion protein of Claim 73 wherein the second polypeptide is the Fc portion of an Ig selected from the group consisting of: IgG1, IgG2, IgG3, IgG4, IgE, IgA, IgD, or IgM.

25 75. The heterologous fusion protein of Claim 73 or 74 wherein the second polypeptide is the Fc portion of an Ig selected from the group consisting of: IgG1, IgG2, IgG3, and IgG4.

30 76. The heterologous fusion protein of Claim 75 wherein the second polypeptide is the Fc portion of an IgG1 immunoglobulin.

- 119 -

77. The heterologous fusion protein of Claim 75 wherein the second polypeptide is the Fc portion of an IgG4 immunoglobulin.

5 78. The heterologous fusion protein of Claims 73 through 78 wherein the IgG is human.

79. The heterologous fusion protein of Claim 75 wherein the second polypeptide has the sequence of SEQ ID NO:
10 32.

80. The heterologous fusion protein of any one of Claims 69 through 79 wherein the GLP-1 compound is selected from the group consisting of Gly⁸-GLP-1(7-37), Val⁸-Tyr¹²-GLP-1(7-37), Val⁸-Tyr¹²-GLP-1(7-36), Val⁸-Trp¹²-GLP-1(7-37), Val⁸-Leu¹⁶-GLP-1(7-37), Val⁸-Val¹⁶-GLP-1(7-37), Val⁸-Tyr¹⁶-GLP-1(7-37), Gly⁸-Glu²²-GLP-1(7-37), Val⁸-Glu²²-GLP-1(7-37), Val⁸-Leu²⁵-GLP-1(7-37), Val⁸-Tyr¹²-Tyr¹⁶-GLP-1(7-37), Val⁸-Trp¹²-Glu²²-GLP-1(7-37), Val⁸-Tyr¹²-Glu²²-GLP-1(7-37), Val⁸-Tyr¹⁶-Phe¹⁹-GLP-1(7-37), Val⁸-Tyr¹⁶-Glu²²-GLP-1(7-37), Val⁸-Trp¹⁶-Glu²²-GLP-1(7-37), Val⁸-Leu¹⁶-Glu²²-GLP-1(7-37), Val⁸-Ile¹⁶-Glu²²-GLP-1(7-37), Val⁸-Phe¹⁶-Glu²²-GLP-1(7-37), Val⁸-Trp¹⁶-Glu²²-GLP-1(7-37), Val⁸-Tyr¹⁸-Glu²²-GLP-1(7-37), Val⁸-Phe¹⁸-Glu²²-GLP-1(7-37), Val⁸-Ile¹⁸-Glu²²-GLP-1(7-37), Val⁸-Lys¹⁸-Glu²²-GLP-1(7-37), Val⁸-Trp¹⁹-Glu²²-GLP-1(7-37), Val⁸-Phe¹⁹-Glu²²-GLP-1(7-37), Val⁸-Phe²⁰-Glu²²-GLP-1(7-37), Val⁸-Glu²²-Leu²⁵-GLP-1(7-37), Val⁸-Glu²²-Ile²⁵-GLP-1(7-37), Val⁸-Glu²²-Val²⁵-GLP-1(7-37), Val⁸-Glu²²-Ile²⁷-GLP-1(7-37), Val⁸-Glu²²-Ala²⁷-GLP-1(7-37), Val⁸-Glu²²-Ile³³-GLP-1(7-37), Val⁸-Asp⁹-Ile¹¹-Tyr¹⁶-Glu²²-GLP-1(7-37), Val⁸-Tyr¹⁶-Trp¹⁹-Glu²²-GLP-1(7-37), Val⁸-Trp¹⁶-Glu²²-Val²⁵-Ile³³-GLP-1(7-37), Val⁸-Trp¹⁶-Glu²²-Ile³³-GLP-1(7-37), Val⁸-Glu²²-Val²⁵-Ile³³-GLP-1(7-37), Val⁸-Trp¹⁶-

1/17

Fig. 1

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

[SEQ ID NO: 32]

3/17

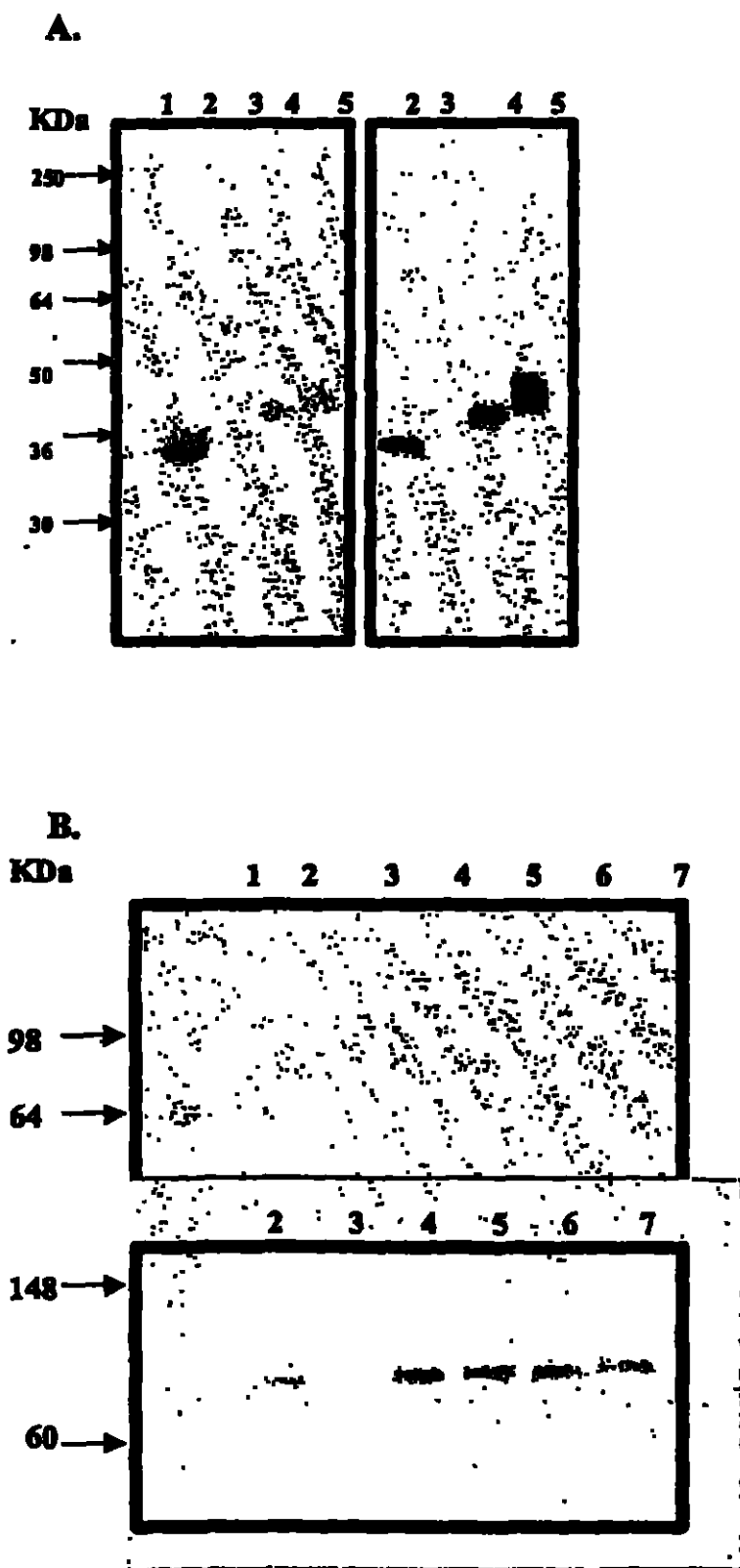
Fig. 2 Continued

			490					495					500				
Cys	Phe	Ser	Ala	Leu	Glu	Val	Asp	Glu	Thr	Tyr	Val	Pro	Lys	Glu	Phe	Asn	Ala
505					510					515					520		
Glu	Thr	Phe	Thr	Phe	His	Ala	Asp	Ile	Cys	Thr	Leu	Ser	Glu	Lys	Glu	Arg	Gln
		525					530					535					540
Ile	Lys	Lys	Gln	Thr	Ala	Leu	Val	Glu	Leu	Val	Lys	His	Lys	Pro	Lys	Ala	Thr
				545					550					555			
Lys	Glu	Gln	Leu	Lys	Ala	Val	Met	Asp	Asp	Phe	Ala	Ala	Phe	Val	Glu	Lys	Cys
	560					565					570					575	
Cys	Lys	Ala	Asp	Asp	Lys	Glu	Thr	Cys	Phe	Ala	Glu	Glu	Gly	Lys	Lys	Leu	Val
		580						585									
Ala	Ala	Ser	Gln	Ala	Ala	Leu	Gly	Leu									

[SEQ ID NO: 34]

4/17

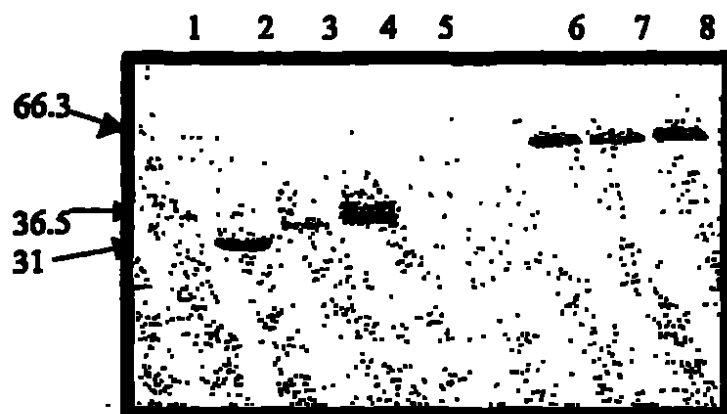
Fig. 3



134

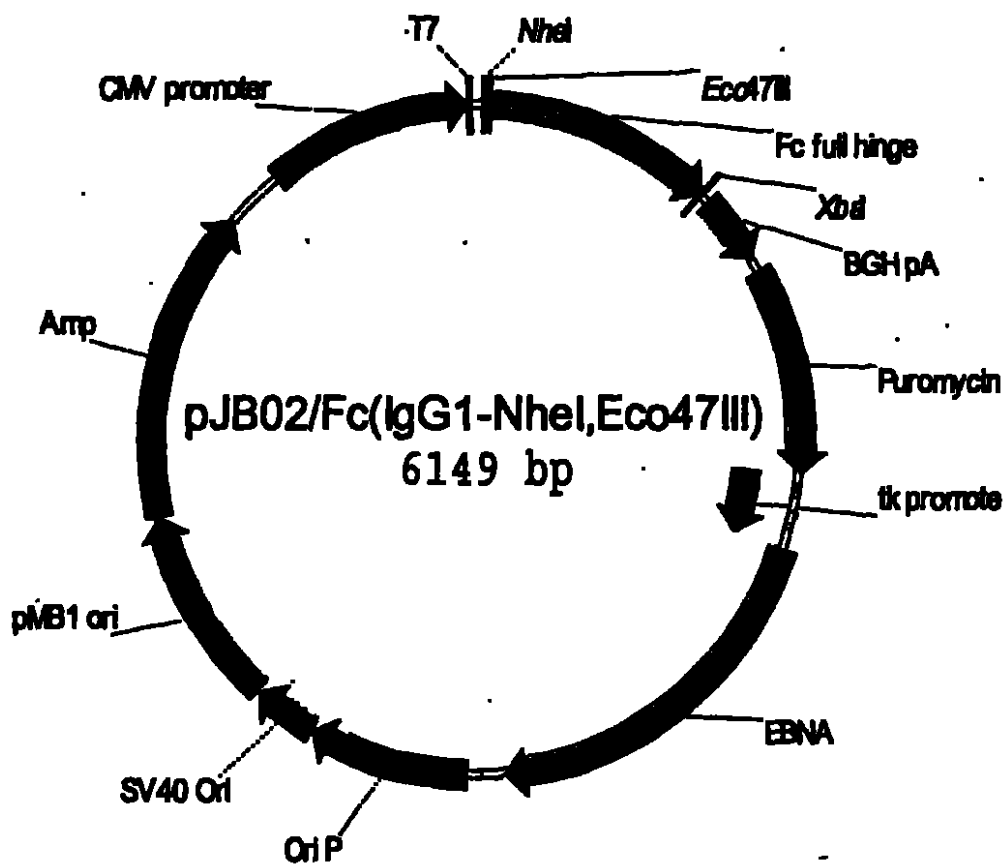
5/17

Fig. 4



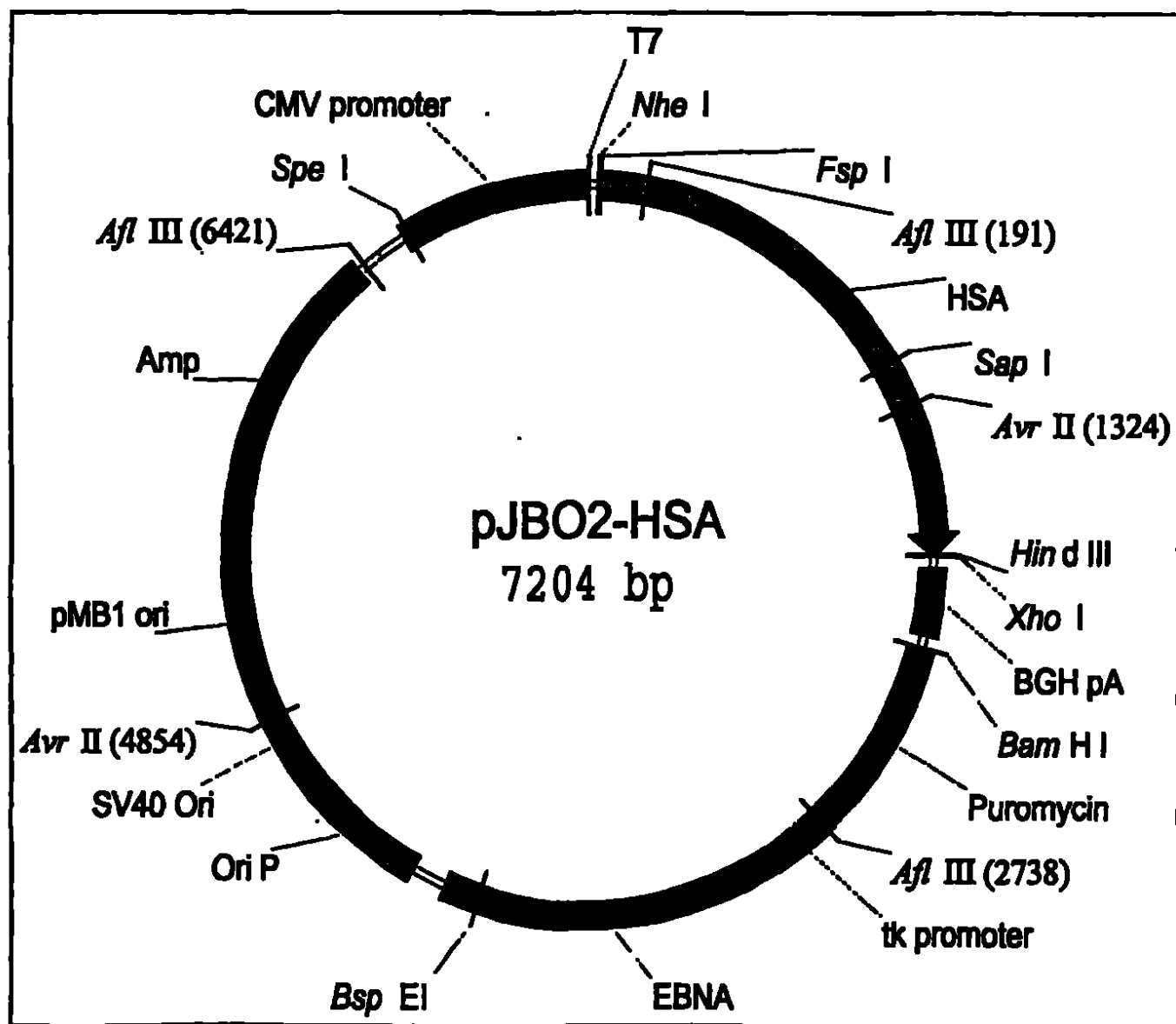
6/17

Fig. 5



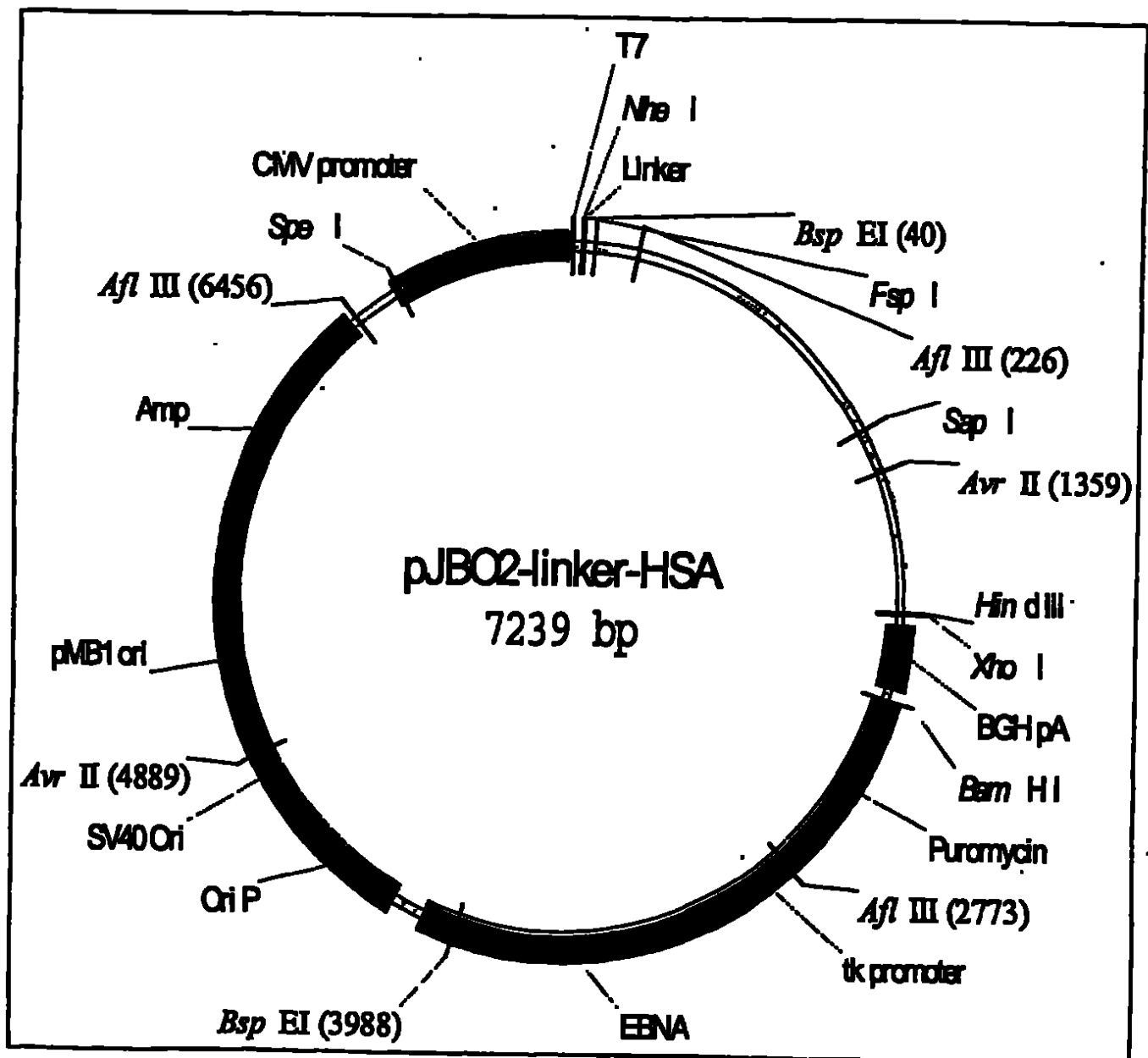
7/17

Fig. 6



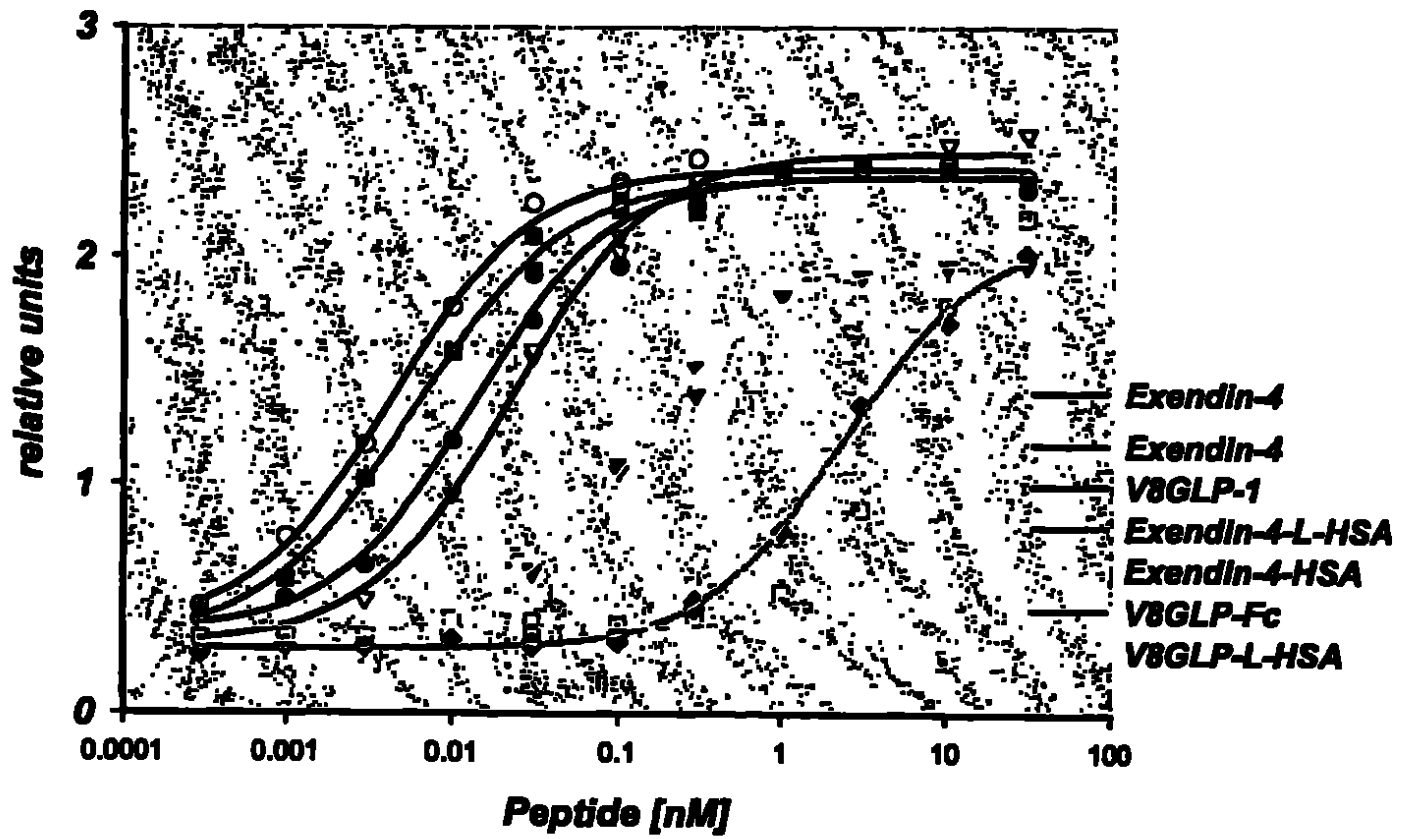
8/17

Fig. 7



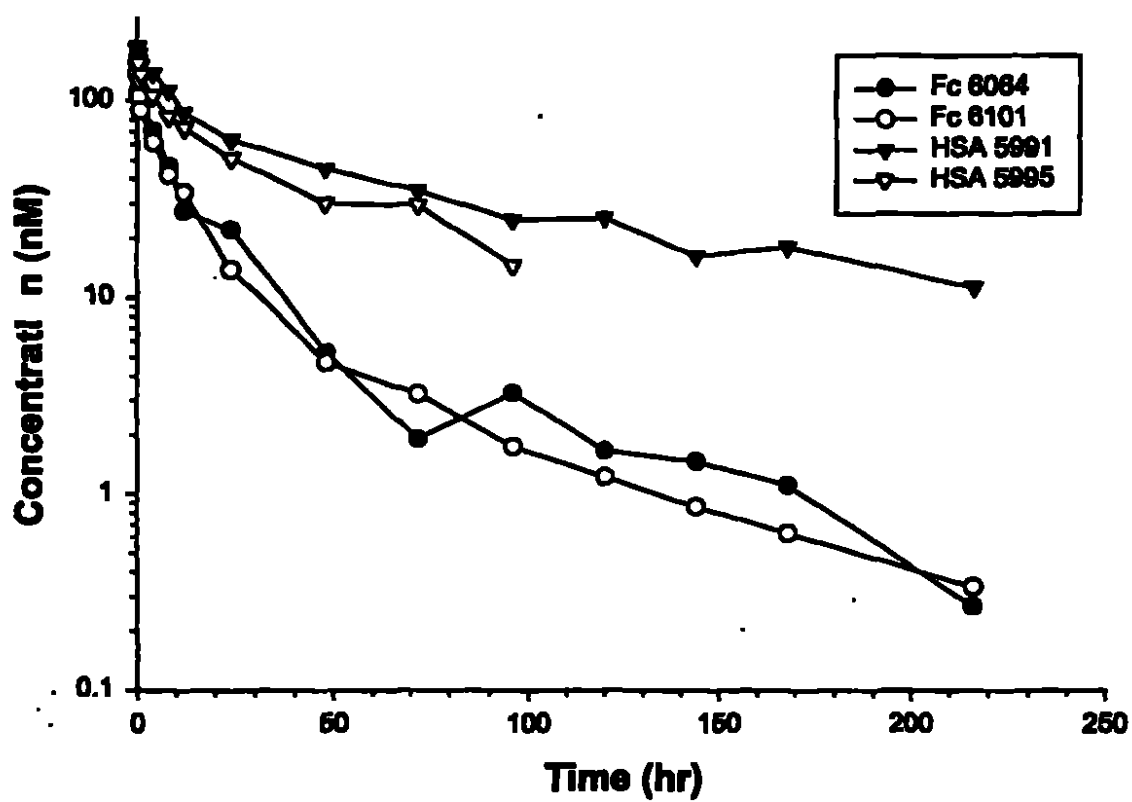
9/17

Fig. 8



10/17

Fig. 9

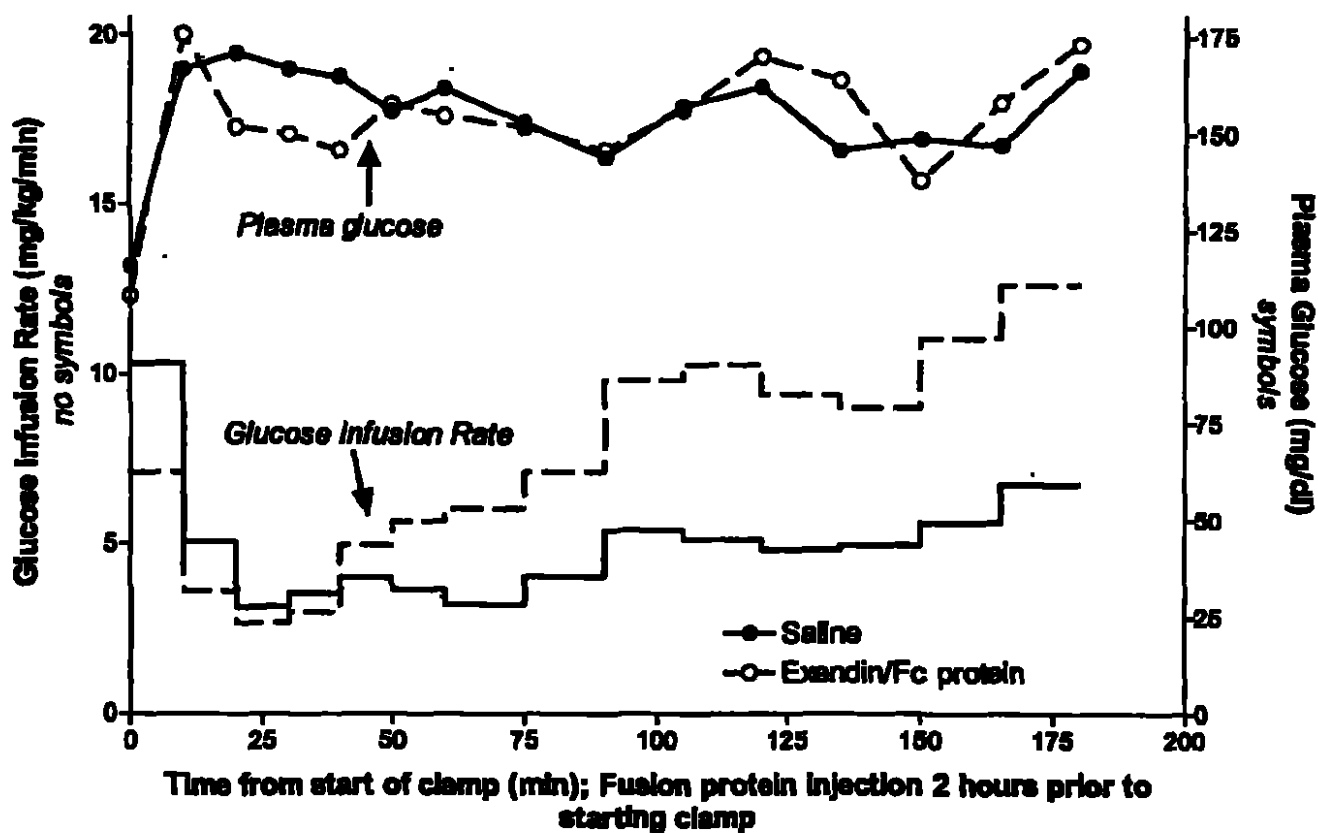


43
140

11/17

Fig. 10

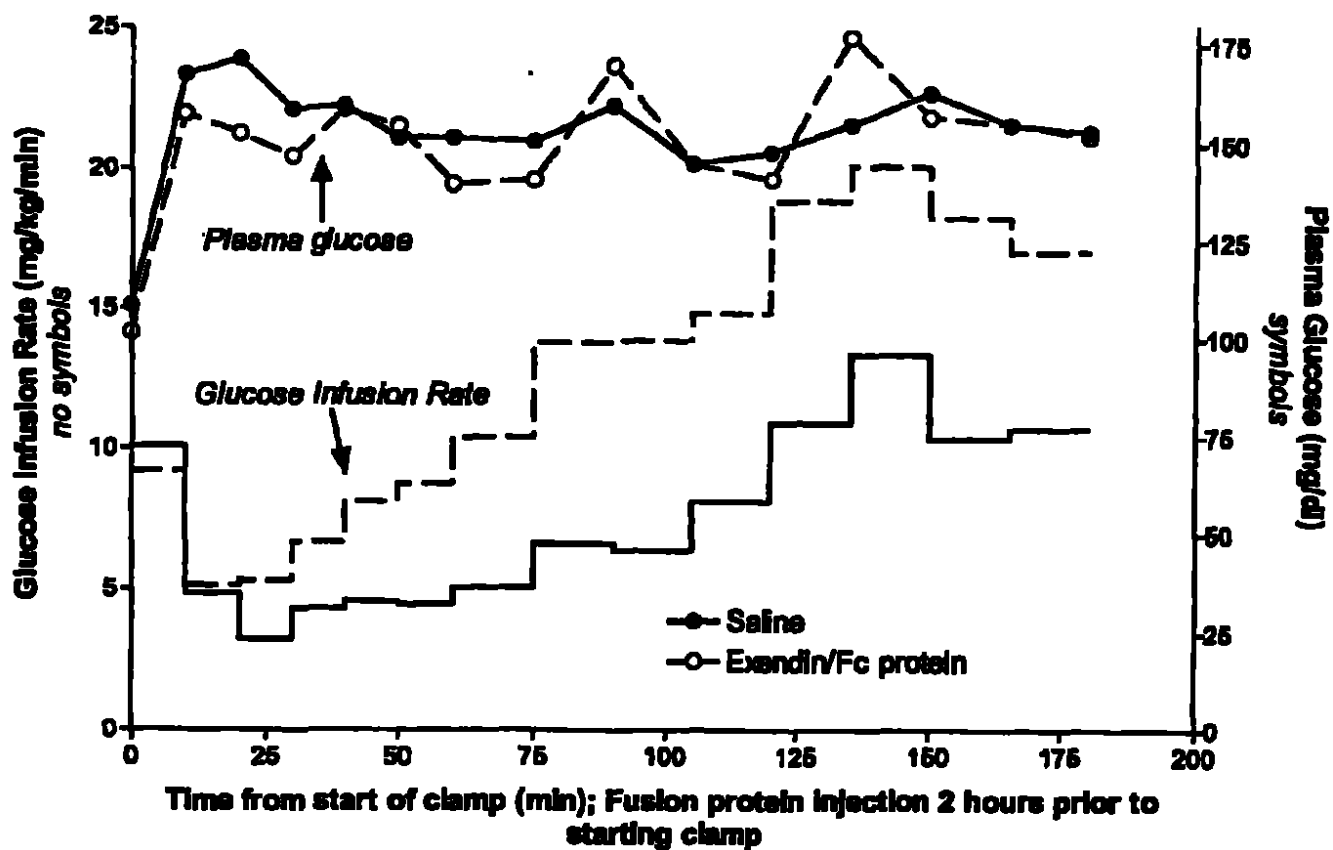
A.



12/17

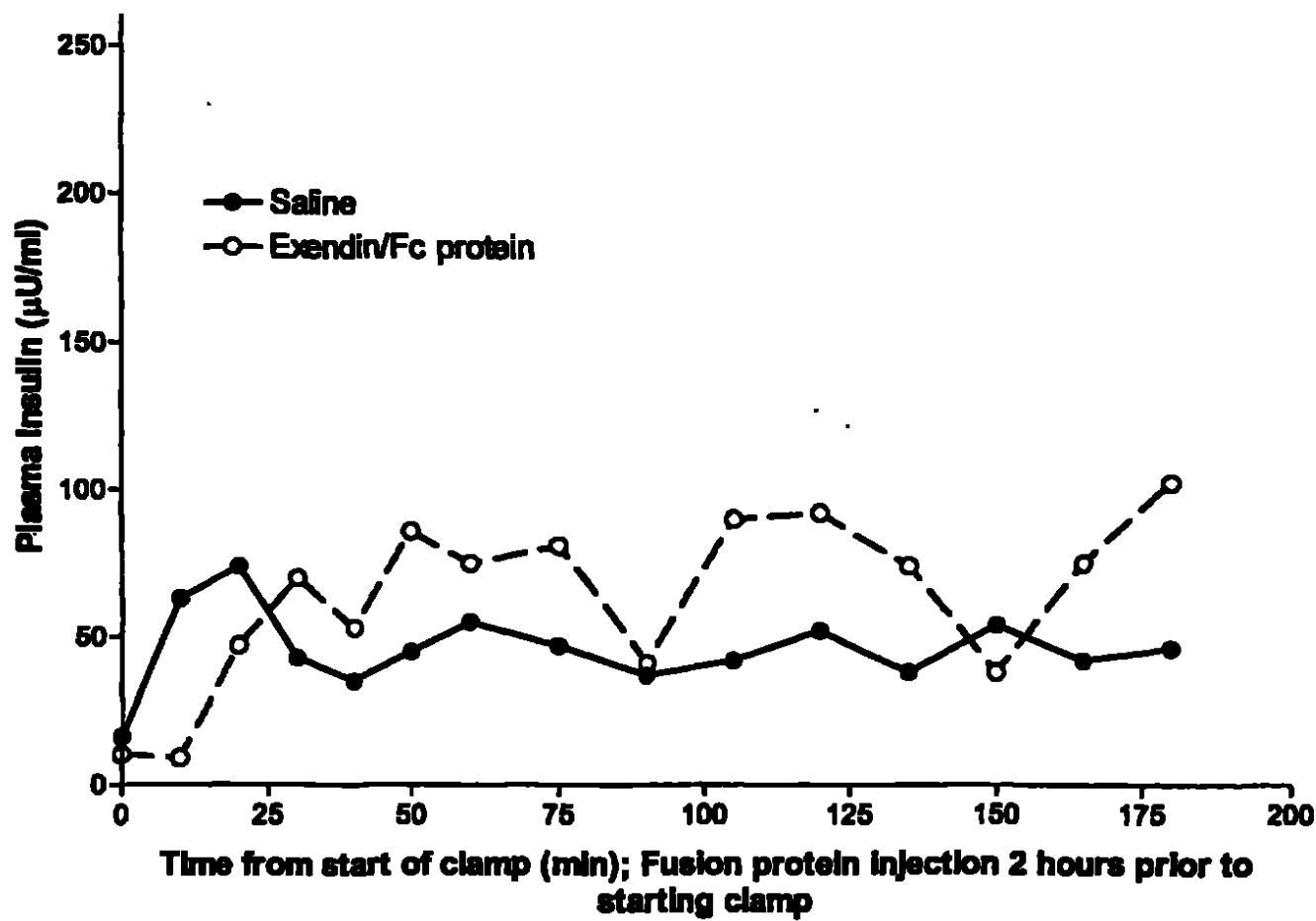
Fig. 10

B.



13/17
Fig. 11

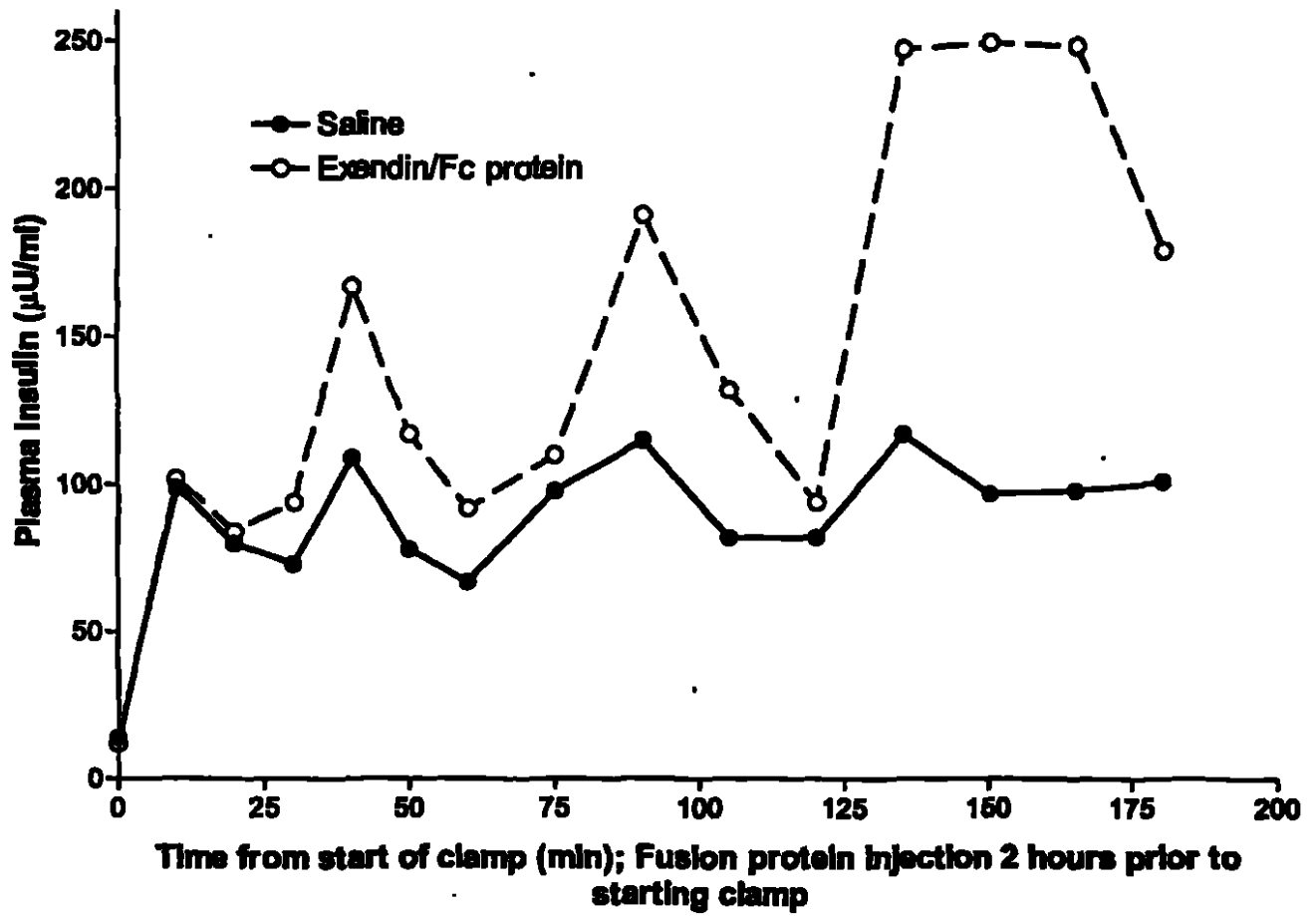
A.



14/17

Fig. 11

B.



15/17

Fig. 12

1
GAGCCCAAATCTTGTGACAAACTCACACATGCCACCGTGCCAGCACC
50
TGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCAAACCCAAGG
100
ACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGAC
150
GTGAGCCACGAAGACCCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGT
200
GGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCA
250
CGTACCGTGTGGTCAGCGTCCTCACCCTCCTGCACCAGGACTGGCTGAAT
300
GGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCAGCCCCCAT
350
CGAGAAAACCATCTCCAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGT
400
ACACCCTGCCCCCATCCCGGGAGGAGATGACCAAGAACCAGGTCAGCCTG
450
ACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGA
500
GAGCAATGGGCAGCCGGAGAACAACCTACAAGACCACGCCTCCCGTGCTGG
550
ACTCCGACGGCTCCTTCTTCCTCTATAGCAAGCTCACCGTGGACAAGAGC
600
AGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCT
650
GCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAT
700
AGT [SEQ ID NO: 33]

F4 8
145

16/17

Fig. 13

1
GATGCCGACAAGAGTGAGGTTGCTCATCGGTTTAAAGATTGGGAGAAGA
50
AAATTTCAAAGCCTTGGTGTGATTGCCTTTGCTCAGTATCTTCAGCAGT
100
GTCCATTTGAAGATCATGTAAAATTAGTGAATGAAGTAACTGAATTTGCA
150
AAAACATGTGTGCTGATGAGTCAGCTGAAAATTGTGACAAATCACTTCA
200
TACCCTTTTTGGAGACAAATTATGCACAGTTGCAACTCTTCGTGAAACCT
250
ATGGTGAAATGGCTGACTGCTGTGCAAAACAAGAACCTGAGAGAAATGAA
300
TGCTTCTTGCAACACAAAGATGACAACCCAAACCTCCCCGATTGGTGAG
350
ACCAGAGGTTGATGTGATGTGCACTGCTTTTCATGACAATGAAGAGACAT
400
TTTTGAAAAATACTTATATGAAATTGCCAGAAGACATCCTTACTTTTAT
450
GCCCCGGAACCTCTTTTCTTTGCTAAAAGGTATAAAGCTGCTTTTACAGA
500
ATGTTGCCAAGCTGCTGATAAAGCTGCCTGCCTGTTGCCAAGCTCGATG
550
AACTTCGGGATGAAGGGAAGGCTTCGTCTGCCAAACAGAGACTCAAGTGT
600
GCCAGTCTCCAAAAATTTGGAGAAAGAGCTTTCAAAGCATGGGCAGTAGC
650
TCGCCTGAGCCAGAGATTTCCCAAAGCTGAGTTTGCGAGAAGTTTCCAAGT
700
TAGTGACAGATCTTACCAAAGTCCACACGGAATGCTGCCATGGAGATCTG
750
CTTGAATGTGCTGATGACAGGGCGGACCTTGCCAAGTATATCTGTGAAAA
800
TCAAGATTCGATCTCCAGTAAACTGAAGGAATGCTGTGAAAAACCTCTGT
850
TGGAAAAATCCCACTGCATTGCCGAAGTGGAATATGATGAGATGCCTGCT
900
GACTTGCCTTCATTAGCTGCTGATTTTGTGAAAGTAAGGATGTTTGCAA
950
AAACTATGCTGAGGCAAAGGATGTCTTCCTGGGCATGTTTTGTATGAAT
1000
ATGCAAGAAGGCATCCTGATTACTCTGTGCTGCTGCTGAGACTTGCC

17/17

Fig. 13 Continued

1050
AAGACATATGAAACCACTCTAGAGAAGTGCTGTGCCGCTGCAGATCCTCA
1100
TGAATGCTATGCCAAAGTGTTGATGAATTTAAACCTCTTGTGGAAGAGC
1150
CTCAGAATTTAATCAAACAAAATTGTGAGCTTTTTGAGCAGCTTGGAGAG
1200
TACAAATTCCAGAATGCGCTATTAGTTCGTTACACCAAGAAAGTACCCCA
1250
AGTGTCAACTCCAACCTCTGTAGAGGTCTCAAGAAACCTAGGAAAAGTGG
1300
GCAGCAAATGTTGTAAACATCCTGAAGCAAAAAGAATGCCCTGTGCAGAA
1350
GACTATCTATCCGTGGTCCTGAACCAGTTATGTGTGTTGCATGAGAAAAC
1400
GCCAGTAAGTGACAGAGTCACCAAATGCTGCACAGAATCCTTGGTGAACA
1450
GGCGACCATGCTTTTCAGCTCTGGAAGTCGATGAAACATACGTTCCCAA
1500
GAGTTTAATGCTGAAACATTCACCTTCCATGCAGATATATGCACACTTTC
1550
TGAGAAGGAGAGACAAATCAAGAAACAACTGCACTTGTTGAGCTCGTGA
1600
AACACAAGCCCAAGGCAACAAAAGAGCAACTGAAAGCTGTTATGGATGAT
1650
TTCCGAGCTTTTGTAGAGAAGTGCTGCAAGGCTGACGATAAGGAGACCTG
1700
CTTTGCCGAGGAGGGTAAAAAAGTTGTTGCTGCAAGTCAAGCTGCCTTAG
1750
GCTTATAATGAC [SEQ ID NO: 35]

X-13991.ST25.txt
SEQUENCE LISTING150
149

<110> Eli Lilly and Company

<120> GLP-1 FUSION PROTEINS

<130> X-13991

<150> US 60/251,954

<151> 2000-06-12

<160> 35

<170> PatentIn version 3.1

<210> 1

<211> 31

<212> PRT

<213> Homo sapiens

<400> 1

His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly
20 25 30

<210> 2

<211> 39

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

X-13991.ST25.txt

151
148

<222> (2)..(2)

<223> Xaa at position 2 is Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (3)..(3)

<223> Xaa at position 3 is Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (5)..(5)

<223> Xaa at position 5 is Thr, Ala, Gly, Ser, Leu, Ile, Val, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (8)..(8)

<223> Xaa at position 8 is Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> Xaa at position 10 is Val, Ala, Gly, Ser, Thr, Leu, Ile, Tyr, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (11)..(11)

<223> Xaa at position 11 is Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (12)..(12)

152
149

X-13991.ST25.txt

<223> Xaa at position 12 is Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, Lys, Trp or Tyr;

<220>

<221> MISC_FEATURE

<222> (13)..(13)

<223> Xaa at position 13 is Tyr, Phe, Trp, Glu, Asp, Gln, or Lys;

<220>

<221> MISC_FEATURE

<222> (14)..(14)

<223> Xaa at position 14 is Leu, Ala, Gly, Ser, Thr, Ile, Val, Glu, Asp, Met, Lys, Trp or Tyr;

<220>

<221> MISC_FEATURE

<222> (15)..(15)

<223> Xaa at position 15 is Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (16)..(16)

<223> Xaa at position 16 is Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, Trp or Lys;

<220>

<221> MISC_FEATURE

<222> (17)..(17)

<223> Xaa at position 17 is Gln, Asn, Arg, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (18)..(18)

H3
150

X-13991.ST25.txt

<223> Xaa at position 18 is Ala, Gly, Ser, Thr, Leu, Ile, Val, Arg, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (19)..(19)

<223> Xaa at position 19 is Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (20)..(20)

<223> Xaa at position 20 is Lys, Arg, Gln, Glu, Asp, or His;

<220>

<221> MISC_FEATURE

<222> (21)..(21)

<223> Xaa at position 21 is Leu, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (24)..(24)

<223> Xaa at position 24 is Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (25)..(25)

<223> Xaa at position 25 is Trp, Phe, Tyr, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

<222> (26)..(26)

X-13991.ST25.txt

<223> Xaa at position 26 is Leu, Gly, Ala, Ser, Thr, Ile, Val, Glu, Asp
, or Lys;

<220>

<221> MISC_FEATURE

<222> (27)..(27)

<223> Xaa at position 27 is Val, Gly, Ala, Ser, Thr, Leu, Ile, Glu, Asp
, or Lys;

<220>

<221> MISC_FEATURE

<222> (28)..(28)

<223> Xaa at position 28 is Asn, Lys, Arg, Glu, Asp, or His;

<220>

<221> MISC_FEATURE

<222> (29)..(29)

<223> Xaa at position 29 is Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp
, or Lys;

<220>

<221> MISC_FEATURE

<222> (30)..(30)

<223> Xaa at position 30 is Gly, Arg, Lys, Glu, Asp, or His;

<220>

<221> MISC_FEATURE

<222> (31)..(31)

<223> Xaa at position 31 is Pro, Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu
, Asp, or Lys, or is deleted;

<220>

<221> MISC_FEATURE

X-13991.ST25.txt

15f
152

<222> (32)..(32)

<223> Xaa at position 32 is Ser, Arg, Lys, Glu, Asp, or His, or is deleted;

<220>

<221> MISC_FEATURE

<222> (33)..(33)

<223> Xaa at position 33 is Ser, Arg, Lys, Glu, Asp, or His, or is deleted;

<220>

<221> MISC_FEATURE

<222> (34)..(34)

<223> Xaa at position 34 is Gly, Asp, Glu, or Lys, or is deleted;

<220>

<221> MISC_FEATURE

<222> (35)..(35)

<223> Xaa at position 35 is Ala, Phe, Trp, Tyr, Glu, Asp, or Lys, or is deleted;

<220>

<221> MISC_FEATURE

<222> (36)..(36)

<223> Xaa at position 36 is Ser, Pro, Lys, Glu, or Asp, or is deleted;

<220>

<221> MISC_FEATURE

<222> (37)..(37)

<223> Xaa at position 37 is Ser, Pro, Glu, Asp, or Lys, or is deleted;

<220>

<221> MISC_FEATURE

X-13991.ST25.txt

<222> (38)..(38)

<223> Xaa at position 38 is Gly, Pro, Glu, Asp, or Lys, or is deleted;

<220>

<221> MISC_FEATURE

<222> (39)..(39)

<223> Xaa at position 39 is Ala, Ser, Val, Glu, Asp, or Lys, or is deleted;

<400> 2

His Xaa Xaa Gly Xaa Phe Thr Xaa Asp Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Phe Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35

<210> 3

<211> 32

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> Xaa at position 1 is L-histidine, D-histidine, or is deleted.

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa at position 2 is Gly, Ala, Val, Leu, Ile, Ser, or Thr;

<220>

X-13991.ST25.txt

~~157~~
154

<221> MISC_FEATURE

<222> (3)..(3)

<223> xaa at position 3 is Thr, Ser, Arg, Lys, Trp, Phe, Tyr, Glu, or His;

<220>

<221> MISC_FEATURE

<222> (5)..(5)

<223> xaa at position 5 is Asp, Glu, Arg, Thr, Ala, Lys, or His;

<220>

<221> MISC_FEATURE

<222> (6)..(6)

<223> xaa at position 6 is His, Trp, Phe, or Tyr;

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> xaa at position 10 is Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Tyr, Glu, or Ala;

<220>

<221> MISC_FEATURE

<222> (12)..(12)

<223> xaa at position 12 is His, Pro, Asp, Glu, Arg, Ser, Ala, or Lys;

<220>

<221> MISC_FEATURE

<222> (13)..(13)

<223> xaa at position 13 is Gly, Asp, Glu, Gln, Asn, Lys, Arg, or Cys;

<220>

158
155

X-13991.ST25.txt

<221> MISC_FEATURE

<222> (17)..(17)

<223> Xaa at position 17 is His, Asp, Lys, Glu, Gln, or Arg;

<220>

<221> MISC_FEATURE

<222> (18)..(18)

<223> Xaa at position 18 is Glu, Arg, Ala, or Lys;

<220>

<221> MISC_FEATURE

<222> (20)..(20)

<223> Xaa at position 20 is Trp, Tyr, Phe, Asp, Lys, Glu, or His;

<220>

<221> MISC_FEATURE

<222> (21)..(21)

<223> Xaa at position 21 is Ala, Glu, His, Phe, Tyr, Trp, Arg, or Lys;

<220>

<221> MISC_FEATURE

<222> (24)..(24)

<223> Xaa at position 24 is Ala, Glu, Asp, Ser, or His;

<220>

<221> MISC_FEATURE

<222> (25)..(25)

<223> Xaa at position 25 is Asp, Glu, Ser, Thr, Arg, Trp, or Lys;

<220>

<221> MISC_FEATURE

X-13991.ST25.txt

156

<222> (27)..(27)

<223> Xaa at position 27 is Asp, Arg, Val, Lys, Ala, Gly, or Glu;

<220>

<221> MISC_FEATURE

<222> (28)..(28)

<223> Xaa at position 28 is Glu, Lys, or Asp;

<220>

<221> MISC_FEATURE

<222> (29)..(29)

<223> Xaa at position 29 is Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, or Glu;

<220>

<221> MISC_FEATURE

<222> (30)..(30)

<223> Xaa at position 30 is Thr, Ser, Asp, Trp, Tyr, Phe, Arg, Glu, or His;

<220>

<221> MISC_FEATURE

<222> (31)..(31)

<223> Xaa at position 31 is Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, or is deleted.

<220>

<221> MISC_FEATURE

<222> (32)..(32)

<223> Xaa at position 31 is Pro or is deleted.

<400> 3

Xaa Xaa Xaa Gly Xaa Xaa Thr Ser Asp Xaa Ser Xaa Xaa Leu Glu Gly
1 5 10 15Xaa Xaa Ala Xaa Xaa Phe Ile Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30

X-13991.ST25.txt

~~110~~
159

<210> 4

<211> 32

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> Xaa at position 1 is L-histidine, D-histidine, or is deleted.

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa at position 2 is Gly, Ala, Val, Leu, Ile, Ser, or Thr;

<220>

<221> MISC_FEATURE

<222> (5)..(5)

<223> Xaa at position 5 is Asp, Glu, Arg, Thr, Ala, Lys, or His;

<220>

<221> MISC_FEATURE

<222> (6)..(6)

<223> Xaa at position 6 is His, Trp, Phe, or Tyr;

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> Xaa at position 10 is Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu, or Ala;

158
H

X-13991.ST25.txt

<220>

<221> MISC_FEATURE

<222> (16)..(16)

<223> xaa at position 16 is Gly, Asp, Glu, Gln, Asn, Lys, Arg, or Cys;

<220>

<221> MISC_FEATURE

<222> (17)..(17)

<223> xaa at position 17 is His, Asp, Lys, Glu, or Gln;

<220>

<221> MISC_FEATURE

<222> (18)..(18)

<223> xaa at position 18 is Glu, His, Ala, or Lys;

<220>

<221> MISC_FEATURE

<222> (19)..(19)

<223> xaa at position 19 is Asp, Lys, Glu, or His;

<220>

<221> MISC_FEATURE

<222> (21)..(21)

<223> xaa at position 21 is Ala, Glu, His, Phe, Tyr, Trp, Arg, or Lys;

<220>

<221> MISC_FEATURE

<222> (24)..(24)

<223> xaa at position 24 is Ala, Glu, Asp, Ser, or His;

159
~~162~~

X-13991.ST25.txt

<220>

<221> MISC_FEATURE

<222> (27)..(27)

<223> Xaa at position 27 is Asp, Arg, Val, Lys, Ala, Gly, or Glu;

<220>

<221> MISC_FEATURE

<222> (28)..(28)

<223> Xaa at position 28 is Glu, Lys, or Asp;

<220>

<221> MISC_FEATURE

<222> (29)..(29)

<223> Xaa at position 29 is Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, or Glu;

<220>

<221> MISC_FEATURE

<222> (30)..(30)

<223> Xaa at position 30 is Arg, Glu, or His;

<220>

<221> MISC_FEATURE

<222> (31)..(31)

<223> Xaa at position 31 is Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, or is deleted.

<220>

<221> MISC_FEATURE

<222> (32)..(32)

<223> Xaa at position 32 is Pro, or is deleted.

<400> 4

Xaa Xaa Glu Gly Xaa Xaa Thr Ser Asp Xaa Ser Ser Tyr Leu Glu Xaa
1 5 10 15

X-13991.ST25.txt

160

~~163~~

Xaa Xaa Xaa Lys Xaa Phe Ile Xaa Trp Leu Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30

<210> 5

<211> 32

<212> PRT

<213> Artificial sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> Xaa at position 1 is L-histidine, D-histidine, or is deleted.

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa at position 2 is Gly, Ala, Val, Leu, Ile, Ser, Met, or Thr;

<220>

<221> MISC_FEATURE

<222> (6)..(6)

<223> Xaa at position 6 is His, Trp, Phe, or Tyr;

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> Xaa at position 10 is Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu, or Ala;

<220>

<221> MISC_FEATURE

<222> (16)..(16)

161
+69

X-13991.ST25.txt

<223> Xaa at position 16 is Gly, Asp, Glu, Gln, Asn, Lys, Arg, or Cys;

<220>

<221> MISC_FEATURE

<222> (17)..(17)

<223> Xaa at position 17 is His, Asp, Lys, Glu, or Gln;

<220>

<221> MISC_FEATURE

<222> (20)..(20)

<223> Xaa at position 20 is Asp, Lys, Glu, or His;

<220>

<221> MISC_FEATURE

<222> (24)..(24)

<223> Xaa at position 24 is Ala, Glu, Asp, Ser, or His;

<220>

<221> MISC_FEATURE

<222> (29)..(29)

<223> Xaa at position 29 is Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His, or Glu;

<220>

<221> MISC_FEATURE

<222> (31)..(31)

<223> Xaa at position 31 is Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, or is deleted.

<220>

<221> MISC_FEATURE

<222> (32)..(32)

<223> Xaa at position 32 is Pro or is deleted.

X-13991.ST25.txt

<400> 5

Xaa Xaa Glu Gly Thr Xaa Thr Ser Asp Xaa Ser Ser Tyr Leu Glu Xaa
1 5 10 15

Xaa Ala Ala Xaa Glu Phe Ile Xaa Trp Leu Val Lys Xaa Arg Xaa Xaa
20 25 30

<210> 6

<211> 32

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> Xaa at position 1 is L-histidine, D-histidine, or is deleted.

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa at position 2 is Gly, Ala, Val, Leu, Ile, Ser, or Thr;

<220>

<221> MISC_FEATURE

<222> (16)..(16)

<223> Xaa at position 16 is Gly, Asp, Glu, Gln, Asn, Lys, Arg, or Cys;

<220>

<221> MISC_FEATURE

<222> (17)..(17)

<223> Xaa at position 17 is His, Asp, Lys, Glu, or Gln;

<220>

<221> MISC_FEATURE

162
165

X-13991.ST25.txt

163
Ht

<222> (18)..(18)

<223> xaa at position 18 is Ala, Glu, His, Phe, Tyr, Trp, Arg, or Lys;

<220>

<221> MISC_FEATURE

<222> (24)..(24)

<223> xaa at position 24 is Ala, Glu, Asp, Ser, or His;

<220>

<221> MISC_FEATURE

<222> (31)..(31)

<223> xaa at position 31 is Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, or is deleted.

<220>

<221> MISC_FEATURE

<222> (32)..(32)

<223> xaa at position 32 is Pro or is deleted.

<400> 6

xaa xaa Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu xaa
1 5 10 15xaa xaa Ala Lys Glu Phe Ile xaa Trp Leu Val Lys Gly Arg xaa xaa
20 25 30

<210> 7

<211> 31

<212> PRT

<213> Artificial sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

X-13991.ST25.txt

<222> (1)..(1)

<223> Xaa at position 1 is L-histidine, D-histidine, or is deleted.

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa at position 2 is Ala, Gly, Val, Thr, Ile, and alpha-methyl-Ala;

<220>

<221> MISC_FEATURE

<222> (15)..(15)

<223> Xaa at position 15 is Glu, Gln, Ala, Thr, Ser, and Gly;

<220>

<221> MISC_FEATURE

<222> (21)..(21)

<223> Xaa at position 21 is Glu, Gln, Ala, Thr, Ser, and Gly;

<400> 7

Xaa Xaa Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Xaa Gly
1 5 10 15Gln Ala Ala Lys Xaa Phe Ile Ala Trp Leu Val Lys Gly Arg Gly
20 25 30

<210> 8

<211> 30

<212> PRT

<213> Artificial sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (19)..(19)

X-13991.ST25.txt

<223> xaa at position 19 is Lys or Arg;

<220>

<221> MOD_RES

<222> (27)..(27)

<223> ACETYLATION

<220>

<221> MISC_FEATURE

<222> (30)..(30)

<223> xaa at position 30 is Gly;

<220>

<221> MOD_RES

<222> (30)..(30)

<223> AMIDATION

<400> 8

Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly Gln
1 5 10 15Ala Ala Xaa Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Xaa
20 25 30

<210> 9

<211> 39

<212> PKT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 9

His Ser Asp Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

X-13991.ST25.txt

Ser Gly Ala Pro Pro Pro Ser
35

<210> 10

<211> 39

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 10

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser
35

<210> 11

<211> 39

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> xaa at position 1 is L-histidine, D-histidine, or is deleted.

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> xaa at position 2 is Gly, Ala, or Val;

<220>

X-13991.ST25.txt

<221> MISC_FEATURE

<222> (10)..(10)

<223> Xaa at position 10 is Leu or Val;

<220>

<221> MISC_FEATURE

<222> (12)..(12)

<223> Xaa at position 12 is Lys or Ser;

<220>

<221> MISC_FEATURE

<222> (13)..(13)

<223> Xaa at position 13 is Gln or Tyr;

<220>

<221> MISC_FEATURE

<222> (14)..(14)

<223> Xaa at position 14 is Met or Leu;

<220>

<221> MISC_FEATURE

<222> (16)..(16)

<223> Xaa at position 16 is Glu or Gln;

<220>

<221> MISC_FEATURE

<222> (17)..(17)

<223> Xaa at position 17 is Glu or Gln;

<220>

<221> MISC_FEATURE

<222> (19)..(19)

<223> Xaa at position 19 is Val or Ala;

x-13991.ST25.txt

<220>

<221> MISC_FEATURE

<222> (20)..(20)

<223> Xaa at position 20 is Arg or Lys;

<220>

<221> MISC_FEATURE

<222> (21)..(21)

<223> Xaa at position 21 is Leu or Glu;

<220>

<221> MISC_FEATURE

<222> (24)..(24)

<223> Xaa at position 24 is Glu or Ala;

<220>

<221> MISC_FEATURE

<222> (27)..(27)

<223> Xaa at position 27 is Val or Lys;

<220>

<221> MISC_FEATURE

<222> (28)..(28)

<223> Xaa at position 28 is Asn or Lys;

<220>

<221> MISC_FEATURE

<222> (30)..(30)

<223> Xaa at position 30 is Gly or Arg; and

<220>

<221> MISC_FEATURE

<222> (31)..(31)

<223> Xaa at position 31 is Gly, or Pro;

X-13991.ST25.txt

<220>

<221> MISC_FEATURE

<222> (32)..(32)

<223> xaa at position 32 is Ser, or is absent.

<220>

<221> MISC_FEATURE

<222> (33)..(33)

<223> xaa at position 33 is Ser, or is absent.

<220>

<221> MISC_FEATURE

<222> (34)..(34)

<223> xaa at position 34 is Gly, or is absent.

<220>

<221> MISC_FEATURE

<222> (35)..(35)

<223> xaa at position 35 is Ala, or is absent.

<220>

<221> MISC_FEATURE

<222> (36)..(36)

<223> xaa at position 36 is Pro, or is absent.

<220>

<221> MISC_FEATURE

<222> (37)..(37)

<223> xaa at position 37 is Pro, or is absent.

<220>

<221> MISC_FEATURE

<222> (38)..(38)

170

X-13991.ST25.txt

<223> Xaa at position 38 is Pro, or is absent.

<220>

<221> MISC_FEATURE

<222> (39)..(39)

<223> Xaa at position 39 is Pro, or is absent.

<220>

<221> MISC_FEATURE

<222> (39)..(39)

<223> Xaa at position 39 is Ser, or is absent.

<400> 11

Xaa Xaa Glu Glu Thr Phe Thr Ser Asp Xaa Ser Xaa Xaa Xaa Glu Xaa
1 5 10 15

Xaa Ala Xaa Xaa Xaa Phe Ile Xaa Trp Leu Xaa Xaa Gly Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35

<210> 12

<211> 31

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> Xaa at position 1 is L-histidine, D-histidine, or is deleted.

<220>

<221> MISC_FEATURE

<222> (2)..(2)

171

X-13991.ST25.txt

<223> Xaa at position 2 is Ala, Gly, Val, Leu, Ile, Ser, or Thr;

<220>

<221> MISC_FEATURE

<222> (6)..(6)

<223> Xaa at position 6 is Phe, Trp, or Tyr;

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> Xaa at position 10 is Val, Trp, Ile, Leu, Phe, or Tyr;

<220>

<221> MISC_FEATURE

<222> (12)..(12)

<223> Xaa at position 12 is Ser, Trp, Tyr, Phe, Lys, Ile, Leu, Val;

<220>

<221> MISC_FEATURE

<222> (13)..(13)

<223> Xaa at position 13 is Tyr, Trp, or Phe;

<220>

<221> MISC_FEATURE

<222> (14)..(14)

<223> Xaa at position 14 is Leu, Phe, Tyr, or Trp;

<220>

<221> MISC_FEATURE

<222> (16)..(16)

<223> Xaa at position 16 is Gly, Glu, Asp, or Lys;

<220>

<221> MISC_FEATURE

X-13991.ST25.txt

<222> (19)..(19)

<223> Xaa at position 19 is Ala, Val, Ile, or Leu;

<220>

<221> MISC_FEATURE

<222> (21)..(21)

<223> Xaa at position 21 is Glu, Ile, or Ala;

<220>

<221> MISC_FEATURE

<222> (24)..(24)

<223> Xaa at position 24 is Ala or Glu;

<220>

<221> MISC_FEATURE

<222> (27)..(27)

<223> Xaa at position 27 is Val or Ile; and

<220>

<221> MOD_RES

<222> (30)..(30)

<223> AMIDATION

<220>

<221> MISC_FEATURE

<222> (31)..(31)

<223> Xaa at position 31 is Gly, His or is absent.

<400> 12

Xaa Xaa Glu Gly Thr Xaa Thr Ser Asp Xaa Ser Xaa Xaa Xaa Glu Xaa
1 5 10 15Gln Ala Xaa Lys Xaa Phe Ile Xaa Trp Leu Xaa Lys Gly Arg Xaa
20 25 30

<210> 13

173

X-13991.ST25.txt

<211> 616

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 13

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Asp
20 25 30Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu
35 40 45Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln
50 55 60Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe
65 70 75 80Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser
85 90 95Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg
100 105 110Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu
115 120 125Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro
130 135 140Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp
145 150 155 160Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg
165 170 175His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr
180 185 190Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys
195 200 205Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser Ser
210 215 220
27

X-13991.ST25.txt

Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg
225 230 235 240

Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys
245 250 255

Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val
260 265 270

His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg
275 280 285

Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser
290 295 300

Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys
305 310 315 320

Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu
325 330 335

Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu
340 345 350

Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg
355 360 365

His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr
370 375 380

Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys
385 390 395 400

Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln
405 410 415

Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr
420 425 430

Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln
435 440 445

Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val
450 455 460

Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala
465 470 475 480

Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu
485 490 495

x-13991.ST25.txt

Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu
500 505 510

Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr
515 520 525

Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile
530 535 540

Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu
545 550 555 560

Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys
565 570 575

Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala
580 585 590

Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala
595 600 605

Ala Ser Gln Ala Ala Leu Gly Leu
610 615

<210> 14

<211> 631

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 14

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Gly
20 25 30

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ala
35 40 45

His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn
50 55 60

Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Cys
65 70 75 80

X-13991.ST25.txt

Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala
85 90 95

Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu
100 105 110

His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu
115 120 125

Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg
130 135 140

Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg
145 150 155 160

Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp Asn
165 170 175

Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His
180 185 190

Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys
195 200 205

Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu
210 215 220

Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser Ser Ala
225 230 235 240

Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala
245 250 255

Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala
260 265 270

Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val His
275 280 285

Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala
290 295 300

Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys
305 310 315 320

Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile
325 330 335

Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala
340 345 350

X-13991.ST25.txt

Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala
 355 360 365
 Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His
 370 375 380
 Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu
 385 390 395 400
 Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr
 405 410 415
 Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln Asn
 420 425 430
 Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys
 435 440 445
 Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val
 450 455 460
 Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly
 465 470 475 480
 Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu
 485 490 495
 Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys
 500 505 510
 Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val
 515 520 525
 Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val
 530 535 540
 Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys
 545 550 555 560
 Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val
 565 570 575
 Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala
 580 585 590
 Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp
 595 600 605
 Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala
 610 615 620
 31

X-13991.ST25.txt

Ser Gln Ala Ala Leu Gly Leu
625 630

<210> 15

<211> 640

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 15

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly
35 40 45

Gly Ser Gly Gly Gly Gly Ser Asp Ala His Lys Ser Glu Val Ala His
50 55 60

Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile
65 70 75 80

Ala Phe Ala Gln Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val Lys
85 90 95

Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu
100 105 110

Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys
115 120 125

Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp
130 135 140

Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His
145 150 155 160

Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp
165 170 175

Val Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys
180 185 190

1 32

X-13991.ST25.txt

179

Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu
195 200 205

Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys
210 215 220

Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu
225 230 235 240

Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala
245 250 255

Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala
260 265 270

Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys
275 280 285

Leu Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp
290 295 300

Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys
305 310 315 320

Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys
325 330 335

Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu
340 345 350

Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys
355 360 365

Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met
370 375 380

Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu
385 390 395 400

Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys
405 410 415

Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe
420 425 430

Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu
435 440 445

Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val
450 455 460

X-13991.ST25.txt

Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu
465 470 475 480

Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro
485 490 495

Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu
500 505 510

Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val
515 520 525

Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser
530 535 540

Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu
545 550 555 560

Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg
565 570 575

Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro
580 585 590

Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala
595 600 605

Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala
610 615 620

Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
625 630 635 640

<210> 16

<211> 624

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 16

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

X-13991.ST25.txt

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

X-13991.ST25.txt

Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys
 305 310 315 320
 Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu
 325 330 335
 Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys
 340 345 350
 Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met
 355 360 365
 Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu
 370 375 380
 Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys
 385 390 395 400
 Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe
 405 410 415
 Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu
 420 425 430
 Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val
 435 440 445
 Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu
 450 455 460
 Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro
 465 470 475 480
 Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu
 485 490 495
 Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val
 500 505 510
 Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser
 515 520 525
 Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu
 530 535 540
 Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg
 545 550 555 560
 Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro
 565 570 575

X-13991.ST25.txt

Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala
580 585 590

Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala
595 600 605

Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
610 615 620

<210> 17

<211> 640

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 17

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly
35 40 45

Gly Ser Gly Gly Gly Gly Ser Asp Ala His Lys Ser Glu Val Ala His
50 55 60

Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile
65 70 75 80

Ala Phe Ala Gln Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val Lys
85 90 95

Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu
100 105 110

Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys
115 120 125

Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp
130 135 140

Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His
145 150 155 160

37

X-13991.ST25.txt

Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp
 165 170 175
 Val Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys
 180 185 190
 Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu
 195 200 205
 Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys
 210 215 220
 Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu
 225 230 235 240
 Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala
 245 250 255
 Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala
 260 265 270
 Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys
 275 280 285
 Leu Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp
 290 295 300
 Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys
 305 310 315 320
 Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys
 325 330 335
 Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu
 340 345 350
 Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys
 355 360 365
 Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met
 370 375 380
 Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu
 385 390 395 400
 Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys
 405 410 415
 Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe
 420 425 430

X-13991.ST25.txt

Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu
435 440 445

Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val
450 455 460

Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu
465 470 475 480

Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro
485 490 495

Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu
500 505 510

Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val
515 520 525

Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser
530 535 540

Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu
545 550 555 560

Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg
565 570 575

Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro
580 585 590

Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala
595 600 605

Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala
610 615 620

Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
625 630 635 640

<210> 18

<211> 264

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

186

X-13991.ST25.txt

<400> 18

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

x-13991.ST25.txt

187

<210> 19

<211> 272

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 19

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
 20 25 30

Ser Gly Ala Pro Pro Pro Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr
 35 40 45

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
 50 55 60

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
 65 70 75 80

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
 85 90 95

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
 100 105 110

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
 115 120 125

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
 130 135 140

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
 145 150 155 160

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
 165 170 175

Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
 180 185 190

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
 195 200 205

191
188

X-13991.ST25.txt

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
210 215 220

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
225 230 235 240

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
245 250 255

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
260 265 270

<210> 20

<211> 264

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 20

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ala
20 25 30

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
35 40 45

Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
50 55 60

Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
65 70 75 80

Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
85 90 95

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
100 105 110

Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln
115 120 125

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
130 135 140

X-13991.ST25.txt

142
189

Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
145 150 155 160

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
165 170 175

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
180 185 190

Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
195 200 205

Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
210 215 220

Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
225 230 235 240

Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
245 250 255

Ser Leu Ser Leu Ser Pro Gly Lys
260

<210> 21

<211> 272

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 21

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr
35 40 45

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
50 55 60

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
65 70 75 80

43

X-13991.ST25.txt

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
85 90 95
Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
100 105 110
Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
115 120 125
Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
130 135 140
Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
145 150 155 160
Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
165 170 175
Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
180 185 190
Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
195 200 205
Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
210 215 220
Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
225 230 235 240
Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
245 250 255
Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
260 265 270

<210> 22

<211> 272

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 22

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

44

X-13991.ST25.txt

194
191

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Ser Ser Gly Ala Ala Glu Pro Lys Ser Cys Asp Lys Thr
35 40 45

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
50 55 60

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
65 70 75 80

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
85 90 95

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
100 105 110

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
115 120 125

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
130 135 140

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
145 150 155 160

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
165 170 175

Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
180 185 190

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
195 200 205

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
210 215 220

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
225 230 235 240

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
245 250 255

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
260 265 270

<210> 23

X-13991.ST25.txt

<211> 287

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 23

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
35 40 45

Ser Gly Gly Gly Gly Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr His
50 55 60

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
65 70 75 80

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
85 90 95

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu
100 105 110

Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
115 120 125

Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser
130 135 140

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
145 150 155 160

Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
165 170 175

Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
180 185 190

Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
195 200 205

Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
210 215 220

X-13991.ST25.txt

191
193

Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
225 230 235 240

Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
245 250 255

Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
260 265 270

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
275 280 285

<210> 24

<211> 284

<212> PRT

<213> Artificial sequence

<220>

<223> synthetic construct ..

<400> 24

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Ser Gly Gly Gly Gly
35 40 45

Ser Gly Gly Gly Gly Ser Ala Glu Ser Lys Tyr Gly Pro Pro Cys Pro
50 55 60

Ser Cys Pro Ala Pro Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe
65 70 75 80

Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val
85 90 95

Thr Cys Val Val Val Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe
100 105 110

Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro
115 120 125

Arg Glu Glu Gln Phe Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr
130 135 140 47

X-13991.ST25.txt

194

Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val
145 150 155 160

Ser Asn Lys Gly Leu Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala
165 170 175

Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln
180 185 190

Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly
195 200 205

Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro
210 215 220

Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser
225 230 235 240

Phe Phe Leu Tyr Ser Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu
245 250 255

Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His
260 265 270

Tyr Thr Gln Lys Ser Leu Ser Leu Ser Leu Gly Lys
275 280

<210> 25

<211> 302

<212> PRT

<213> Artificial sequence

<220>

<223> synthetic construct

<400> 25

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
35 40 45

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
50 55 60

48

X-13991.ST25.txt

Gly Gly Gly Gly Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr His Thr
65 70 75 80

Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe
85 90 95

Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro
100 105 110

Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val
115 120 125

Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr
130 135 140

Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val
145 150 155 160

Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys
165 170 175

Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser
180 185 190

Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro
195 200 205

Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val
210 215 220

Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly
225 230 235 240

Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp
245 250 255

Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp
260 265 270

Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His
275 280 285

Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
290 295 300

<210> 26

<211> 294

<212> PRT

199
196

X-13991.ST25.txt

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 26

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

X-13991.ST25.txt

220
197

Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys
245 250 255

Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys
260 265 270

Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu
275 280 285

Ser Leu Ser Pro Gly Lys
290

<210> 27

<211> 280

<212> PRT

<213> Artificial sequence

<220>

<223> synthetic construct

<400> 27

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Ser Ser Gly Ala Pro Pro Pro Ser Ala
35 40 45

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
50 55 60

Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
65 70 75 80

Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
85 90 95

Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
100 105 110

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
115 120 125

Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln
130 135 140
51

X-13991.ST25.txt

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
145 150 155 160

Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
165 170 175

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
180 185 190

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
195 200 205

Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
210 215 220

Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
225 230 235 240

Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
245 250 255

Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
260 265 270

Ser Leu Ser Leu Ser Pro Gly Lys
275 280

<210> 28

<211> 287

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 28

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Val Lys Glu Phe Ile Ala Trp Leu Ile Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
35 40 45

Ser Gly Gly Gly Gly Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr His
50 55 60 52

X-13991.ST25.txt

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
65 70 75 80

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
85 90 95

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu
100 105 110

Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
115 120 125

Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser
130 135 140

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
145 150 155 160

Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
165 170 175

Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
180 185 190

Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
195 200 205

Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
210 215 220

Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
225 230 235 240

Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
245 250 255

Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
260 265 270

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
275 280 285

<210> 29

<211> 272

<212> PKT

<213> Artificial sequence

X-13991.ST25.txt

<220>

<223> synthetic construct

<400> 29

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr
35 40 45

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
50 55 60

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
65 70 75 80

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
85 90 95

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
100 105 110

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
115 120 125

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
130 135 140

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
145 150 155 160

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
165 170 175

Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
180 185 190

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
195 200 205

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
210 215 220

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
225 230 235 240

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
245 250 255

X-13991.ST25.txt

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
260 265 270

<210> 30

<211> 272

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 30

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Ser Ser Gly Ala Ala Glu Pro Lys Ser Cys Asp Lys Thr
35 40 45

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
50 55 60

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
65 70 75 80

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
85 90 95

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
100 105 110

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
115 120 125

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
130 135 140

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
145 150 155 160

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
165 170 175

Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
180 185 190

55

X-13991.ST25.txt

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
195 200 205

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
210 215 220

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
225 230 235 240

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
245 250 255

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
260 265 270

<210> 31

<211> 287

<212> PRT

<213> Artificial Sequence

<220>

<223> synthetic construct

<400> 31

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
35 40 45

Ser Gly Gly Gly Gly Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr His
50 55 60

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
65 70 75 80

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
85 90 95

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu
100 105 110

Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
115 120 125

56

X-13991.ST25.txt

Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser
130 135 140

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
145 150 155 160

Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
165 170 175

Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
180 185 190

Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
195 200 205

Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
210 215 220

Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
225 230 235 240

Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
245 250 255

Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
260 265 270

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
275 280 285

<210> 32

<211> 232

<212> PRT

<213> Homo sapiens

<400> 32

Ala Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro
1 5 10 15

Ala Pro Glu Lys Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
20 25 30

Lys Asp Thr Lys Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
35 40 45

Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
50 55 60

57

X-13991.ST25.txt

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
65 70 75 80

Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln
85 90 95

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
100 105 110

Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
115 120 125

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
130 135 140

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
145 150 155 160

Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
165 170 175

Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
180 185 190

Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
195 200 205

Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
210 215 220

Ser Leu Ser Leu Ser Pro Gly Lys
225 230

<210> 33

<211> 703

<212> DNA

<213> Homo sapiens

<400> 33

gagcccaaat cttgtgacaa aactcacaca tgcccaccgt gcccagcacc tgaactcctg 60

gggggaccgt cagtcttctt cttcccccca aaaccaagg acacctcat gatctcccg 120

acccttgagg tcacatgcgt ggtggtggac gtgagccacg aagaccctga ggtcaagttc 180

aactggtacg tggacggcgt ggaggtgcat aatgccaaga caaagccgcg ggaggagcag 240

tacaacagca cgtaccgtgt ggtcagcgtc ctcaccgtcc tgcaccagga ctggctgaat 300

ggcaaggagt acaagtgcaa ggtctccaac aaagccctcc cagcccccat cgagaaaacc 360

X-13991.ST25.txt

atctccaaag ccaaagggca gccccgagaa ccacaggtgt acaccctgcc cccatcccgg 420
 gaggagatga ccaagaacca ggtcagcctg acctgcctgg tcaaaggctt ctatcccagc 480
 gacatcgccg tggagtggga gagcaatggg cagccggaga acaactacaa gaccacgcct 540
 cccgtgctgg actccgacgg ctcttcttct ctctatagca agctcaccgt ggacaagagc 600
 aggtggcagc aggggaacgt cttctcatgc tccgtgatgc atgaggctct gcacaaccac 660
 tacacgcaga agagcctctc cctgtctccg ggtaaatagat agt 703

<210> 34

<211> 585

<212> PRT

<213> Homo sapiens

<400> 34

Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
 1 5 10 15

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
 20 25 30

Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
 35 40 45

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
 50 55 60

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
 65 70 75 80

Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
 85 90 95

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
 100 105 110

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
 115 120 125

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
 130 135 140

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
 145 150 155 160

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
 165 170 175

X-13991.ST25.txt

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser
 180 185 190
 Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
 195 200 205
 Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro
 210 215 220
 Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys
 225 230 235 240
 Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp
 245 250 255
 Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser
 260 265 270
 Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His
 275 280 285
 Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser
 290 295 300
 Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala
 305 310 315 320
 Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg
 325 330 335
 Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr
 340 345 350
 Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu
 355 360 365
 Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro
 370 375 380
 Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Asn Leu Gly Glu
 385 390 395 400
 Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro
 405 410 415
 Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys
 420 425 430
 Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys
 435 440 445

X-13991.ST25.txt

207

Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His
450 455 460

Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser
465 470 475 480

Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr
485 490 495

Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp
500 505 510

Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
515 520 525

Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
530 535 540

Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
545 550 555 560

Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
565 570 575

Ala Ala Ser Gln Ala Ala Leu Gly Leu
580 585

<210> 35

<211> 1762

<212> DNA

<213> Homo sapiens

<400> 35
gatgcgcaca agagtgaggt tgctcatcgg tttaaagatt tgggagaaga aaatttcaaa 60
gccttggtgt tgattgcctt tgctcagtat cttcagcagt gtccatttga agatcatgta 120
aaattagtga atgaagtaac tgaatttgca aaaacatgtg ttgctgatga gtcagctgaa 180
aattgtgaca aatcacttca tacccttttt ggagacaaat tatgcacagt tgcaactctt 240
cgtgaaacct atggtgaaat ggctgactgc tgtgcaaaac aagaacctga gagaaatgaa 300
tgcttcttgc aacacaaaga tgacaaccca aacctcccc gattggtgag accagaggtt 360
gatgtgatgt gcactgcttt tcatgacaat gaagagacat ttttgaaaaa atacttatat 420
gaaattgcca gaagacatcc ttacttttat gccccggaac tccttttctt tgctaaaagg 480
tataaagctg cttttacaga atgttgccaa gctgctgata aagctgcctg cctgttgcca 540
aagctcgatg aacttcggga tgaaggggaag gcttcgtctg ccaaacagag actcaagtgt 600
1 51

X-13991.ST25.txt

208

gccagtctcc	aaaaatttgg	agaaagagct	ttcaaagcat	gggcagtagc	tcgcctgagc	660
cagagatttc	ccaaagctga	gtttgcagaa	gtttccaagt	tagtgacaga	tcttaccaa	720
gtccacacgg	aatgctgcc	tggagatctg	cttgaatgtg	ctgatgacag	ggcggacctt	780
gccaagtata	tctgtgaaa	tcaagattcg	atctccagta	aactgaagga	atgctgtgaa	840
aaacctctgt	tggaaaaatc	ccactgcatt	gccgaagtgg	aaaatgatga	gatgcctgct	900
gacttgcctt	cattagctgc	tgattttgtt	gaaagtaagg	atgtttgcaa	aaactatgct	960
gaggcaaagg	atgtcttcct	gggcatgttt	ttgtatgaat	atgcaagaag	gcacccctgat	1020
tactctgtcg	tgctgctgct	gagacttgcc	aagacatatg	aaaccactct	agagaagtgc	1080
tgtgccgctg	cagatcctca	tgaatgctat	gccaaagtgt	tcgatgaatt	taaacctctt	1140
gtggaagagc	ctcagaattt	aatcaacaa	attgtgagc	tttttgagca	gcttggagag	1200
tacaaattcc	agaatgcgct	attagttcgt	tacaccaaga	aagtacccca	agtgtcaact	1260
ccaactcttg	tagaggtctc	aagaaaccta	ggaaaagtgg	gcagcaaattg	ttgtaaacad	1320
cctgaagcaa	aaagaatgcc	ctgtgcagaa	gactatctat	ccgtgggtcct	gaaccagtta	1380
tgtgtgttgc	atgagaaaac	gccagtaagt	gacagagtca	ccaaatgctg	cacagaatcc	1440
ttggtgaaca	ggcgaccatg	cttttcagct	ctggaagtgc	atgaaacata	cgttcccaaa	1500
gagtttaatg	ctgaaacatt	caccttccat	gcagatatat	gcacactttc	tgagaaggag	1560
agacaaatca	agaaacaaac	tgacttggtt	gagctcgtga	aacacaagcc	caaggcaaca	1620
aaagagcaac	tgaaagctgt	tatggatgat	ttcgagctt	ttgtagagaa	gtgctgcaag	1680
gctgacgata	aggagacctg	ctttgccgag	gagggtaaaa	aacttggtgc	tgcaagtcaa	1740
gctgccttag	gcttataatg	ac				1762

(12) SOLICITUD INTERNACIONAL PUBLICADA EN VIRTUD DEL TRATADO DE COOPERACIÓN EN MATERIA DE PATENTES (PCT)

(19) Organización Mundial de la Propiedad Intelectual
Oficina Internacional

(43) Fecha de Publicación Internacional
13 de junio de 2002 (13.06.2002)

(Código de Barras)
WIPO
OMPI

(10) Número de Publicación Internacional
WO 02/48227 A2

PCT

(81) Clasificación internacional de Patente : C07K 14/805,
14/765, C12N 15/62, A81K 38/38, 38/28,
C07K 19/00, A81P 3/10.

(21) Número de Solicitud Internacional: PCT / US 01 / 42928

(22) Fecha de Presentación Internacional:
29 de noviembre de 2001 (29.11.2001)

(23) Idioma de presentación: Inglés

(28) Idioma de publicación: Inglés

(30) Datos de prioridad:

60/251,854 7 de diciembre de 2000 (7.12.2000) US

(71) Solicitante (designado para todos los países excepto Estados Unidos) ELI
LILLY AND COMPANY [E.E.U.U./E.E.U.U.]: Lilly Corporate Center, Indianapolis,
IN 46286 (E.E.U.U.)

(72) Inventores e;

(73) Inventores/solicitantes (sólo para Estados Unidos): GLAEBNER,
Wolfgang [DE/US]; 7512 Fieldstone Court, Indianapolis, IN 46254 (US).
MICANOVIC, Radmila [US/US]; 7128, White Oak Trail, Indianapolis, IN 46236
(US). TSCHANG, Sheng-Hung, Rainbow [US/US]; 4883 Riley Mews, Carmel, IN
46033 (US)

(74) Agentes : STEWART, Mark, J. et al.; Eli Lilly and Company, Lilly
Corporate Center, Drop Code 1104, Indianapolis, IN 46285 (US)

(81) Estados Designados (fase nacional) : AE, AG, AL, AM, AT, AT (modelo de
utilidad), AU, AZ, BA, BB, BF, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, CZ
(modelo de utilidad), DE, DK, DK, (modelo de utilidad), DM, DZ, EC, EE, EE,
(modelo de utilidad), ES, FI, FI (modelo de utilidad), GB, GD, GE, GH, GM, HR,
HU, ID, IL, IN, IS, JP, DE, DG, DP, DR, DZ, LZ, LK, LR, LS, LT, LU, LV, MA, MD,
MG, MKM MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK,
SK (modelo de utilidad), SL, TJ, TM, TR, TT, TZ, UA, UF, US, UZ, VN, YU, ZA,
ZW.

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

Declaraciones bajo la regla 4.17:

-Por lo que se refiere al poder del titular para solicitar y que le sea concedida
una patente (Regla 4.17 (I)) para las siguientes designaciones

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, DE, DG, DP, DR, DZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN,
MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR,
TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, Patente ARIPO (GH, GM, KE, LS, MW,
MZ, SD, SL, SZ, TZ, UG, ZM, ZW) Patente EUROASIÁTICA (AM, AZ, BY, KG,
KZ, MD, RU, TJ, TM) Patente EUROPEA (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR) Patente OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

- Por lo que se refiere al poder del titular para reivindicar la prioridad de una
solicitud anterior (Regla 4.17 (II)) para las siguientes designaciones: 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, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR,
TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, Patente ARIPO (GH, GM, KE, LS, MW,
MZ, SD, SL, SZ, TZ, UG, ZM, ZW) Patente EUROASIÁTICA (AM, AZ, BY, KG,
KZ, MD, RU, TJ, TM), Patente EUROPEA (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR). Patente OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

Así como la capacidad del solicitante para reivindicar la prioridad de
aplicaciones anteriores (Regla 4.17(III)) para todas las designaciones.

Publicada:

- Sin Informe de búsqueda internacional y para ser republicada a la recepción de
dicho informe.

Para los códigos de dos letras y otras abreviaturas, consultar con la Guía de
Códigos y Abreviaturas" que aparecen al comienzo de cada informe
periódico de la Gaceta de PCT.

(84) Título : FUSION DE PROTEINAS GLP-1

(57) Resumen : La presente invención se refiere a compuestos de tipo glucacón condensados con proteínas que presentan el efecto de prolongar a
semivida in vivo de los péptidos. Estas proteínas de fusión se pueden usar para tratar la diabetes mellitus no insulina-dependiente así como una
diversidad de otras afecciones.

PROTEÍNAS DE FUSIÓN DE GLP-1

La presente invención se refiere a péptidos de tipo glucagón que incluyen análogos y derivados de los mismos condensados con proteínas que tienen el efecto de ampliar *in vivo* la semivida de los péptidos. Estas proteínas de fusión se pueden usar
5 para tratar diabetes mellitus no insulino dependiente así como una diversidad de otros afecciones.

El péptido tipo glucagón 1 (abreviadamente GLP - 1) es un péptido de 37 aminoácidos que lo secretan las células L en el intestino en respuesta a la ingestión de
10 alimentos. Se ha encontrado que estimula la secreción de insulina (acción insulínica), que ocasiona por lo tanto la captación de glucosa por las células y disminuye los niveles de glucosa en suero [véase, por ejemplo, Mojsov, S., (1992, *Int. J. Peptide Protein Research*, 40: 333 - 343)]. Sin embargo, el GLP - 1 es escasamente activo. Una posterior escisión endógena entre la 6ª y 7ª posición produce un péptido
15 GLP 1 (7-37) OH más potente biológicamente activo. Se conocen numerosos análogos y derivados de GLP - 1 y se indican en esta memoria descriptiva como "compuestos GLP - 1". Estos análogos de GLP - 1 incluyen las exendinas que son péptidos encontrados en el veneno del monstruo - GILA. Las exendinas tienen homología de secuencias con el GLP - 1 nativo y pueden unirse al receptor de GLP - 1 e iniciar la
20 cascada de señal de transducción responsable de las numerosas actividades que se han atribuido a GLP - 1 (7 - 37) OH.

Los compuestos GLP - 1 tienen una diversidad de actividades fisiológicamente significativas. Por ejemplo, se ha mostrado que los GLP - 1 estimula la liberación de insulina, reducen la secreción de glucagón, inhiben el vaciado gástrico y potencian la
25 utilización de la glucosa. [Nauck, M. A. y col (1993) *Diabetologia* 36: 741 - 744;

Gutniak, M., y col (1992) *New England J. of Med.* 326: 1316 – 1322; Nauck, M. A. y col., (1993) *J. Clin. Invest.* 91: 301 - 307].

El GLP – 1 representa la mayor promesa como un tratamiento de la diabetes mellitus no insulino dependiente [abreviadamente NIDDM]. Hay numerosos fármacos orales en el mercado para tratar la resistencia a la insulina asociada con la NIDDM. Sin embargo, a medida que la enfermedad progresa los pacientes se deben cambiar a tratamientos que estimulen la liberación de insulina y eventualmente a tratamientos que impliquen inyecciones de insulina. Sin embargo, los fármacos actuales que estimulan la liberación de insulina pueden también ocasionar hipoglucemia como también puede ocasionar hipoglucemia la administración real de insulina. Sin embargo, la actividad de GLP – 1 se controla mediante los niveles de glucosa en sangre. Cuando los niveles caen a ciertos niveles umbral, el GLP – 1 no es activo. De este modo, no hay riesgo de hipoglucemia asociada con el tratamiento que implica GLP - 1.

Sin embargo, la utilidad de la terapia que implica péptidos GLP – 1 se ha limitado por sus rápidas eliminaciones y cortas semividas. Por ejemplo, GLP – (7 – 37) tiene una semivida en suero de 3 – 5 minutos solamente. La amida de GLP – 1 (7 – 36) tiene un tiempo de acción de aproximadamente 50 minutos cuando se administra subcutáneamente. Incluso los análogos y derivados que son resistentes a la escisión endógena con proteasa no tienen semividas suficientemente largas para evitar administraciones repetidas durante un período de 24 horas. Es inconveniente la eliminación rápida del agente terapéutico en casos en los que se desee mantener un alto nivel en sangre del agente durante un período prolongado de tiempo ya que entonces serán necesarias administraciones repetidas. Además, un compuesto de acción prolongada es particularmente importante para pacientes diabéticos cuyo régimen de tratamiento pasado ha implicado tomar sólo medicación oral. A menudo estos pacientes

tienen un tiempo extremadamente difícil de transición a un régimen que implique inyecciones múltiples de medicación.

La presente invención supera los problemas asociados con la administración de un compuesto que tiene una corta semivida en plasma. Los compuestos de la presente
5 invención comprenden compuestos de GLP - 1 condensados con otra proteína con una larga semivida de circulación tal como la porción Fc de una inmunoglobulina o albúmina.

En general, péptidos terapéuticos pequeños son difíciles de manipular porque incluso ligeros cambios en su estructura pueden afectar a la estabilidad y/o actividad
10 biológica. Esto ha sido especialmente cierto para los compuestos de GLP - 1 actualmente en desarrollo. Por ejemplo, GLP - 1 (7 - 37) OH tiene una tendencia a experimentar un cambio de conformación desde una estructura principalmente de hélice alfa a una estructura principalmente de hoja beta. Esta forma de hoja beta da como resultado un material agregado que se piensa que es inactivo. Por lo tanto, fue
15 sorprendente que se pudieran desarrollar proteínas de fusión de GLP - 1 biológicamente activas las semividas aumentadas. Fue especialmente inesperado dada la dificultad de trabajar con GLP - 1 (7 - 37) OH solo y el gran tamaño del copartícipe en la fusión pequeño péptido GLP - 1 adjunto.

20 Los compuestos de la presente invención incluyen proteínas de fusión heterólogas que comprenden un primer polipéptido con un extremo N terminal y un extremo C terminal condensados a un segundo polipéptido con extremo N terminal y un extremo C terminal en las que el primer polipéptido es un compuesto GLP - 1 y el segundo polipéptido se selecciona del grupo constituido por

25 a) albúmina humana;

- b) análogos de albúmina humana; y
- c) fragmentos de albúmina humana,

y en las que el extremo C terminal del primer polipéptido se condensa al extremo N terminal del segundo polipéptido.

5 Los compuestos de la presente también incluyen una proteína de fusión heteróloga que comprende un primer polipéptido con un extremo N terminal y extremo C terminal condensado con un segundo polipéptido con un extremo N terminal y extremo C terminal en la que el primer polipéptido es un compuesto GLP - 1 y el segundo polipéptido se selecciona de un grupo constituido por

- 10
- (a) albúmina humana;
 - (b) análogos de albúmina humana; y
 - (c) fragmentos de albúmina humana,

y en la que el extremo C terminal del primer polipéptido se condensa con el extremo N terminal del segundo polipéptido mediante un engarce de péptidos. Se prefiere que el

15 engarce de péptidos se seleccione del grupo constituido por:

- a) un péptido rico en glicina;
- b) un péptido que tiene la secuencia [Gly - Gly - Gly - Gly - Ser]_n en la que n es 1, 2, 3, 4, 5, ó 6; y
- c) un péptido que tiene la secuencia [Gly - Gly - Gly - Gly - Ser]₃.

20 Los compuestos adicionales de la presente invención incluyen una proteína de fusión heteróloga que comprende un primer polipéptido con un extremo N terminal y un extremo C terminal condensado con un segundo polipéptido con un extremo N terminal y un extremo C terminal en el que el primer polipéptido es un compuesto GLP - 1 y el segundo polipéptido se selecciona del grupo constituido por

- 25
- a) la porción Fc de una inmunoglobulina;

b) el análogo de la porción Fc de una inmunoglobulina; y

c) fragmentos de la porción Fc de una inmunoglobulina, y estando el extremo C terminal del primer polipéptido condensado con el extremo N terminal del segundo polipéptido. El compuesto GLP - 1 se puede condensar con el segundo polipéptido mediante un engarce de péptidos. Es preferible que el engarce de péptidos se seleccione entre un grupo constituido por:

a) un péptido rico en glicina;

b) un péptido que tiene la secuencia [Gly - Gly - Gly - Gly - Ser]_n en la que n es 1, 2, 3, 4, 5, ó 6; y

c) un péptido que tiene la secuencia [Gly - Gly - Gly - Gly - Ser]₃.

Generalmente se prefiere que el compuesto GLP - 1 que es parte de la proteína de fusión heteróloga no tenga más de 6 aminoácidos que sean diferentes del correspondiente aminoácido en GLP - 1 (7 - 37) OH, GLP - 1 (7 - 36) OH, o exendina - 4. Incluso se prefiere más que el compuesto GLP - 1 no tenga más de 5 aminoácidos diferentes del correspondiente aminoácido en GLP - 1 (7 - 37) OH, GLP - 1 (7 - 36) OH, o exendina - 4. Se prefiere más que el compuesto GLP - 1 no tenga más de 4, 3, ó 2 aminoácidos diferentes del correspondiente aminoácido en GLP - 1 (7 - 37) OH, GLP - 1 (7 - 36) OH, o exendina - 4. Preferiblemente, un compuesto de GLP - 1 que es parte de la proteína de fusión heteróloga tiene glicina o valina en la posición 8.

La presente invención también incluye polinucleótidos que codifican la proteína de fusión heteróloga descrita en esta memoria descriptiva, vectores que comprenden estos polinucleótidos y células huésped transfectadas o transformadas con los vectores descritos en esta memoria descriptiva. También se incluye un proceso para producir una proteína de fusión heteróloga que comprende las etapas de transcripción y traducción de

un polinucleótido descrito en esta memoria descriptiva en condiciones en las que la proteína de fusión heteróloga se expresa en cantidades detectables.

La presente invención también comprende un procedimiento para normalizar los niveles de glucosa en sangre en un mamífero que lo requiera que comprende la
5 administración de una cantidad terapéuticamente eficaz de una proteína de fusión heteróloga descrita en esta memoria descriptiva.

Además la invención también se ilustra con referencia a los siguientes dibujos:

Figura 1: secuencia de aminoácidos de IgG₁ Fc que comprende los dominios de
10 la región flexible, CH₂ y CH₃.

Figura 2: secuencia de aminoácidos de albúmina de suero humano.

Figura 3: A. Gel de SDS – PAGE e inmunotransferencia del mismo gel que ilustra el peso molecular de proteínas de fusión de IgG₁ - Fc y GLP – 1 – Fc (franja 1, patrones PM; franja 2, Fc purificada; franja 3, medio transfectado de Mock; franja 4, Val⁸ – GLP – 1 – Fc; franja 5, Exendina – 4-Fc) B. Gel de SDS - PAGE e
15 inmunotransferencia del mismo gel que ilustran el peso molecular de HSA humana y proteínas de fusión de GLP – 1 – HSA (franja 1, patrones PM; franja 2, HSA purificada; franja 3, medio transfectado de Mock; franja 4, Val⁸ – GLP – 1 – HSA; franja 5, val⁸ – GLP – 1 – [Gly – Gly – Gly – Gly – Ser]₃ – HSA; Franja 6, exendina - 4 – HSA; Franja
20 7, exendina – 4 – [Gly – Gly – Gly – Gly – Ser]₃ – HSA).

Figura 4: gel de SDS – PAGE de Fc purificado, albúmina y proteínas de fusión de GLP – 1 (franja 1, patrones de PM; franja 2, Fc purificado; franja 3, Val⁸ – GLP – 1 – Fc; franja 4, exendina – 4 – Fc; franja 5, patrón PM; franja 6, Val⁸ – GLP – 1 – HSA; franja 7, exendina – 4 – HSA; franja 8, exendina – 4 - [Gly – Gly – Gly – Gly – Ser]₃ –
25 HSA).

Figura 5: vector de expresión de clonaje que contiene las regiones Fc ilustradas en la figura 1.

Figura 6: vector de expresión de clonaje que contiene la secuencia de albúmina ilustrada en la figura 2.

- 5 Figura 7: vector de expresión de clonaje que contiene DNA que codifica un engarce de 15 aminoácidos condensado en fase y la posición 5' de la secuencia de albúmina ilustrada en la figura 2.

Figura 8: actividad de respuesta a la dosis *in vitro* de las proteínas de fusión de GLP - 1.

- 10 Figura 9: farmacocinética de las proteínas de fusión GLP-1-Fc y GLP-1-HSA.

Figura 10: respuesta glucodinámica a exendina - Fc en dos perros en ayunas normales.

Figura 11: respuesta insulínica a exendina - Fc en dos perros en ayunas normales.

- 15 Figura 12. secuencia de DNA que codifica una región de IgG1 Fc humana.

Figura 13: secuencia de DNA que codifica una proteína de albúmina humana.

- Las proteínas de fusión heterólogas de la presente invención comprenden un compuesto GLP - 1 condensado a albúmina humana, un análogo de albúmina humana, un fragmento de albúmina humana, la porción Fc de una inmunoglobulina, un análogo de la porción Fc de una inmunoglobulina, o un fragmento de la porción Fc de una inmunoglobulina. El extremo C terminal del compuesto GLP - 1 se puede condensar directamente, o se puede condensar por medio de un engarce de péptidos, con el extremo N terminal de una proteína albúmina Fc. Estas proteínas de fusión heterólogas son biológicamente activas y tienen una semivida aumentada comparada
- 20
- 25

con el GPL - 1 nativo.

Se prefiere que los compuestos GLP - 1 que forman parte de la proteína de fusión heteróloga comprendan polipéptidos que tienen entre aproximadamente veinticinco y aproximadamente treinta y nueve aminoácidos de origen natural o de origen no natural que tienen suficiente homología con el GLP - 1 (7 - 37) OH nativo de modo que muestran actividad insulínica mediante la unión al receptor de GLP - 1 en las células β del páncreas. Típicamente un compuesto GLP - 1 comprende un polipéptido que tiene la secuencia de aminoácidos de GLP - 1 (7 - 37) OH, un análogo de GLP - 1 (7 - 37) OH, un fragmento de GLP - 1 (7 - 37) OH o un fragmento de un análogo de GLP - 1 (7 - 37) OH. El GLP - 1 (7 - 37) OH tiene la secuencia de aminoácidos de la SEC ID N° 1:

	7	8	9	10	11	12	13	14	15	16	17
	His	Ala	Glu	Gly	Thr	Phe	Thr	Ser	Asp	Val	Ser
	18	19	20	21	22	23	24	25	26	27	28
15	Ser	Tyr	Leu	Glu	Gly	Gln	Ala	Ala	Lys	Glu	Phe
	29	30	31	32	33	34	35	36	37		
	Ile	Ala	Trp	Leu	Val	Lys	Gly	Arg	Gly		
	(SEC ID N° 1)										

Es costumbre en la técnica que al extremo amino terminal de GLP - 1 (7 - 37) OH se le asigne el residuo número 7 y al extremo carboxilo terminal el número 37. Los otros aminoácidos en el polipéptido se numeran consecutivamente como se muestra en la SEC ID N° 1. Por ejemplo, la posición 12 es fenilalanina y la posición 22 es glicina.

Los compuestos GLP - 1 también comprenden "fragmentos de GLP - 1". Un fragmento de GLP - 1 es un polipéptido obtenido después del truncamiento de uno o

22
218

más aminoácidos del extremo N terminal y/o extremo C terminal de GLP - 1 (7 - 37) OH o un análogo o un derivado del mismo. La nomenclatura utilizada para describir GLP - 1 (7 - 37) OH es también aplicable a los fragmentos de GLP - 1. Por ejemplo, GLP - 1 (9 - 36) OH designa un fragmento obtenido mediante el truncamiento de dos aminoácidos del extremo N terminal y un aminoácido del extremo C terminal. Los aminoácidos en el fragmento se designan mediante el mismo número que el correspondiente aminoácido en GLP - 1 (7 - 37) OH. Por ejemplo, el ácido glutámico del extremo N terminal en GLP - 1 (9 - 36) OH está en la posición 9; la posición 12 está ocupada por fenilalanina; y la posición 22 está ocupada por glicina, como en GLP - 1 (7 - 37) OH. Para GLP - 1 (7 - 37) OH, la glicina de la posición 37 de GLP - 1 (7 - 37) OH está suprimida.

Los compuestos GLP - 1 también incluyen polipeptidos en los que se han añadido uno o más aminoácidos al extremo N terminal y/o extremo C terminal de GLP - 1 (7 - 37) OH, o fragmentos o análogos de los mismos. Es preferible que los compuestos GLP - 1 de este tipo tengan hasta aproximadamente treinta y nueve aminoácidos. Los aminoácidos en el compuesto GLP - 1 "extendido" se designan mediante el mismo número que el correspondiente aminoácido en GLP - 1 (7 - 37) OH. Por ejemplo, el aminoácido del extremo N terminal de compuesto GLP - 1 obtenido mediante la adición de dos aminoácidos al terminal N de GLP - 1 (7 - 37) OH está en la posición 5; y el aminoácido del extremo C terminal del compuesto GLP - 1 obtenido mediante la adición de un aminoácido al extremo C terminal de GLP - 1 (7 - 37) OH está en la posición 38. De este modo, la posición 12 está ocupada por fenilalanina y la posición 22 está ocupada por glicina en ambos de estos compuestos de GLP - 1 "extendidos" como en GLP - 1 (7 - 37) OH. Los aminoácidos 1 - 6 de un compuesto de GLP - 1 extendido son preferiblemente los mismos o una sustitución conservadora del

227
219

mismo aminoácido en la posición correspondiente de GLP - 1 (1 - 37) OH. Los aminoácidos 38 - 45 de un compuesto de GLP - 1 extendido son preferiblemente los mismos o una sustitución conservadora del aminoácido en la posición correspondiente de glucagón o extendina - 4.

- 5 Los compuestos de GLP - 1 de la presente invención comprenden "análogos de GLP - 1". Un análogo de GLP - 1 tiene una homología suficiente con GLP - 1 (7 - 37) OH o un fragmento de GLP - 1 (7 - 37) OH de forma que el análogo tiene actividad insulínica. Preferiblemente, un análogo de GLP - 1 tiene la secuencia de aminoácidos de GLP - 1 (7 - 37) OH o un fragmento de la misma, modificada de
- 10 manera que de uno, dos, tres, cuatro, o cinco aminoácidos difieren del aminoácido en la posición correspondiente de GLP - 1 (7 - 37) OH o un fragmento de GLP - 1 (7 - 37) OH. En la nomenclatura usada en esta memoria descriptiva para designar los compuestos GLP - 1, el aminoácido sustituyente y su posición se indican antes que la estructura de partida. Por ejemplo, Glu²² - GLP - 1 (7 - 37) OH designa un compuesto
- 15 de GLP - 1 en el que la glicina que normalmente se encuentra en la posición 22 de GLP - 1 (7 - 37) OH se ha sustituido con ácido glutámico; Val⁸ - Glu²² - GLP - 1 (7 - 37) OH designa un compuesto de GLP - 1 en el que la alanina que normalmente se encuentra en la posición 8 y la glicina que normalmente se encuentra en la posición 22 de GLP - 1 (7 - 37) OH se han sustituido con valina y ácido glutámico respectivamente.
- 20 Los compuestos GLP - 1 de la presente invención también incluyen "derivados de GLP - 1". Un derivado de GLP - 1 se define como una molécula que tiene la secuencia de aminoácidos de GLP - 1 o de un análogo de GLP - 1, pero que tiene adicionalmente una modificación química de uno o más de sus grupos laterales de aminoácidos, átomos de carbono α , grupo amino terminal o grupo ácido carboxílico
- 25 terminal. Una modificación química incluye, pero no está limitada a, la adición de

restos químicos, creación de nuevos enlaces y eliminación de restos químicos. Las modificaciones en los grupos laterales de aminoácidos incluyen, sin limitación, acilación de los grupos amino ϵ de lisina, N-alquilación de arginina, histidina, o lisina, alquilación de grupos ácidos carboxílicos glutámico o aspártico y desamidación de glutamina o asparagina. Las modificaciones del grupo amino terminal incluyen, sin limitación, las modificaciones de desamino, N-alquilo inferior, N-di-alquilo inferior, y N-acilo. Las modificaciones del grupo carboxi terminal incluyen, sin limitación, la amida, amida de alquilo inferior, dialquilamida y modificaciones de ester de alquilo inferior. Alquilo inferior es alquilo ($C_1 - C_4$). Además, uno o más grupos laterales, o grupos terminales, se pueden proteger mediante grupos protectores conocidos por los químicos de proteínas expertos habituales. El carbono α de un aminoácido puede ser mono- o dimetilado.

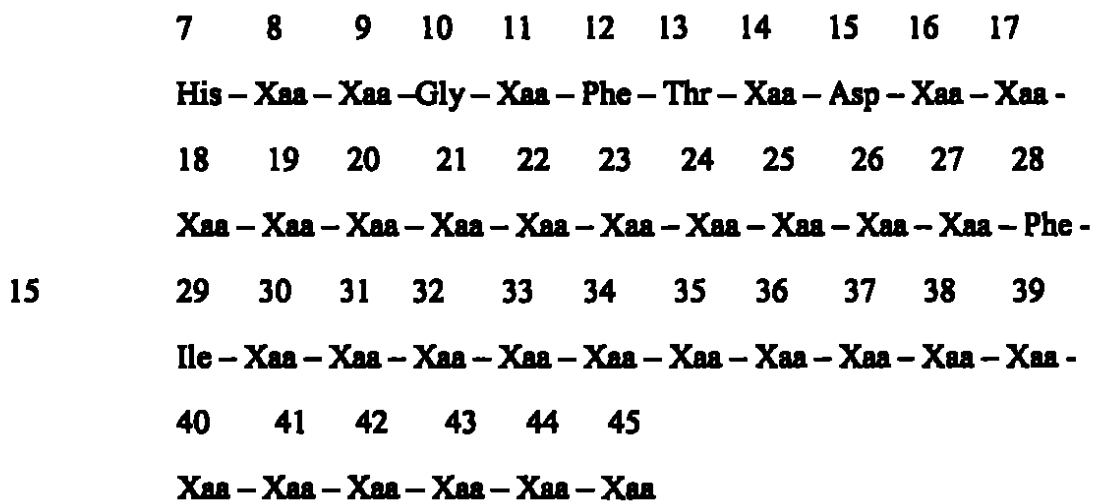
Cualquier compuesto GLP - 1 puede ser parte de las proteínas de fusión heterólogas de la presente invención mientras que el compuesto GLP - 1 sí mismo sea capaz de unirse e inducir señalización mediante el receptor de GLP - 1. La unión al receptor de GLP - 1 y la transducción de la señal se puede valorar usando ensayos *in vitro* tales como los descritos en el documento EP 619.322 y en la patente de Estados Unidos 5.120.712 respectivamente.

Se conocen en la técnica numerosos fragmentos análogos y derivados de GLP - 1 activos y cualquiera de estos análogos y derivados también pueden ser parte de las proteínas de fusión heterólogas de la presente invención. Se proporcionan en esta memoria descriptiva algunos ejemplos de análogos de GLP - 1 nuevos así como análogos y derivados de GLP - 1 conocidos en la técnica.

Algunos análogos de GLP - 1 y fragmentos de GLP - 1 conocidos en la técnica incluyen, por ejemplo, GLP - 1(7 - 34) y GLP - 1(7 - 35), GLP - 1(7 - 36), Gln⁹ -

GLP - 1(7 - 37), D - Gln⁹ - GLP - 1(7 - 37), Thr¹⁶ - Lys¹⁸ - GLP - 1 (7 - 37) y Lys¹⁸ - GLP - 1 (7 - 37). Los análogos de GLP - 1 tales como GLP - 1 (7 - 34) y GLP - 1 (7 - 35) se describen en la patente de Estados Unidos 5.118.666. Se conocen también las formas biológicamente elaboradas de GLP - 1 que tienen propiedades insulínótropicas, tales como GLP - 1 (7 - 36). Otros compuestos conocidos de GLP - 1 biológicamente activos se describen en la patente de Estados Unidos N° 5.997.071 de Hoffmann y col., la patente de Estados Unidos N° 5.545.618 de Buckley y col., y Adelhorst, y col., *J. Biol. Chem.* 269: 6275 (1994).

Un grupo preferido de análogos de GLP - 1 está compuesto de análogos de GLP - 1 de fórmula I (SEC ID N° 2)



Fórmula (SEC ID N° 2)

20 en la que:

Xaa en la posición 8 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

Xaa en la posición 9 es Glu, Asp, o Lys;

Xaa en la posición 11 es Thr, Ala, Gly, Ser, Leu, Ile, Val, Glu, Asp o Lys;

Xaa en la posición 14 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp o Lys;

25 Xaa en la posición 16 es Val, Ala, Gly, Ser, Thr, Leu, Ile, Tyr, Glu, Asp, Trp o

Lys;

Xaa en la posición 17 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp o Lys;

Xaa en la posición 18 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, Trp, Tyr o Lys;

5 Xaa en la posición 19 es Tyr, Phe, Trp, Glu, Asp, Gln, o Lys;

Xaa en la posición 20 es Leu, Ala, Gly, Ser, Thr, Ile, Val, Glu, Asp, Met, Trp, Tyr o Lys;

Xaa en la posición 21 es Glu, Asp, o Lys;

Xaa en la posición 22 es Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, o Lys

10 Xaa en la posición 23 es Gln, Asn, Arg, Glu, Asp, o Lys;

Xaa en la posición 24 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Arg, Glu, Asp o Lys;

Xaa en la posición 25 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

Xaa en la posición 26 es Lys, Arg, Gln, Glu, Asp, o His;

Xaa en la posición 27 es Leu, Glu, Asp o Lys;

15 Xaa en la posición 30 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

Xaa en la posición 31 es Trp, Phe, Tyr, Glu, Asp o Lys;

Xaa en la posición 32 es Leu, Gly, Ala, Ser, Thr, Ile, Val, Glu, Asp o Lys;

Xaa en la posición 33 es Val, Gly, Ala, Ser, Thr, Leu, Ile, Glu, Asp o Lys;

Xaa en la posición 34 es Asn, Lys, Arg, Glu, Asp o His;

20 Xaa en la posición 35 es Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

Xaa en la posición 36 es Gly, Arg, Lys, Glu, Asp o His;

Xaa en la posición 37 es Pro, Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys o

se ha suprimido;

Xaa en la posición 38 es Ser, Arg, Lys, Glu, Asp o His o se ha suprimido;

25 Xaa en la posición 39 es Ser, Arg, Lys, Glu, Asp o His, o se ha suprimido;

Xaa en la posición 40 es Gly, Asp, Glu o Lys o se ha suprimido;

Xaa en la posición 41 es Ala, Phe, Trp, Tyr, Glu, Asp o Lys, o se ha suprimido;

Xaa en la posición 42 es Ser, Pro, Lys, Glu o Asp, o se ha suprimido;

Xaa en la posición 43 es Ser, Pro, Glu, Asp o Lys, o se ha suprimido;

5 Xaa en la posición 44 es Gly, Pro, Glu, Asp o Lys, o se ha suprimido;

y

Xaa en la posición 45 es Ala, Ser, Val, Glu, Asp o Lys, o se ha suprimido;

con la condición de que cuando el aminoácido en la posición 37, 38, 39, 40, 41,
42, 43 ó 44 se ha suprimido, entonces cada aminoácido cadema abajo de ese
10 aminoácido está también suprimido.

Se prefiere que el compuesto GLP - 1 de fórmula I contenga menos de seis
aminoácidos diferentes del aminoácido correspondiente en GLP - 1 (7 - 37) OH o
exendina - 4. Se prefiere más que menos de cinco aminoácidos sean diferentes del
correspondiente aminoácido en GLP - 1 (7 - 37) OH o exendina - 4. Se prefiere incluso
15 más que menos de cuatro aminoácidos sean diferentes del correspondiente aminoácido
en GLP - 1 (7 - 37) OH o exendina - 4.

Los compuestos GLP - 1 de la presente invención incluyen derivados de
fórmula I tales como ésteres de C₁₋₆, o amida, o alquil C₁₋₆ amida, o dialquil C₁₋₆
amida de los mismos. El documento WO 99/43706 describe derivados de compuestos
20 GLP - 1 de fórmula I y se incorpora en esta memoria descriptiva como referencia en su
totalidad. Los compuestos de fórmula I derivatizados como se describe en el documento
WO 99/43706 y no derivatizados se incluyen en la presente invención.

Otro grupo preferido de los compuestos GLP - 1 está compuesto de análogos de
25 GLP - 1 de fórmula II (SEC ID N° 3):

7 8 9 10 11 12 13 14 15 16 17
 Xaa – Xaa – Xaa – Gly – Xaa – Xaa – Thr – Ser – Asp – Xaa – Ser –
 18 19 20 21 22 23 24 25 26 27 28
 5 Xaa – Xaa – Leu – Glu – Gly – Xaa – Xaa – Ala – Xaa – Xaa – Phe –
 29 30 31 32 33 34 35 36 37
 Ile – Xaa – Xaa – Leu – Xaa – Xaa – Xaa – Xaa – R

Fórmula II (SEC ID Nº 3)

en la que:

- 10 Xaa en la posición 7 es: L - histidina, D – histidina, desaminohistidina, 2 - aminohistidina, β – hidroxihistidina, homohistidina, α – fluorometilhistidina o α – metilhistidina;
- Xaa en la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser o Thr;
 Xaa en la posición 9 es: Thr, Ser, Arg, Lys, Trp, Phe, Tyr, Glu o His;
- 15 Xaa en la posición 11 es: Asp, Glu, Arg, Thr, Ala, Lys o His;
 Xaa en la posición 12 es: His, Trp, Phe o Tyr;
 Xaa en la posición 16 es: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Tyr, Glu o Ala;
 Xaa en la posición 18 es: His, Pro, Asp, Glu, Arg, Ser, Ala o Lys;
 Xaa en la posición 29 es: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;
- 20 Xaa en la posición 23 es: His, Asp, Lys, Glu, Gln o Arg;
 Xaa en la posición 24 es: Glu, Arg, Ala o Lys;
 Xaa en la posición 26 es: Trp, Tyr, Phe, Asp, Lys, Glu o His;
 Xaa en la posición 27 es: Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;
 Xaa en la posición 30 es: Ala, Glu, Asp, Ser o His;
- 25 Xaa en la posición 31 es: Asp, Glu, Ser, Thr, Arg, Trp o Lys;

225

Xaa en la posición 33 es: Asp, Arg, Val, Lys, Ala, Gly o Glu;

Xaa en la posición 34 es: Glu, Lys o Asp;

Xaa en la posición 35 es: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o Glu;

5 Xaa en la posición 36 es: Thr, Ser, Asp, Trp, Tyr, Phe, Arg, Glu o His;

R en la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly – Pro o se ha suprimido.

Otro grupo preferido de los compuestos de GLP – 1 está compuesto de análogos
10 de GLP – 1 de fórmula III (SEC ID N° 4):

	7	8	9	10	11	12	13	14	15	16	17
	Xaa	Xaa	Glu	Gly	Xaa	Xaa	Thr	Ser	Asp	Xaa	Ser
	18	19	20	21	22	23	24	25	26	27	28
15	Ser	Tyr	Leu	Glu	Xaa	Xaa	Xaa	Xaa	Lys	Xaa	Phe
	29	30	31	32	33	34	35	36	37		
	Ile	Xaa	Trp	Leu	Xaa	Xaa	Xaa	Xaa	R		

Fórmula III (SEC ID N° 4)

en la que:

20 Xaa en la posición 7 es: L - histidina, D - histidina, desaminohistidina, 2 - aminohistidina, β - hidroxihistidina, homohistidina, α - fluorometilhistidina o α - metilhistidina;

Xaa en la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser o Thr;

Xaa en la posición 11 es: Asp, Glu, Arg, Thr, Ala, Lys o His;

25 Xaa en la posición 12 es: His, Trp, Phe o Tyr;

Xaa en la posición 16 es: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu o Ala;

Xaa en la posición 22 es: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;

Xaa en la posición 23 es: His, Asp, Lys, Glu o Gln;

Xaa en la posición 24 es: Glu, His, Ala o Lys;

5 Xaa en la posición 25 es: Asp, Lys; Glu o His;

Xaa en la posición 27 es: Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;

Xaa en la posición 30 es: Ala, Glu, Asp, Ser o His;

Xaa en la posición 33 es: Asp, Arg, Val, Lys, Ala, Gly o Glu;

Xaa en la posición 34 es: Glu, Lys o Asp;

10 Xaa en la posición 35 es: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o
Glu;

Xaa en la posición 36 es: Asp, Glu o His;

R en la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly,
Gly – Pro o se ha suprimido.

15

Otro grupo preferido de los compuestos de GLP – 1 está compuesto de análogos
de GLP – 1 de fórmula IV (SEC ID N° 5):

	7	8	9	10	11	12	13	14	15	16	17
20	Xaa	Xaa	Glu	Gly	Thr	Xaa	Thr	Ser	Asp	Xaa	Ser
	18	19	20	21	22	23	24	25	26	27	28
	Ser	Tyr	Leu	Glu	Xaa	Xaa	Ala	Ala	Xaa	Glu	Phe
	29	30	31	32	33	34	35	36	37		
25	Ile	Xaa	Trp	Leu	Val	Lys	Xaa	Arg	R		

Fórmula IV (SEC ID N° 5)

en la que:

Xaa en la posición 7 es: L - histidina, D - histidina, desaminohistidina, 2 - aminohistidina, β - hidroxihistidina, homohistidina, α - fluorometilhistidina o α - metilhistidina;

Xaa en la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser, Met o Thr;

Xaa en la posición 12 es: His, Trp, Phe o Tyr;

Xaa en la posición 16 es: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu o Ala;

Xaa en la posición 22 es: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;

10 Xaa en la posición 23 es: His, Asp, Lys, Glu o Gln;

Xaa en la posición 26 es: Asp, Lys, Glu o His;

Xaa en la posición 30 es: Ala, Glu, Asp, Ser o His;

Xaa en la posición 35 es: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o Glu;

15 R en la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly - Pro o se ha suprimido.

Otro grupo preferido de los compuestos de GLP - 1 está compuesto de análogos de GLP - 1 de fórmula V (SEC ID N° 6):

20	7	8	9	10	11	12	13	14	15	16	17
	Xaa	- Xaa	- Glu	- Gly	- Thr	- Phe	- Thr	- Ser	- Asp	- Val	- Ser
	18	19	20	21	22	23	24	25	26	27	28
	Ser	- Tyr	- Leu	- Glu	- Xaa	- Xaa	- Xaa	- Ala	- Lys	- Glu	- Phe
	29	30	31	32	33	34	35	36	37		
25	Ile	- Xaa	- Trp	- Leu	- Val	- Lys	- Gly	- Arg	- R		

Fórmula V (SEC ID N° 6)

en la que:

Xaa en la posición 7 es: L - histidina, D - histidina, desaminohistidina, 2 - aminohistidina, β - hidroxihistidina, homohistidina, α - fluorometilhistidina o α - metilhistidina;

Xaa en la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser o Thr;

Xaa en la posición 22 es: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;

Xaa en la posición 23 es: His, Asp, Lys, Glu o Gln;

Xaa en la posición 24 es: Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;

10 Xaa en la posición 30 es: Ala, Glu, Asp, Ser o His;

R en la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly - Pro o se ha suprimido.

Los compuestos preferidos GLP - 1 de fórmulas I, II, III, IV y V comprenden

15 análogos de GLP - 1 o fragmentos de análogos de GLP - 1 en los que los análogos o fragmentos contienen un aminoácido distinto de la alanina en la posición 8 (análogos de posición 8). Es preferible que estos análogos de posición 8 contengan uno o más cambios adicionales en las posiciones 9, 11, 12, 16, 18, 22, 23, 24, 26, 27, 30, 31, 33, 34, 35, 36 y 37 comparados con el correspondiente aminoácido de GLP - 1 (7 - 37) OH

20 nativo. También es preferible que estos análogos tengan 6 o menos cambios comparados con los correspondientes aminoácidos en GLP - 1 (7 - 37) OH o GLP - 1 (7 - 36) OH nativos. Análogos más preferidos tienen 5 ó menos cambios comparados con los correspondientes aminoácidos en GLP - 1 (7 - 37) OH o GLP - 1 (7 - 36) OH nativos o tienen 4 ó menos cambios comparados con los correspondientes aminoácidos

25 en GLP - 1 (7 - 37) OH o GLP - 1 (7 - 36) OH nativos. Incluso se prefiere aún más

que estos análogos tengan 3 o menos cambios comparados con los correspondientes aminoácidos en GLP - 1 (7 - 37) OH o GLP - 1 (7 - 36) OH nativos. Es lo más preferible que estos análogos tengan 2 ó menos cambios comparados con los correspondientes aminoácidos en GLP - 1 (7 - 37) OH nativo.

- 5 Se ha encontrado que los compuestos de fórmula II, III, IV y V tienen una propensión reducida a agregarse y generar formas insolubles. Esto también es importante en el contexto de una proteína de fusión en la que el péptido relativamente pequeño GLP - 1 debe mantener una conformación activa a pesar de estar condensado con una proteína mucho más grande. Los compuestos preferidos de GLP - 1 de fórmula
- 10 II, III, IV y V comprendidos por las proteínas de fusión de la presente invención comprenden análogos de GLP - 1 o fragmentos de análogos de GLP - 1 en los que la glicina en la posición 22 y preferiblemente la alanina en la posición 8 se han sustituido con otro aminoácido.

- Cuando la posición 22 es ácido aspártico, ácido glutámico, arginina o lisina, la
- 15 posición 8 es preferiblemente glicina, valina, leucina, isoleucina, serina, treonina o metionina y más preferiblemente valina o glicina. Cuando la posición 22 es un ácido sulfónico tal como ácido cisteico, la posición 8 es preferiblemente glicina, valina, leucina, isoleucina, serina, treonina o metionina y más preferiblemente valina o glicina.

- Otros compuestos preferidos de GLP - 1 incluyen análogos de GLP - 1 de
- 20 fórmula IV (SEC ID N° 5) en la que los análogos tienen la secuencia de GLP - 1 (7 - 37) OH excepto que el aminoácido en la posición 8 es preferiblemente glicina, valina, leucina, isoleucina, serina, treonina o metionina y más preferiblemente valina o glicina y la posición 30 es ácido glutámico, ácido aspártico, serina o histidina y más preferiblemente ácido glutámico.

- 25 Otros compuestos preferidos GLP - 1 incluyen análogos de GLP - 1 de fórmula

IV (SEC ID N° 5) en la que los análogos tienen la secuencia de GLP – 1 (7 – 37) OH excepto que el aminoácido en la posición 8 es preferiblemente glicina, valina, leucina, isoleucina, serina, treonina o metionina y más preferiblemente valina o glicina y el aminoácido en la posición 37 es histidina, lisina, arginina, treonina, serina, ácido glutámico, ácido aspártico, triptófano, tirosina, fenilalanina y más preferiblemente histidina.

Otros compuestos preferidos GLP – 1 incluyen análogos de GLP – 1 de fórmula IV (SEC ID N° 5) en la que los análogos tienen la secuencia de GLP – 1 (7 – 37) OH excepto que el aminoácido en la posición 8 es preferiblemente glicina, valina, leucina, isoleucina, serina, treonina o metionina y más preferiblemente valina o glicina y el aminoácido en la posición 22 es ácido glutámico, lisina, ácido aspártico o arginina y más preferiblemente ácido glutámico o lisina y el aminoácido en la posición 23 es lisina, arginina, ácido glutámico, ácido aspártico e histidina y más preferiblemente lisina o ácido glutámico.

Otros compuestos preferidos de GLP – 1 incluyen análogos de GLP – 1 de fórmula V (SEC ID N° 6) en la que los análogos tienen la secuencia de GLP – 1 (7 – 37) OH excepto que el aminoácido en la posición 8 es preferiblemente glicina, valina, leucina, isoleucina, serina, treonina o metionina y más preferiblemente valina o glicina y el aminoácido en la posición 22 es ácido glutámico, lisina, ácido aspártico o arginina y más preferiblemente ácido glutámico o lisina y la posición 27 es alanina, lisina, arginina, triptófano, tirosina, fenilalanina o histidina y más preferiblemente alanina.

Otros compuestos preferidos GLP – 1 incluyen análogos de GLP – 1 de fórmula II en los que la análogos tienen la secuencia de GLP – 1(7 – 37) OH excepto que el aminoácido en la posición 8 es uno, dos o tres aminoácidos seleccionados del grupo constituido por la posición 9, posición 11, posición 12, posición 16, posición 18,

posición 22, posición 23, posición 24, posición 26, posición 27, posición 30, posición 31, posición 33, posición 34, posición 35, posición 36 y posición 37, difieren del aminoácido en la posición correspondiente de GLP-1 (7-37) OH nativo.

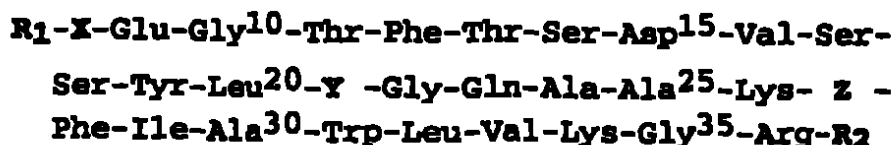
Otros compuestos GLP-1 preferidos de fórmula II incluyen:

5

Val⁸-GLP-1(7-37)OH, Gly⁸-GLP-1(7-37)OH, Glu²²-GLP-1(7-37)OH, Asp²²-GLP-1(7-37)OH, Arg²²-GLP-1(7-37)OH, Lys²²-GLP-1(7-37)OH, Cys²²-GLP-1(7-37)OH, Val⁸-Glu²²-GLP-1(7-37)OH, Val⁸-Asp²²-GLP-1(7-37)OH, Val⁸-Arg²²-GLP-1(7-37)OH, Val⁸-Lys²²-GLP-1(7-37)OH, Val⁸-Cys²²-GLP-1(7-37)OH, Gly⁸-Glu²²-GLP-1(7-37)OH, Gly⁸-Asp²²-GLP-1(7-37)OH, Gly⁸-Arg²²-GLP-1(7-37)OH, Gly⁸-Lys²²-GLP-1(7-37)OH, Gly⁸-Cys²²-GLP-1(7-37)OH, Glu²²-GLP-1(7-36)OH, Asp²²-GLP-1(7-36)OH, Arg²²-GLP-1(7-36)OH, Lys²²-GLP-1(7-36)OH, Cys²²-GLP-1(7-36)OH, Val⁸-Glu²²-GLP-1(7-36)OH, Val⁸-Asp²²-GLP-1(7-36)OH, Val⁸-Arg²²-GLP-1(7-36)OH, Val⁸-Lys²²-GLP-1(7-36)OH, Val⁸-Cys²²-GLP-1(7-36)OH, Gly⁸-Glu²²-GLP-1(7-36)OH, Gly⁸-Asp²²-GLP-1(7-36)OH, Gly⁸-Arg²²-GLP-1(7-36)OH, Gly⁸-Lys²²-GLP-1(7-36)OH, Gly⁸-Cys²²-GLP-1(7-36)OH, Lys²³-GLP-1(7-37)OH, Val⁸-Lys²³-GLP-1(7-37)OH, Gly⁸-Lys²³-GLP-1(7-37)OH, His²⁴-GLP-1(7-37)OH, Val⁸-His²⁴-GLP-1(7-37)OH, Gly⁸-His²⁴-GLP-1(7-37)OH, Lys²⁴-GLP-1(7-37)OH, Val⁸-Lys²⁴-GLP-1(7-37)OH, Gly⁸-Lys²³-GLP-1(7-37)OH, Glu³⁰-GLP-1(7-37)OH, Val⁸-Glu³⁰-GLP-1(7-37)OH, Gly⁸-Glu³⁰-GLP-1(7-37)OH, Asp³⁰-GLP-1(7-37)OH, Val⁸-Asp³⁰-GLP-1(7-37)OH, Gly⁸-Asp³⁰-GLP-1(7-37)OH, Gln³⁰-GLP-1(7-37)OH, Val⁸-Gln³⁰-GLP-1(7-37)OH, Gly⁸-Gln³⁰-GLP-1(7-37)OH, Tyr³⁰-

GLP-1(7-37)OH, Val⁸-Tyr³⁰-GLP-1(7-37)OH, Gly⁸-Tyr³⁰-GLP-1(7-37)OH, Ser³⁰-GLP-1(7-37)OH, Val⁸-Ser³⁰-GLP-1(7-37)OH, Gly⁸-Ser³⁰-GLP-1(7-37)OH, His³⁰-GLP-1(7-37)OH, Val⁸-His³⁰-GLP-1(7-37)OH, Gly⁸-His³⁰-GLP-1(7-37)OH, Glu³⁴-GLP-1(7-37)OH, Val⁸-Glu³⁴-GLP-1(7-37)OH, Gly⁸-Glu³⁴-GLP-1(7-37)OH, Ala³⁴-GLP-1(7-37)OH, Val⁸-Ala³⁴-GLP-1(7-37)OH, Gly⁸-Ala³⁴-GLP-1(7-37)OH, Gly³⁴-GLP-1(7-37)OH, Val⁸-Gly³⁴-GLP-1(7-37)OH, Gly⁸-Gly³⁴-GLP-1(7-37)OH, Ala³⁵-GLP-1(7-37)OH, Val⁸-Ala³⁵-GLP-1(7-37)OH, Gly⁸-Ala³⁵-GLP-1(7-37)OH, Lys³⁵-GLP-1(7-37)OH, Val⁸-Lys³⁵-GLP-1(7-37)OH, Gly⁸-Lys³⁵-GLP-1(7-37)OH, His³⁵-GLP-1(7-37)OH, Val⁸-His³⁵-GLP-1(7-37)OH, Gly⁸-His³⁵-GLP-1(7-37)OH, Pro³⁵-GLP-1(7-37)OH, Val⁸-Pro³⁵-GLP-1(7-37)OH, Gly⁸-Pro³⁵-GLP-1(7-37)OH, Glu³⁵-GLP-1(7-37)OH, Val⁸-Glu³⁵-GLP-1(7-37)OH, Gly⁸-Glu³⁵-GLP-1(7-37)OH, Val⁸-Ala²⁷-GLP-1(7-37)OH, Val⁸-His³⁷-GLP-1(7-37)OH, Val⁸-Glu²²-Lys²³-GLP-1(7-37)OH, Val⁸-Glu²²-Glu²³-GLP-1(7-37)OH, Val⁸-Glu²²-Ala²⁷-GLP-1(7-37)OH, Val⁸-Gly³⁴-Lys³⁵-GLP-1(7-37)OH, Val⁸-His³⁷-GLP-1(7-37)OH, Gly⁸-His³⁷-GLP-1(7-37)OH, Val⁸-Glu²²-Ala²⁷-GLP-1(7-37)OH, Gly⁸-Glu²²-Ala²⁷-GLP-1(7-37)OH, Val⁸-Lys²²-Glu²³-GLP-1(7-37)OH, y Gly⁸-Lys²²-Glu²³-GLP-1(7-37)OH.

Otro grupo preferido de los análogos y derivados de GLP - 1 para uso en la presente invención se compone de moléculas de fórmula VI (SEC ID N° 7)



fórmula VI (SEC ID N° 7)

en los que: R₁ se selecciona del grupo constituido por L - histidina, D - histidina, desaminohistidina, 2 - aminohistidina, β - hidroxihistidina, homohistidina, alfa - fluorometilhistidina y alfa - metilhistidina; X se selecciona del grupo constituido por Ala, Gly, Val, Thr, Ile y alfa - metil - Ala; Y se selecciona del grupo constituido por Glu, Gln, Ala, Thr, Ser y Gly; Z se selecciona del grupo constituido por Glu, Gln, Ala, Thr, Ser y Gly; y R₂ es Gly - OH.

Otro grupo preferido de los compuestos GLP - 1 para uso en la presente invención se describe en el documento WO 91/11457 y consiste esencialmente en GLP - 1 (7 - 34), GLP - 1 (7 - 35), GLP - 1 (7 - 36) o GLP - 1 (7 - 37), o la amida de los mismos, y las sales farmacéuticamente aceptables de los mismos, que tienen al menos una modificación seleccionada del grupo constituido por:

- (a) sustitución de lisina en la posición 26 y/o posición 34 por glicina, serina, cisteína, treonina, asparagina, glutamina, tirosina, alanina, valina, isoleucina, leucina, metionina, fenilalanina, arginina o D - lisina o sustitución de glicina, serina, cisteína, treonina, asparagina, glutamina, tirosina, alanina, valina, isoleucina, leucina, metionina, fenilalanina lisina o una D - arginina,
- (b) sustitución de triptófano en la posición 31 por un aminoácido resistente a la oxidación;
- (c) sustitución de al menos uno de: valina en la posición 16; por tirosina

la serina en la posición 18 por lisina; ácido glutámico en la posición 21 por ácido aspártico; glicina en la posición 22 por serina; glutamina en la posición 23 por arginina; alanina en la posición 24 por arginina y lisina en la posición 26 por glutamina; y

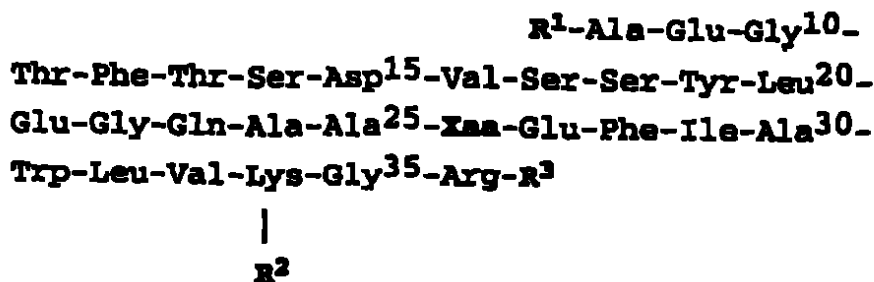
- 5 (d) sustitución de al menos uno de: alanina en la posición 8 por glicina, serina o cisteína, ácido glutámico en la posición 9 por ácido aspártico, glicina, serina, cisteína, treonina, asparagina, glutamina, tirosina, alanina, valina, isoleucina, leucina, metionina o fenilalanina; glicina en la posición 10 por serina, cisteína, treonina, asparagina, glutamina, tirosina, alanina, valina, isoleucina, leucina, metionina o fenilalanina y ácido aspártico en la posición 15 por ácido glutámico; y
- 10
- (e) sustitución de histidina en la posición 7 por glicina, serina, cisteína, treonina, asparagina, glutamina, tirosina, alanina, valina, isoleucina, leucina, metionina o fenilalanina, o la forma D - o N - acilada o alquilada de histidina; en la que en las sustitucion es (a), (b), (d) y (e), los aminoácidos sustituidos pueden opcionalmente estar en la forma D- y los aminoácidos sustituidos en la posición 7 pueden opcionalmente estar en la forma N - acilada o N - alquilada.
- 15

20 Ya que la enzima, dipeptidilpeptidasa IV (abreviadamente DPP IV), puede ser responsable de la rápida inactivación observada *in vivo* de GLP - 1 administrado, [véase, por ejemplo, Mentlein, R., y col., *Eur. J. Biochem.*, 214: 829 - 835 (1993)], se prefieren los análogos y derivados que están protegidos de la actividad de DDP IV en el contexto de una proteína de fusión, y se prefieren

25 más las proteínas de fusión en las que el compuesto GLP - 1 es Gly³ - GLP - 1

(7 - 37) OH, Val⁸ - GLP - 1 (7 - 37) OH, α - metil - Ala⁸ - GLP - 1 (7 - 37) OH, o Gly⁸ - Gln²¹ - GLP - 1 (7 - 37) OH.

Otro grupo preferido de los compuestos GLP - 1 para uso en la presente invención está constituido por los compuestos de fórmula VII (SEC ID N° 8) reivindicado en la patente de Estados Unidos N° 5.512.549 que está expresamente incorporada en esta memoria descriptiva como referencia.



fórmula VII (SEC ID N° 8)

en la que R¹ se selecciona del grupo constituido por 4-imidazopropionilo, 4-imidazoacetilo o 4-imidazo- α , α dimetilacetilo; R² se selecciona del grupo constituido por acilo (C₆ - C₁₀) no ramificado o está ausente; R³ se selecciona del grupo constituido por Gly - OH o NH₂; y Xaa es Lys o Arg.

Los compuestos más preferidos de fórmula IV para uso en la presente invención son aquellos en los que Xaa es Arg y R² es acilo (C₆ - C₁₀) no ramificado. Incluso compuestos más preferidos de fórmula IV para uso en la presente invención son aquellos en los que Xaa es Arg, R² es acilo (C₆ - C₁₀) no ramificado y R³ es Gly - OH. Otros compuestos altamente preferidos de fórmula IV para uso en la presente invención son aquellos en los que Xaa es Arg, R² es acilo (C₆ - C₁₀) no ramificado, R³ es Gly - OH y R¹ es 4-imidazopropionilo. Un compuesto especialmente preferido de fórmula IV para uso en la presente invención es aquel en el que Xaa es Arg, R² es acilo C₈ no ramificado, R³ es Gly - OH y R¹ es 4-imidazopropionilo.

Preferiblemente, los compuestos GLP – 1 comprenden análogos de GLP – 1 en los que la estructura de dichos análogos o fragmentos contiene un aminoácido distinto de alanina en la posición 8 (análogos de posición 8). La estructura también puede incluir L – histidina, D – histidina o formas modificadas de histidina tales como
5 desaminohistidina, 2-aminohistidina, β -hidroxihistidina, homohistidina α -fluorometilhistidina o α -metilhistidina en la posición 7. Es preferible que estos análogos de posición 8 contengan uno o más cambios adicionales en las posiciones 12, 16, 18, 19, 20, 22, 25, 27, 30, 33 y 37 comparados con el correspondiente aminoácido de GLP – 1 (7 – 37) OH nativo. Se prefiere más que estos análogos de posición 8 contengan uno o
10 más cambios adicionales en las posiciones 16, 18, 22, 25, y 33 comparados con el correspondiente aminoácido de GLP – 1 (7 – 37) OH nativo.

En una realización preferida, el análogo de GLP – 1 es GLP – 1 (7 – 37) OH en el que el aminoácido en la posición 12 se selecciona del grupo constituido por triptófano o tirosina. Es más preferido que además de la sustitución en la posición 12, el
15 aminoácido en la posición 8 se sustituya con glicina, valina, leucina, isoleucina, serina, treonina o metionina y más preferiblemente valina o glicina. Es incluso más preferido que además de las sustituciones en las posiciones 12 y 8, el aminoácido en la posición 22 esté sustituido con ácido glutámico.

En otra realización preferida, el análogo de GLP – 1 es GLP – 1 (7 – 37) OH en
20 el que el aminoácido en la posición 16 se selecciona del grupo constituido por triptófano, isoleucina, leucina, fenilalanina o tirosina. Se prefiere más que además de la sustitución en la posición 16, el aminoácido en la posición 8 se sustituya con glicina, valina, leucina, isoleucina, serina, treonina o metionina y más preferiblemente valina o glicina. Incluso se prefiere más que además de las sustituciones en las posiciones 16 y 8,
25 el aminoácido en la posición 22 se sustituya con ácido glutámico. También se prefiere

que además de las sustituciones en las posiciones 16 y 8, el aminoácido en la posición 30 se sustituya con ácido glutámico. También se prefiere que además de las sustituciones en las posiciones 16 y 8, el aminoácido en la posición 37 se sustituya con histidina.

- 5 En otra realización preferida, el análogo de GLP – 1 es GLP – 1 (7 – 37) OH en el que el aminoácido en la posición 18 se selecciona del grupo constituido por triptófano, tirosina, fenilalanina, lisina, leucina o isoleucina, preferiblemente triptófano, tirosina e isoleucina. Se prefiere más que además de la sustitución en la posición 18, el aminoácido en la posición 8 se sustituya con glicina, valina, leucina, isoleucina, serina,
- 10 treonina o metionina y más preferiblemente valina o glicina. Incluso se prefiere más que además de las sustituciones en las posiciones 18 y 8, el aminoácido en la posición 22 se sustituya con ácido glutámico. También se prefiere que además de las sustituciones en las posiciones 18 y 8, el aminoácido en la posición 30 se sustituya con ácido glutámico. También se prefiere que además de las sustituciones en las posiciones 18 y 8, el
- 15 aminoácido en la posición 37 se sustituya con histidina.

- En otra realización preferida, el análogo de GLP – 1 es GLP – 1 (7 – 37) OH en el que el aminoácido en la posición 19 se selecciona del grupo constituido por triptófano o fenilalanina, preferiblemente triptófano. Se prefiere más que además de la sustitución en la posición 19, el aminoácido en la posición 8 se sustituya con glicina, valina,
- 20 leucina, isoleucina, serina, treonina o metionina y más preferiblemente valina o glicina. Incluso se prefiere más que además de las sustituciones en las posiciones 19 y 8, el aminoácido en la posición 22 se sustituya con ácido glutámico. También se prefiere que además de las sustituciones en las posiciones 19 y 8, el aminoácido en la posición 30 se sustituya con ácido glutámico. También se prefiere que además de las sustituciones en
- 25 las posiciones 19 y 8, el aminoácido en la posición 37 se sustituya con histidina.

En otra realización preferida, el análogo de GLP – 1 es GLP – 1 (7 – 37) OH en el que el aminoácido en la posición 20 es fenilalanina, tirosina o triptófano. Se prefiere más que además de la sustitución en la posición 20, el aminoácido en la posición 8 se sustituya con glicina, valina, leucina, isoleucina, serina, treonina o metionina y más preferiblemente valina o glicina. Incluso se prefiere más que además de las sustituciones en la posiciones 20 y 8, el aminoácido en la posición 22 se sustituya con ácido glutámico. También se prefiere que además de las sustituciones en las posiciones 20 y 8, el aminoácido en la posición 30 se sustituya con ácido glutámico. También se prefiere que además de las sustituciones en las posiciones 20 y 8, el aminoácido en la posición 37 se sustituya con histidina.

En otra realización preferida, el análogo de GLP – 1 es GLP – 1 (7 – 37) OH en el que el aminoácido en la posición 25 se selecciona del grupo constituido por valina, isoleucina y leucina, preferiblemente valina. Se prefiere más que además de la sustitución en la posición 25, el aminoácido en la posición 8 se sustituya con glicina, valina, leucina, isoleucina, serina, treonina o metionina y más preferiblemente valina o glicina. Incluso se prefiere más que además de las sustituciones en la posiciones 25 y 8, el aminoácido en la posición 22 se sustituya con ácido glutámico. También se prefiere que además de las sustituciones en las posiciones 25 y 8, el aminoácido en la posición 30 se sustituya con ácido glutámico. También se prefiere que además de las sustituciones en las posiciones 25 y 8, el aminoácido en la posición 37 se sustituya con histidina.

En otra realización preferida, el análogo de GLP – 1 es GLP – 1 (7 – 37) OH en el que el aminoácido en la posición 27 se selecciona del grupo constituido por isoleucina o alanina. Se prefiere más que además de la sustitución en la posición 27, el aminoácido en la posición 8 se sustituya con glicina, valina, leucina, isoleucina, serina,

treonina o metionina y más preferiblemente valina o glicina. Incluso se prefiere más que además de las sustituciones en las posiciones 27 y 8, el aminoácido en la posición 22 se sustituya con ácido glutámico. También se prefiere que además de las sustituciones en las posiciones 27 y 8, el aminoácido en la posición 30 se sustituya con ácido glutámico.

- 5 También se prefiere que además de las sustituciones en las posiciones 27 y 8, el aminoácido en la posición 37 se sustituya con histidina.

- En otra realización preferida, el análogo de GLP - 1 es GLP - 1 (7 - 37) OH en el que el aminoácido en la posición 33 es isoleucina. Se prefiere más que además de la sustitución en la posición 33, el aminoácido en la posición 8 se sustituya con glicina, valina, leucina, isoleucina, serina, treonina o metionina y más preferiblemente valina o glicina. Incluso se prefiere más que además de las sustituciones en las posiciones 33 y 8, el aminoácido en la posición 22 se sustituya con ácido glutámico. También se prefiere que además de las sustituciones en las posiciones 33 y 8, el aminoácido en la posición 30 se sustituya con ácido glutámico. También se prefiere que además de las sustituciones en las posiciones 33 y 8, el aminoácido en la posición 37 se sustituya con histidina.

- Los compuestos de GLP - 1 tienen modificaciones en una o más de las siguientes posiciones: 8, 12, 16, 18, 19, 20, 22, 25, 27, 30, 33 y 37. Los compuestos de GLP - 1 muestran potencia aumentada comparada con GLP - 1 (7 - 37) OH y comprenden la secuencia de aminoácidos de fórmula IX (SEC ID N° 12).

**Xaa₇-Xaa₈-Glu-Gly-Thr-Xaa₁₁-Thr-Ser-Asp-Xaa₁₆-Ser-
Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Gln-Ala-Xaa₂₃-Lys-Xaa₂₇-
Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Lys-Gly-Arg-Xaa₃₇**

Fórmula IX (SEC ID N° 12)

en la que:

Xaa₇ es: L - histidina, D - histidina, desaminohistidina, 2 - aminohistidina, β -

5 hidroxihistidina, homohistidina, α - fluorometilhistidina o α - metilhistidina;

Xaa₈ es: Ala, Gly, Val, Leu, Ile, Ser o Thr;

Xaa₁₂ es: Phe, Trp o Tyr;

Xaa₁₆ es: Val, Trp, Ile, Leu, Phe o Tyr;

Xaa₁₈ es: Ser, Trp, Tyr, Phe, Lys, Ile, Leu, Val;

10 Xaa₁₉ es: Tyr, Trp o Phe;

Xaa₂₀ es: Leu, Phe, Tyr o Trp;

Xaa₂₂ es: Gly, Glu, Asp o Lys;

Xaa₂₅ es: Ala, Val, Ile, o Leu;

Xaa₂₇ es: Glu, Ile o Ala;

15 Xaa₃₀ es: Ala o Glu;

Xaa₃₃ es: Val o Ile; y

Xaa₃₇ es: Gly, His, NH₂ o está ausente.

Algunos compuestos preferidos de GLP - 1 de fórmula IX incluyen

GLP-1(7-37)OH, GLP-1(7-36)-NH₂, Gly⁸-GLP-1(7-37)OH, Gly⁸-GLP-1(7-36)NH₂, Val⁸-GLP-1(7-37)OH, Val⁸-GLP-1(7-36)NH₂, Leu⁸-GLP-1(7-37)OH, Leu⁸-GLP-1(7-36)NH₂, Ile⁸-GLP-1(7-37)OH, Ile⁸-GLP-1(7-36)NH₂, Ser⁸-GLP-1(7-37)OH, Ser⁸-GLP-1(7-36)NH₂, Thr⁸-GLP-1(7-37)OH, Thr⁸-GLP-1(7-36)NH₂, Val⁸-Tyr¹²-GLP-1(7-37)OH, Val⁸-Tyr¹²-GLP-1(7-36)NH₂, Val⁸-Tyr¹⁶-GLP-1(7-37)OH, Val⁸-Tyr¹⁶-GLP-1(7-36)NH₂, Val⁸-Glu²²-GLP-1(7-37)OH, Val⁸-Glu²²-GLP-1(7-36)NH₂, Gly⁸-Glu²²-GLP-1(7-37)OH, Gly⁸-Glu²²-GLP-1(7-36)NH₂, Val⁸-Asp²²-GLP-1(7-37)OH, Val⁸-Asp²²-GLP-1(7-36)NH₂, Gly⁸-Asp²²-GLP-1(7-37)OH, Gly⁸-Asp²²-GLP-1(7-36)NH₂, Val⁸-Lys²²-GLP-1(7-37)OH, Val⁸-Lys²²-GLP-1(7-36)NH₂, Gly⁸-Lys²²-GLP-1(7-37)OH, Gly⁸-Lys²²-GLP-1(7-36)NH₂, Leu⁸-Glu²²-GLP-1(7-37)OH, Leu⁸-Glu²²-GLP-1(7-36)NH₂, Ile⁸-Glu²²-GLP-1(7-

37)OH, Ile⁸-Glu²²-GLP-1(7-36)NH₂, Leu⁸-Asp²²-GLP-1(7-37)OH, Leu⁸-Asp²²-GLP-1(7-36)NH₂, Ile⁸-Asp²²-GLP-1(7-37)OH, Ile⁸-Asp²²-GLP-1(7-36)NH₂, Leu⁸-Lys²²-GLP-1(7-37)OH, Leu⁸-Lys²²-GLP-1(7-36)NH₂, Ile⁸-Lys²²-GLP-1(7-37)OH, Ile⁸-Lys²²-GLP-1(7-36)NH₂, Ser⁸-Glu²²-GLP-1(7-37)OH, Ser⁸-Glu²²-GLP-1(7-36)NH₂, Thr⁸-Glu²²-GLP-1(7-37)OH, Thr⁸-Glu²²-GLP-1(7-36)NH₂, Ser⁸-Asp²²-GLP-1(7-37)OH, Ser⁸-Asp²²-GLP-1(7-36)NH₂, Thr⁸-Asp²²-GLP-1(7-37)OH, Thr⁸-Asp²²-GLP-1(7-36)NH₂, Ser⁸-Lys²²-GLP-1(7-37)OH, Ser⁸-Lys²²-GLP-1(7-36)NH₂, Thr⁸-Lys²²-GLP-1(7-37)OH, Thr⁸-Lys²²-GLP-1(7-36)NH₂, Glu²²-GLP-1(7-37)OH, Glu²²-GLP-1(7-36)NH₂, Asp²²-GLP-1(7-37)OH, Asp²²-GLP-1(7-36)NH₂, Lys²²-GLP-1(7-37)OH, Lys²²-GLP-1(7-36)NH₂, Val⁸-Ala²⁷-GLP-1(7-37)OH, Val⁸-Glu²²-Ala²⁷-GLP-1(7-37)OH, Val⁸-Glu³⁰-GLP-1(7-37)OH, Val⁸-Glu³⁰-GLP-1(7-36)NH₂, Gly⁸-Glu³⁰-GLP-1(7-37)OH, Gly⁸-Glu³⁰-GLP-1(7-36)NH₂, Leu⁸-Glu³⁰-GLP-1(7-37)OH, Leu⁸-Glu³⁰-GLP-1(7-36)NH₂, Ile⁸-Glu³⁰-GLP-1(7-37)OH, Ile⁸-Glu³⁰-GLP-1(7-36)NH₂, Ser⁸-Glu³⁰-GLP-1(7-37)OH, Ser⁸-Glu³⁰-GLP-1(7-36)NH₂, Thr⁸-Glu³⁰-GLP-1(7-37)OH, Thr⁸-Glu³⁰-GLP-1(7-36)NH₂, Val⁸-His³⁷-GLP-1(7-37)OH, Val⁸-His³⁷-GLP-1(7-36)NH₂, Gly⁸-His³⁷-GLP-1(7-37)OH, Gly⁸-His³⁷-GLP-1(7-36)NH₂, Leu⁸-His³⁷-GLP-1(7-37)OH, Leu⁸-His³⁷-GLP-1(7-36)NH₂, Ile⁸-His³⁷-GLP-1(7-37)OH, Ile⁸-His³⁷-GLP-1(7-36)NH₂, Ser⁸-His³⁷-GLP-1(7-37)OH, Ser⁸-His³⁷-GLP-1(7-36)NH₂, Thr⁸-His³⁷-GLP-1(7-37)OH, Thr⁸-His³⁷-GLP-1(7-36)NH₂.

Algunos compuestos preferidos GLP - 1 de fórmula IX que tienen sustituciones múltiples incluyen GLP - 1 (7 - 37) OH en el que la posición 8 es valina o glicina, la posición 22 es ácido glutámico, la posición 16 es tirosina, leucina o triptófano, la

5 posición 18 es tirosina, triptófano o isoleucina, la posición 25 es valina y la posición 33 es isoleucina. Otros compuestos GLP - 1 preferidos incluyen los siguientes: Val⁸-Try¹⁶-GLP-1 (7 -

37)OH, Val⁸-Tyr¹²-Glu²²-GLP-1(7-37)OH, Val⁸-Tyr¹⁶-Phe¹⁹-GLP-1(7-37)OH, Val⁸-Tyr¹⁶-Glu²²-GLP-1(7-37)OH, Val⁸-Trp¹⁶-Glu²²-GLP-1(7-37)OH, Val⁸-Leu¹⁶-Glu²²-GLP-1(7-37)OH, Val⁸-Ile¹⁶-

Glu²²-GLP-1(7-37)OH, Val⁸-Phe¹⁶-Glu²²-GLP-1(7-37)OH; Val⁸-Trp¹⁶-Glu²²-GLP-1(7-37)OH, Val⁸-Tyr¹⁸-Glu²²-GLP-1(7-37)OH, Val⁸-Phe¹⁸-Glu²²-GLP-1(7-37)OH, and Val⁸-Ile¹⁸-Glu²²-GLP-1(7-37)OH.

Los compuestos GLP - 1 de la presente invención también comprenden compuestos de exendina. La exendina - 3 y la exendina - 4 son péptidos biológicamente activos aislados primero de los venenos del *Helodermatidae* y se ha demostrado que se unen al receptor de GLP - 1 y estimulan la producción de H⁺ cAMP dependiente en células parietales de mamíferos. La exendina - 3 y la exendina - 4 son ambos péptidos de 39 aminoácidos que son aproximadamente homólogos en un 53% al GLP -1. Actúan como potentes agonistas de la actividad de GLP - 1. Especialmente, un derivado truncado del extremo N terminal de exendina, conocido como exendina (aminoácidos 9 - 39) es un inhibidor de exendina - 3, exendina - 4 y GLP - 1.

Un compuesto de exendina típicamente comprende un polipéptido que tiene la secuencia de aminoácidos de exendina - 3, exendina - 4 o un análogo o fragmento de las mismas. La exendina - 3 y la exendina - 4 se describen en la patente de Estados Unidos N° 5.424.286.

La exendina - 3 tiene la secuencia de aminoácidos de SEC ID N° 9:

7 8 9 10 11 12 13 14 15 16 17

His - Ser - Asp - Gly - Thr - Phe - Thr - Ser - Asp - Leu - Ser -

18 19 20 21 22 23 24 25 26 27 28

20 Lys - Gln - Met - Glu - Glu - Glu - Ala - Val - Arg - Leu - Phe -

29 30 31 32 33 34 35 36 37 38 39
Ile – Glu – Trp – Leu – Lys – Asn – Gly – Gly – Pro – Ser – Ser -

40 41 42 43 44 45
Gly – Ala – Pro – Pro – Pro – Ser

5 (SEC ID N° 9)

La exendina – 4 tiene la secuencia de aminoácidos de SEC ID N° 10:

7 8 9 10 11 12 13 14 15 16 17
His – Gly – Glu – Gly – Thr – Phe – Thr – Ser – Asp – Leu – Ser -

18 19 20 21 22 23 24 25 26 27 28

10 Lys – Gln – Met – Glu – Glu – Glu – Ala – Val – Arg – Leu – Phe -

29 30 31 32 33 34 35 36 37 38 39
Ile – Glu – Trp – Leu – Lys – Asn – Gly – Gly – Pro – Ser – Ser -

40 41 42 43 44 45
Gly – Ala – Pro – Pro – Pro – Ser

15 (SEC ID N° 10)

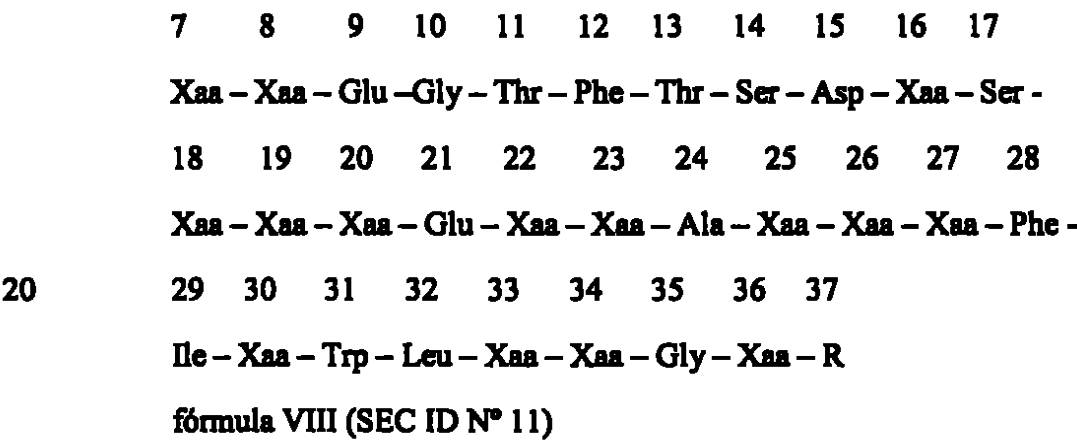
Los compuestos GLP – 1 también incluyen fragmentos de exendina que son polipéptidos obtenidos después del truncamiento de uno o más aminoácidos del extremo N terminal y/o extremo C terminal de exendina o un análogo de exendina. Además, los compuestos de GLP – 1 incluyen polipéptidos de exendina en los que uno o más aminoácidos se han añadido al extremo N terminal y/o extremo C terminal de exendina o fragmentos de los mismos. Los compuestos de exendina de este tipo tienen hasta aproximadamente cuarenta y cinco aminoácidos.

Los compuestos GLP – 1 también incluyen “análogos de exendina”. Un análogo de exendina tiene suficiente homología con la exendina – 4, exendina – 3 o un

fragmento de las mismas de tal manera que el análogo tiene actividad insulínica. La actividad de los fragmentos de exendina y/o análogos se pueden valorar utilizando ensayos *in vitro* tales como los descritos en el documento EP 619.322 y la patente de Estados Unidos 5.120.712.

5 Preferiblemente, un análogo de exendina tiene la secuencia de aminoácidos de la exendina - 4 o un fragmento de la misma, modificada de manera que entre uno, dos, tres, cuatro o cinco aminoácidos son diferentes del aminoácido en la posición correspondiente de la exendina - 4 o el fragmento de exendina - 4. En la nomenclatura utilizada en esta memoria descriptiva para designar compuestos de exendina, el
10 aminoácido sustituyente y su posición se indican antes de la estructura parental. Por ejemplo, Val⁸ - exendina - 4 designa un compuesto de exendina en el que la glicina que normalmente se encuentra en la posición 8 de la exendina - 4 se ha sustituido con valina.

Otro grupo preferido de los compuestos GLP - 1 se compone de GLP - 1/
15 análogos de exendina-4 de fórmula VIII (SEC ID N° 11).



en la que:

Xaa en la posición 7 es L - histidina, D - histidina, desaminohistidina, 2 -
25 aminohistidina, β - hidroxihistidina, homohistidina, α - fluorometilhistidina o α -

metilhistidina;

- Xaa en la posición 8 es: Gly, Ala o Val;
- Xaa en la posición 16 es: Leu o Val;
- Xaa en la posición 18 es: Lys o Ser;
- 5 Xaa en la posición 19 es: Gln o Tyr;
- Xaa en la posición 20 es: Met o Leu;
- Xaa en la posición 22 es: Glu o Gln;
- Xaa en la posición 23 es: Glu o Gln;
- Xaa en la posición 25 es: Val o Ala;
- 10 Xaa en la posición 26 es: Arg o Lys;
- Xaa en la posición 27 es: Leu o Glu;
- Xaa en la posición 30 es: Glu o Ala;
- Xaa en la posición 33 es: Val o Lys;
- Xaa en la posición 34 es: Asn o Lys;
- 15 Xaa en la posición 36 es: Gly o Arg; y

R en la posición 37 es: Gly, Pro, Pro – Ser – Ser – Gly – Ala – Pro – Pro – Pro – Ser o está ausente. La actividad de 18 especies diferentes que caen dentro de este género se proporciona en la tabla 6.

Además los análogos de exendina que son útiles en la presente invención se describen en las publicaciones de las patentes PCT WO 99/25728 (Beeley y col.), WO 20 99/25727 (Beeley y col.), WO 98/05351 (Young y col.), WO 99/40788 (Young y col.), WO 99/07404 (Beeley y col.) y WO 99/43708 (Knudsen y col.).

Las proteínas de fusión de GLP – 1 de la presente invención pueden comprender sitios de glicosilación. La glicosilación es una modificación química en la que restos de 25 azúcar se añaden a la proteína en sitios específicos. La glicosilación de proteínas juega

un papel para asegurar la carga correcta, confirmación y estabilidad la proteína de maduración y puede dirigir la proteína a la superficie de la célula y a una secreción eventual de la proteína. Lo más importante es que, la glicosilación efectúa la tasa de eliminación de muchas proteínas *in vivo*. Los azúcares pueden estar O - unidos o N - unidos- En general, los azúcares O - unidos se añaden al oxígeno del grupo hidroxilo de la serina y treonina, mientras que los azúcares N - unidos se añaden al nitrógeno amida de la asparagina. El sitio de consenso de la N - glicosilación es Asn X1 X2 siendo X1 cualquier aminoácido excepto Pro y X2 es Ser o Thr.

Los compuestos GLP - 1 generalmente no se glicosilan *en vivo*; sin embargo, es de interés que las proteínas de fusión de GLP - 1 de la presente invención que comprenden un compuesto GLP - 1 con una extensión del extremo C terminal condensado con una secuencia Fc se glicosilan en la última serina en la extensión del extremo C terminal (SSGAPPPS*) y en la treonina en la posición 11 en la región del extremo N terminal de Fc (AEPKSCDKTHT*CPPC . . .).

15

Proteínas de fusión Fc heterólogas

Los compuestos GLP - 1 descritos anteriormente se pueden condensar directamente o mediante un engarce de péptidos a la porción Fc de una inmunoglobulina.

20

Las inmunoglobulinas son moléculas que contienen cadenas de polipéptidos que se mantienen juntas por medio de enlaces disulfuro, que tienen típicamente dos cadenas ligeras y dos cadena pesadas. En cada cadena un dominio (V) tiene una secuencia de aminoácidos variable dependiendo de la especificidad de anticuerpo de la molécula. Los otros dominios (C) tienen una secuencia bastante constante común con las moléculas de la misma clase.

25

Como se utiliza en esta memoria descriptiva, la porción Fc de una inmunoglobulina tiene el significado comúnmente dado al término en el campo de la inmunología. Específicamente, este término se refiere a un fragmento de anticuerpo que se obtiene retirando las dos regiones de unión del antígeno (fragmentos Fab) del anticuerpo. Una forma para retirar los fragmentos Fab es digerir la inmunoglobulina con proteasa de papaína. De este modo, la porción Fc se forma a partir de fragmentos de tamaño aproximadamente igual de la región constante de ambas cadenas pesadas, que se asocian a través de interacciones no covalentes y enlaces disulfuro. La porción Fc puede incluir las regiones flexibles y se extienden mediante los dominios CH₂ y CH₃ al extremo C terminal del anticuerpo. Las regiones flexibles representativas de las inmunoglobulinas de seres humanos y ratones se pueden encontrar en Antibody Engineering, A Practical Guide, Borrebaeck, C. A. K., ed., W. H. Freeman and Co., 1992, las enseñanzas de las cuales se incorporan en esta memoria descriptiva como referencia. La porción Fc puede además incluir uno o más sitios de glicosilación. La secuencia de aminoácidos de una proteína de Fc representativa contiene una región flexible, dominios CH₂ y CH₃ y un sitio de N-glicosilación en la posición 82 se muestran en la figura 1.

Hay cinco tipos de regiones Fc de inmunoglobulinas humanas con efecto diferente o y propiedades farmacocinéticas: IgG, IgA, IgM, IgD e IgE. La IgG es la inmunoglobulina más abundante en suero. La IgG también tiene la semivida más larga en suero de todas las inmunoglobulinas (23 días). A diferencia de otras inmunoglobulinas, la IgG se recircula eficientemente después de la unión a un receptor de Fc. Hay cuatro subclases de IgG G1, G2, G3 y G4, cada una de las cuales tiene diferente efecto o funciones. G1, G2 y G3 se pueden unir a C1q y fijar el complemento mientras la G4 no puede. Incluso aunque la G3 es capaz de unirse a C1q más

eficazmente que G1, G1 es más eficaz en la mediación de la lisis celular dirigida por el complemento. G2 fija el complemento muy ineficazmente. El sitio de unión a C1q en IgG se localiza en la región carboxi terminal del dominio CH₂.

Todas las subclases de IgG son capaces de unirse a los receptores de Fc (CD16, CD32, CD64) siendo G1 y G3 más eficaces que G2 y G4. La región de unión al receptor de Fc de IgG se forma mediante restos localizados tanto en la región flexible como en las regiones carboxi terminal del dominio CH₂.

La IgA puede existir tanto en forma monomérica como dimérica que se mantienen unidas por una cadena J. La IgA es la segunda Ig más abundante en el suero, pero tiene una semivida de 6 días solamente. La IgA tiene tres efectos o funciones. Se une a un receptor específico de IgA en macrófagos y eosinófilos, que conducen la fagocitosis y la desgranulación respectivamente. También puede fijar el complemento mediante una vía alternativa desconocida.

La IgM se expresa en forma bien de pentámero o hexámero, ambas formas se mantienen juntas mediante una cadena J. La IgM tiene una semivida en suero de 5 días. Se une débilmente a C1q mediante un sitio de unión localizado en su dominio CH₃. La IgD tiene una semivida de 3 días en suero. No está claro qué efecto o funciones se atribuyen a esta Ig. La IgE es una Ig monomérica y tiene una semivida en suero de 2,5 días. La IgE se une a dos receptores de Fc que conducen la desgranulación y dan como resultado la liberación de agentes proinflamatorios.

Dependiendo del efecto deseado *in vivo*, las proteínas de fusión heterólogas de la presente invención pueden contener cualquiera de los isotipos descritos anteriormente o pueden contener regiones de Fc mutadas en las que las funciones de unión al complemento y/o al receptor de Fc se han alterado. De este modo, las proteínas de fusión heterólogas de la presente invención pueden contener la porción Fc entera de una

inmunoglobulina, los fragmentos de la porción Fc de una inmunoglobulina, o análogos de los mismos condensados con un compuesto GLP - 1.

Las proteínas de fusión de la presente invención pueden constar de proteínas de cadena sencilla o polipéptidos de cadena múltiple. Dos o más proteínas de fusión de Fc
5 se pueden producir de manera que interactúen mediante enlaces disulfuro que se forman naturalmente entre las regiones Fc. Estos multímeros pueden ser homogéneos con relación al compuesto GLP -1 o pueden contener diferentes compuestos GLP - 1 condensados en el extremo N terminal de la porción Fc de la proteína de fusión.

Independientemente de la estructura final de la proteína de fusión, las regiones
10 Fc o de tipo Fc deben servir para prolongar la semivida *in vivo* en plasma del compuesto GLP - 1 condensado al extremo N terminal. Además, el compuesto GLP - 1 condensado debe mantener alguna actividad biológica. Se puede demostrar un aumento en la semivida utilizando el procedimiento descrito en el ejemplo 7 en el que la semivida de la proteína de fusión se compara con la semivida del compuesto de GLP -
15 1 solo. La actividad biológica se puede determinar mediante procedimientos *in vitro* e *in vivo* conocidos en la técnica. Los ensayos biológicos representativos se describen en los ejemplos 6, 8 y 9.

Ya que la región Fc de la IgG producida mediante proteólisis tiene la misma semivida *in vivo* que la molécula intacta de IgG y los fragmentos Fab se degradan
20 rápidamente, se cree que la secuencia relevante para prolongar la semivida reside en los dominios CH₂ y CH₃. Además, se ha mostrado en la bibliografía que las velocidades catabólicas de las variantes de IgG que no se unen al receptor de Fc de alta afinidad o C1q son indistinguibles de la velocidad de eliminación del anticuerpo de tipo natural parental, indicando que el sitio catabólico es distinto de los sitios implicados en la unión
25 al receptor de Fc o C1q. [Wawrzynczak y col., (1992) *Molecular Immunology* 29: 221].

Los estudios de mutagénesis específica puntual que usan una región Fc de IgG1 murina sugerían que el sitio de la región Fc de IgG1 que controla la velocidad metabólica se localiza en la interfase del dominio de CH₂ y CH₃.

En función de estos estudios, las regiones Fc se pueden modificar en el sitio
5 catabólico para optimizar la semivida de las proteínas de fusión. Es preferible que la región Fc utilizada para las proteínas de fusión heterólogas de la presente invención se derive de una región Fc de una IgG1 o de una IgG4. Incluso se prefiere más que la región Fc sea IgG4 o derivada de IgG4. Preferiblemente la región Fc de IgG contiene ambas regiones CH₂ y CH₃ que incluyen la región flexible (área en la molécula de
10 inmunoglobulina que contiene muchos residuos del aminoácido prolina y es donde el fragmento Fc une uno de los dos fragmentos Fab (nota del traductor)).

Proteínas de fusión heterólogas de albúmina

Los compuestos de GLP - 1 descritos anteriormente se pueden condensar
15 directamente o mediante un engarce de péptidos a albúmina o un análogo, fragmento o derivado de la misma.

En general, las proteínas de albúmina que forman parte de las proteínas de fusión de la presente invención se pueden derivar a partir de albúmina clonada de cualquier especie. Sin embargo, se prefieren la albúmina humana y los fragmentos y
20 análogos de la misma para reducir el riesgo de que la proteína de fusión sea inmunogénica en seres humanos. La albúmina de suero humano (abreviadamente HSA) consta de una cadena polipeptídica no glicosilada de 585 aminoácidos con un peso molecular de la fórmula de 66.500. La secuencia de aminoácidos de la HSA humana se muestra en la figura 2. [Véase Meloun y col (1975) FEBS Letters 58: 136; Behrens y
25 col, (1975) Fed. Proc. 34: 591; Lawn y col. (1981) Nucleic Acids Research 9: 6102 -

6114; Minghetti y col., (1986) J. Biol Chem. 261: 6747]. Se han descrito una diversidad de variantes polimórficas así como análogos y fragmentos de albúmina. [Véase Weitkamp y col., (1973) Ann. Hum. Genet. 37: 219]. Por ejemplo, en el documento EP 322.094, los inventores describen diversas formas más pequeñas de HSA. Algunos de estos fragmentos incluyen HSA (1 – 373), HSA (1 – 388), HSA (1 – 389), HSA (1 – 369) y HSA (1 – 419) y fragmentos entre 1 – 369 y 1 – 419. El documento 399.666 describe fragmentos de albúmina que incluyen HSA (1 – 177) y HSA (1 – 200) y fragmentos entre HSA (1 – 177) y HSA (1 – 200).

Se entiende que las proteínas de fusión heterólogas de la presente invención incluyen compuestos GLP – 1 que se acoplan a cualquier proteína de albúmina que incluye fragmentos, análogos y derivados en los que dicha proteína de fusión es biológicamente activa y tiene una semivida en plasma más larga que el compuesto de GLP – 1 solo. De este modo, la porción de albúmina de la proteína de fusión no necesita necesariamente tener una semivida en plasma igual a la de la albúmina humana nativa. Se conocen o se pueden generar los fragmentos, análogos y derivados que tengan semividas más largas o que tengan semividas intermedias a las la de albúmina humana nativa y el compuesto de GLP – 1 de interés.

Las proteínas de fusión heterólogas de la presente invención comprenden proteínas que tienen sustituciones conservadoras de aminoácidos en el compuesto GLP – 1 y/o la porción de Fc o de albúmina de la proteína de fusión. Una “sustitución conservadora” es el reemplazo de un aminoácido por otro aminoácido que tiene la misma carga neta electrónica y aproximadamente el mismo tamaño y forma. Los aminoácidos con cadenas laterales de aminoácidos alifáticas o alifáticas sustituidas tienen aproximadamente el mismo tamaño cuando número total de carbonos y heteroátomos en sus cadena laterales difieren en no más de aproximadamente cuatro.

Tienen aproximadamente la misma forma cuando el número total de ramificaciones en sus cadenas laterales difiere en no más de uno. Los aminoácidos con grupos fenilo o fenilo sustituidos en sus cadenas laterales se considera que tienen aproximadamente el mismo tamaño y forma. Excepto que en esta memoria descriptiva se especifique otra
5 cosa, las sustituciones conservadoras se hacen preferiblemente con aminoácidos de origen natural.

Sin embargo, el término "aminoácido" se utiliza en esta memoria descriptiva en su sentido más amplio e incluye aminoácidos de origen natural así como aminoácidos de origen no natural, incluyendo análogos y derivados de aminoácidos. Esto último
10 incluye moléculas que contienen un resto aminoácido. Un experto en la técnica reconocerá, en vista de esta amplia definición, que la referencia en esta memoria descriptiva de un aminoácido incluye, por ejemplo, aminoácidos L proteogénicos de origen natural; aminoácidos D; aminoácidos químicamente modificados tales como los análogos y derivados de aminoácidos; aminoácidos no proteogénicos de origen natural
15 tales como norleucina, β - alanina, ornitina, GABA, etc.; y compuestos sintetizados químicamente que tienen propiedades conocidas en la técnica que son características de aminoácidos. Como se utiliza en esta memoria descriptiva el término "proteogénico" indica que el aminoácido se puede incorporar a un péptido, polipéptido o proteína en una célula mediante una vía metabólica.

20 La incorporación de aminoácidos no naturales, que incluyen aminoácidos no nativos sintéticos, aminoácidos sustituidos o uno o más aminoácidos D en las proteínas de fusión heterólogas de la presente invención puede ser ventajosa en numerosas formas diferentes. Los péptidos que contienen aminoácidos D etc., exhiben una estabilidad incrementada *in vitro* o *in vivo* comparada con los homólogos que contienen
25 aminoácidos L. De este modo, la construcción de los péptidos, etc., que incorporan

aminoácidos D puede ser particularmente útil cuando se desea o se requiere una estabilidad intracelular mayor. Más específicamente, los péptidos D, etc., son resistentes a peptidasas y proteasas endógenas, proporcionando por lo tanto una biodisponibilidad mejorada de la molécula y vidas prolongadas *in vivo* cuando se desean tales propiedades. Adicionalmente, los péptidos D, etc., no se pueden elaborar eficazmente para una presentación principal de histocompatibilidad de complejo clase II restringido respecto a las células auxiliares T, y por lo tanto es menos probable que induzcan respuestas inmunes humorales en el organismo entero.

Además de los análisis estructura/función de los diversos polipéptidos comprendidos en la presente invención, hay numerosos factores que se pueden considerar cuando se seleccionan los aminoácidos para sustitución. Un factor que se puede considerar para hacer tales cambios es el índice hidropático de los aminoácidos. La importancia del índice hidropático de los aminoácidos que confiere la función interactiva biológica en una proteína la han descrito Kyte y Doolittle (1982, *J. Mol. Biol.*, 157: 105 – 132). Se acepta que el carácter hidropático relativo de los aminoácidos contribuye a la estructura secundaria de la proteína resultante. Esto, a su vez, afecta a la interacción de la proteína con moléculas tales como enzimas, sustratos, receptores, ligandos, DNA, anticuerpos, antígenos, etc. En función de su hidrofobicidad y características de carga, cada aminoácido se le ha asociado un índice hidropático como sigue:

Isoleucina (+ 4,5); valina (+ 4,2); leucina (+3,8); fenilalanina (+ 2,8); cisteína/cistina (+ 2,5); metionina (+ 1,9); alanina (+ 1,8); glicina (- 0,4); treonina (- 0,7); serina (- 0,8); triptófano (- 0,9); tirosina (- 1,3); prolina (- 1,6); histidina (- 3,2); glutamato/glutamina/aspartato/asparagina (- 3,5); lisina (- 3,9); y arginina (- 4,5).

Como se sabe en la técnica, ciertos aminoácidos en un péptido, polipéptido o

258
255

proteína se pueden sustituir con otros aminoácidos que tengan un índice o valor hidropático similar y así producir un péptido resultante, etc, que tenga una actividad biológica similar o incluso mejorada. Cuando se hacen tales cambios es preferible que los aminoácidos que tienen índices hidropáticos entre ± 2 se sustituyan con otro. Las

5 sustituciones más preferidas son aquellas en las que los aminoácidos tienen índices hidropáticos entre ± 1 . Las sustituciones más preferidas son aquellas en las que los aminoácidos tienen índices hidropáticos dentro de $\pm 0,5$.

Los aminoácidos similares se pueden sustituir también en función de la hidrofiliidad. La patente de Estados Unidos N° 4.554.101 describe que la mayor

10 hidrofiliidad media local de una proteína, según se rige por la hidrofiliidad de sus aminoácidos adyacentes, se correlaciona con una propiedad biológica de la proteína. Se han asignado a los aminoácidos los siguientes los siguientes valores de hidrofiliidad: arginina/lisina (+ 3,0); aspartato/glutamato (+ 3,0 $\pm$ 1); serina (+ 0,3); asparagina/glutamina (+ 0,2); glicina (0); treonina (- 0,4); prolina (- 0,5 $\pm$ 1);

15 alanina/histidina (- 0,5); cisteína (- 1,0); metionina (- 1,3); valina (- 1,5); leucina/isleucina (- 1,8); tirosina (- 2,3); fenilalanina (- 2,5); y triptófano (- 3,4). De este modo, un aminoácido en un péptido, polipéptido o proteína se puede sustituir con otro aminoácido que tenga un valor similar de hidrofiliidad y así producir un péptido resultante, etc, que tenga una actividad biológica similar, es decir manteniendo todavía

20 una correcta función biológica. Cuando se hacen tales cambios, los aminoácidos que tienen índices hidropáticos dentro de ± 2 se sustituyen preferiblemente con otro, se prefieren más los que están dentro de ± 1 y los más preferidos son los que están dentro de $\pm 0,5$.

Como se ha indicado anteriormente, las sustituciones de los aminoácidos en las

25 proteínas de fusión de la presente invención se pueden basar en la semejanza relativa de

los sustituyentes de la cadena lateral de los aminoácidos, por ejemplo, su hidrofobicidad, hidrofiliidad, carga, tamaño, etc. Además, las sustituciones se pueden realizar en función de la propensión de la estructura secundaria. Por ejemplo, un aminoácido helicoidal se puede sustituir con un aminoácido que preservara la estructura helicoidal. Las sustituciones ejemplares que tienen en cuenta diversas de las características anteriores con el fin de producir cambios conservadores de aminoácidos que dan como resultado cambios silenciosos en los péptidos presentes, etc., se pueden seleccionar de otros miembros de la clase a la que los aminoácidos de origen natural pertenecen. Los aminoácidos se pueden dividir en los siguientes cuatro grupos: (1) aminoácidos ácidos; (2) aminoácidos básicos; (3) aminoácidos polares neutros; y aminoácidos (4) aminoácidos no polares neutros.

Procedimientos generales para la preparación de proteínas de fusión heterólogas de la presente invención.

Aunque las proteínas de fusión heterólogas de la presente invención se pueden preparar mediante una diversidad de diferentes procedimientos, se prefieren los procedimientos de recombinación. Para propósitos de la presente invención, como se describe y se reivindica en la presente invención, se definen más adelante los siguientes términos y abreviaturas de biología molecular general. Los términos y abreviaturas utilizados en este documento tienen sus significados normales a no ser que se indique otra cosa. Por ejemplo, "°C" se refiere a grados Celsius; "mmol" se refiere a milimol o milimoles; "mg" se refiere a miligramos; "µg" se refiere a microgramos; "ml" se refiere a mililitros; y "µl" se refiere a microlitros. Las abreviaturas de los aminoácidos son como se indica en 37 C. F. R. § 1.822 (b) (2) (1994).

"Pares de bases" o "bp" como se utiliza en esta memoria descriptiva se refiere a

DNA o RNA. Las abreviaturas A, C, G, y T, cuando se presentan en moléculas de DNA, corresponden a las formas 5' – monofosfato de los desoxirribonucleósidos (desoxi) adenosina, (desoxi) citidina, (desoxi) guanosina y timidina respectivamente. Las abreviaturas U, C, G y A, cuando se presentan en moléculas de RNA, corresponden a las formas 5' – monofosfato de los ribonucleósidos uridina, citidina, guanosina y adenosina respectivamente. En el DNA de doble cadena, el par de bases se refiere a la asociación de A con T o C con G. En un DNA/RNA, el par de bases heterodúplex puede referirse a una asociación de A con U o C con G. (Véase la definición de "complementario" más adelante.)

- 10 "Digestión" o "restricción" de DNA se refiere a la escisión catalítica del DNA con una enzima de restricción que actúa solamente en ciertas secuencias en el DNA ("endonucleasas de secuencias específicas"). Las diversas enzimas de restricción utilizadas en esta memoria descriptiva están comercialmente disponibles y sus condiciones de reacción, cofactores y otros requerimientos se utilizaron como conocen
- 15 los expertos en la técnica. Las cantidades de tampones y sustrato adecuadas para las enzimas de restricción particulares se especifican por el fabricante o se pueden encontrar fácilmente en la bibliografía.

- "Ligación" se refiere al procedimiento de formación enlaces fosfodiéster entre dos fragmentos de ácido nucleico de doble cadena. A no ser que se indique otra cosa, la
- 20 ligación se puede llevar a cabo utilizando tampones y condiciones conocidas con una ligasa de DNA, tal como T4 DNA ligasa.

- "Plásmido" se refiere a elemento genético de autorreplicación extracromosomal (usualmente). Los plásmidos se designan generalmente mediante una minúscula "p" seguido de letras y/o números. Los plásmidos de partida en esta memoria descriptiva
- 25 están, o bien comercialmente disponibles, disponibles públicamente en función no

restringida, o se pueden construir a partir de plásmidos disponibles de acuerdo con procedimientos publicados. Además, los plásmidos equivalentes a los descritos se conocen en la técnica y serán evidentes para los expertos en la técnica.

5 “Vector de clonación de DNA recombinante” como se utiliza en esta memoria descriptiva se refiere a cualquier agente de replicación autónomo, incluyendo, pero no limitado a, plásmidos y fagos, que comprenden una molécula de DNA a la que se pueden añadir o se han añadido uno o más segmentos adicionales de DNA.

10 “Vector de expresión de DNA recombinante” como se utiliza en esta memoria descriptiva se refiere a cualquier vector de clonación de DNA recombinante en el que se ha incorporado un promotor para controlar la transcripción del DNA insertado.

 “Transcripción” se refiere al procedimiento por el que la información contenida en una secuencia de nucleótidos de DNA se transfiere a una secuencia complementaria de RNA.

15 “Transfección” se refiere a la captación de un vector de expresión por una célula huésped se expresen o no, de hecho, las secuencias de codificación. Los expertos en la técnica conocen numerosos procedimientos de transfección, por ejemplo, coprecipitación con fosfato cálcico, transfección de liposomas y electroporación. La transfección exitosa se reconoce generalmente cuando se produce cualquier indicación
20 de la operación de este vector en la célula huésped.

 “Transformación” se refiere a la introducción de DNA en un organismo de manera que el DNA es replicable, bien como un elemento extracromosomal o por integración cromosomal. Los procedimientos de transformación de las células huésped eucarióticas y bacterianas se conocen bien en la técnica, muchos de estos
25 procedimientos, tales como inyección nuclear, fusión protoplástica o tratamiento con

Tales técnicas y condiciones son bien conocidas por los expertos en la técnica.

"Secuencia de aminoácidos aislados" se refiere a cualquier secuencia de aminoácidos, construida o sintetizada, que es posicionalmente distinta de la secuencia de origen natural.

- 5 "Compuesto de DNA aislado" se refiere a cualquier secuencia de DNA, construida o sintetizada, que es posicionalmente distinta de su localización natural en DNA genómico.

- 10 "Compuesto de ácido nucleico aislado" se refiere a cualquier secuencia de DNA o RNA, construida o sintetizada, que es posicionalmente distinta de su localización natural.

"Cebador" se refiere a un fragmento de ácido nucleico que funciona como sustrato iniciador de una elongación enzimática o sintética.

"Promotor" se refiere a una secuencia de DNA que dirige la transcripción de DNA a RNA.

- 15 "Sonda" se refiere a un compuesto de ácido nucleico o fragmento del mismo, que se hibrida con otro compuesto de ácido nucleico.

- 20 "Rigurosidad" de las reacciones de hibridación se determina fácilmente por un experto en la técnica, y generalmente es un cálculo empírico dependiente de la longitud de la sonda, temperatura de lavado y concentración de sal. En general, sondas más largas requieren temperaturas más altas para una hibridación apropiada, mientras que sondas pequeñas necesitan temperaturas menores. Generalmente la hibridación depende de la capacidad del DNA desnaturalizado para rehibridarse cuando están presentes cadenas complementarias en un ambiente por debajo de su temperatura de fusión. Cuanto mayor es el grado de homología deseada entre la sonda y la secuencia hibridable
- 25 más alta puede ser la temperatura relativa utilizada. Como resultado, las temperaturas

relativas más altas tenderán a hacer las reacciones más rigurosas, mientras temperaturas inferiores lo hacen menos. Para detalles adicionales y explicación de la rigurosidad de las reacciones de hibridación, véase Ausubel y col., *Current Protocols in Molecular Biology*, Wiley Interscience Publishers, 1995.

- 5 “Condiciones rigurosas” o “condiciones altamente rigurosas” como se definen en esta memoria descriptiva, se pueden identificar como aquellas que (1) emplean intensidad iónica baja y temperatura alta de lavado, por ejemplo, cloruro sódico 15 mM / citrato sódico 1,5 mM / dodecilsulfato sódico al 0,1% a 50°C; (2) emplean un agente desnaturizante durante la hibridación, tal como formamida, por ejemplo, formamida al
- 10 50% (v/v) con albúmina sérica bovina al 0,1% / ficoll al 0,1% / polivinilpirrolidona al 0,1% / tampón fosfato sódico 50 mM a pH 6,5 con cloruro sódico 750 mM / citrato sódico 75 mM a 42°C; o (3) emplean formamida al 50%, 5X SSC (cloruro sódico 750 mM, citrato sódico 75 mM), fosfato sódico 50 mM (pH 6,8), pirofosfato sódico al 0,1%, solución de 5X Denhardt, DNA de esperma de salmón sonificado (50 µg/ml), SDS al
- 15 0,1% y sulfato de dextrano al 10% a 42°C con lavados a 42°C en 0,2X SSC (cloruro sódico 30 mM / citrato sódico 3 mM) y formamida al 50% a 55°C seguido de un lavado de alta rigurosidad que consiste en SSC 0,1X que contiene EDTA a 55°C.

- “Condiciones moderadamente rigurosas” se pueden identificar como se describe en el documento de Sambrook y col., [*Molecular Cloning: A Laboratory Manual*. Nueva York: Cold Spring Harbor Press, (1989)], e incluye el uso de soluciones de lavado y condiciones de hibridación (por ejemplo, temperatura, intensidad iónica y % de SDS) menos rigurosas que las descritas anteriormente. Un ejemplo de condiciones moderadamente rigurosas es la incubación durante una noche a 37°C en una solución que comprende: formamida al 20%, 5X SSC (cloruro sódico 750 mM, citrato sódico
- 25 75 mM), fosfato sódico 50 mM a pH 7,6, solución de Denhardt 5X, sulfato de dextrano

al 10% y 20 mg/ml de DNA de esperma de salmón tundido y desnaturalizado, seguido de un lavado de los filtros una vez con SSC a aproximadamente 37 – 50°C. Los expertos en la técnica reconocerán como ajustar la temperatura, intensidad iónica, etc., según se necesite para acomodar los factores tales como la longitud de la sonda y similares.

- 5 “PCR” se refiere a la ampliamente conocida reacción en cadena con polimerasa que emplea una polimerasa de DNA térmicamente estable.

“Secuencia guía” se refiere a una secuencia de aminoácidos que se puede separar enzimática o químicamente para producir el polipéptido deseado de interés.

- 10 “Secuencia de señal de secreción” se refiere a una secuencia de aminoácidos que generalmente está presentes en la región N terminal de un polipéptido más grande que funciona para iniciar la asociación de ese polipéptido con los compartimientos de la membrana celular como el retículo endoplásmico y la secreción de ese polipéptido a través de la membrana plasmática.

- 15 Construcción de DNA que codifica las proteínas de fusión heterólogas de la presente invención:

- Las proteínas de albúmina de tipo natural e inmunoglobulina se pueden obtener de una diversidad de fuentes. Por ejemplo, estas proteínas se pueden obtener a partir de genotecas de cDNA preparadas a partir de tejido o células que expresan el mRNA de
20 interés a un nivel detectable. Las genotecas se pueden analizar con sondas diseñadas usando el DNA publicado o secuencia de proteínas para la proteína particular de interés. Por ejemplo, las regiones constantes de cadena ligera o pesada de inmunoglobulina se describen en Adams y col., (1980) Biochemistry 19: 2711 – 2719; Goughet y col., (1980) Biochemistry 19: 2702 – 2710; Dolby y col., (1980) Proc. Natl. Acad. Sci. Estados Unidos 77: 6027 – 6031; Rice y col., (1982) Proc. Natl. Acad. Sci. Estados
25

Unidos 79: 7862 – 7862; Falkner y col., (1982) *Nature* 298: 286 – 288; y Morrison y col (1984) *Ann. Rev. Immunol.* 2: 239 – 256. Algunas referencias que describen la proteína de albúmina y secuencias de DNA incluyen Meloun y col., (1975) *FEBS Letters* 58: 136; Behrens y col., (1975) *Fed. Proc.* 34: 591; Lawn y col. (1981) *Nucleic Acids Research* 9: 6102 – 6114; y Minghetti y col., (1986) *J. Biol. Chem.* 261: 6747.

El análisis de una genoteca de cDNA con la sonda seleccionada se puede llevar a cabo usando procedimientos estándar, tales como los descritos en Sambrook y col., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, NY (1989). Un medio alternativo para aislar un gen que codifica una proteína de albúmina o inmunoglobulina es usar la metodología PCR [Sambrook y col., más arriba; Dieffenbach y col., *PCR Primer: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, NY (1995)]. Los cebadores en PCR se pueden diseñar en función de secuencias publicadas.

Generalmente las secuencias de tipo natural que contienen todo el gen clonadas a partir de especies particulares pueden servir como molde para crear análogos, fragmentos y derivados que mantienen la capacidad de conferir una semivida en plasma más larga en el compuesto GLP – 1 que es parte de la proteína de fusión. Se prefiere que las porciones de Fc y albúmina de las proteínas de fusión heterólogas se deriven a partir de la secuencia humana nativa con el fin de reducir el riesgo de inmunogenicidad potencial de la proteína de fusión en seres humanos.

En particular, se prefiere que la porción de inmunoglobulina de una proteína de fusión comprendida por la presente invención contenga solamente un fragmento Fc de la inmunoglobulina. Dependiendo de si se desean efecto o funciones particulares y de las características estructurales en la proteína de fusión, un fragmento Fc puede contener las regiones flexibles junto con los dominios CH₂ y CH₃ o alguna otra combinación de

los mismos. Estos fragmentos Fc se pueden generar utilizando técnicas de PCR con cebadores diseñados para hidridarse a secuencias correspondientes a los extremos deseados del fragmento. De manera similar, si se desean los fragmentos de albúmina, los cebadores en PCR se pueden diseñar para que sean complementarios a las
5 secuencias internas de albúmina. Los cebadores en PCR también se pueden diseñar para crear sitios de enzimas de restricción para facilitar la clonación en los vectores de expresión.

El DNA que codifica los compuestos GLP – 1 de la presente invención se puede preparar mediante una diversidad de procedimientos diferentes que incluyen
10 procedimientos de clonación similares a los descritos anteriormente así como DNA sintetizado químicamente. La síntesis química puede ser atractiva dada la corta longitud del péptido codificado. Se ha publicado la secuencia de aminoácidos de GLP – 1 así como la secuencia del gen de preproglucagón. [López y col., (1983) Proc. Natl. Acad. Sci. , Estados Unidos 80: 5485 – 5489; Bell y col., (1983) Nature, 302: 716 – 718;
15 Heinrich, G. y col (1984) Endocrinol, 115: 2176 – 2181; Ghiglione, M. y col. (1984) Diabetologia 27: 599 - 600]. De este modo, se pueden diseñar cebadores para compuestos de GLP – 1 nativos de PCR y fragmentos de los mismos.

Después el gen que codifica una proteína de fusión se puede construir ligando el DNA que codifica un compuesto GLP – 1 en fase al DNA que codifica una proteína de
20 albúmina o de Fc. El gen que codifica el compuesto de GLP – 1 y el gen que codifica la proteína de albúmina o de Fc también se pueden unir en fase mediante DNA que codifica un engarce peptídico.

La función y estabilidad *in vivo* de las proteínas de fusión heterólogas de la presente invención se pueden optimizar mediante la adición de engarces de péptidos
25 pequeños para evitar potencialmente las interacciones de dominios no deseadas. Aunque

estos engarces pueden potencialmente ser de cualquier longitud y estar constituidos por cualquier combinación de aminoácidos, se prefiere que la longitud no sea mayor que la necesaria para evitar las interacciones no deseadas de dominios y/o optimizar la actividad y/o estabilidad biológica. En general, los engarces no deben contener

5 aminoácidos con cadenas laterales extremadamente voluminosas o aminoácidos apropiados para introducir estructuras secundarias significativas. Se prefiere que el engarce sea rico en serina – glicina y tenga menos de 30 aminoácidos de longitud. Se prefiere más que el engarce no tenga más de 20 aminoácidos de longitud. Incluso se prefiere más que el engarce no tenga más de 15 amino ácidos de longitud. Un engarce

10 preferido contiene repeticiones de la secuencia Gly – Gly – Gly – Gly – Ser. Se prefiere que haya entre 2 y 6 repeticiones de esta secuencia. Incluso se prefiere más que haya entre 3 y 4 repeticiones de esta secuencia.

El DNA que codifica el GLP – 1 de tipo natural, albúmina y polipéptidos de Fc y fragmentos de los mismos se pueden mutar tanto antes de la ligación como en el

15 contexto de un cDNA que codifica un proteína de fusión entera. Se conocen bien en la técnica una diversidad de técnicas de mutagénesis. Por ejemplo, un procedimiento de PCR mutagénico utiliza una extensión de superposición de cadenas para crear mutaciones de bases específicas con el propósito de cambiar una secuencia específica de aminoácidos en la proteína correspondiente. Esta mutagénesis por PCR requiere el uso

20 de cuatro cebadores, dos en la orientación delantera (cebadores A y C) y dos en la orientación inversa (cebadores B y D). Un gen mutado se amplifica a partir del molde de tipo natural en dos etapas diferentes.. La primera reacción amplifica el gen en mitades realizando una reacción de A a B y una reacción separada de C a D en la que los cebadores B y C dirigen el área del gen a mutar. Cuando se alinean estos cebadores

25 con al área diana, pueden contener disparidades en las bases dianas a cambiarse. Una

vez que las reacciones A a B y C a D se han completado, los productos de reacción se aíslan y se mezclan para uso como molde para la reacción A a D. Después esta reacción produce el producto completo mutado.

Una vez que se produce un gen que codifica la proteína de fusión entera se puede clonar en el vector adecuado de expresión. Las estrategias específicas que se pueden emplear para preparar las proteínas de fusión de GLP – 1 de la presente invención se describen en el ejemplo 1.

Procedimientos generales para expresar de forma recombinante las proteínas de fusión heterólogas de la presente invención:

Las células huésped se transfectan o se transforman con vectores de expresión o clonación descritos en esta memoria descriptiva para la producción de proteína de fusión heteróloga y se cultivan en medio nutritivo convencional modificado según sea apropiado para inducir promotores, seleccionar transformantes o amplificar los genes que codifican las secuencias deseadas. Las condiciones de cultivo, tales como medio, temperatura, pH y similares se pueden seleccionar por el experto en la técnica sin experimentación excesiva. En general, los principios, protocolos y técnicas prácticas para maximizar la productividad de cultivos celulares se pueden encontrar en *Mammalian Cell Biotechnology: A Practical Approach*, M. Butler, ed (IRL Press, 1991) y Sambrook y col., más arriba. Los expertos en la técnica conocen los procedimientos de transfección, por ejemplo, CaPO_4 y electroporación. Los aspectos generales de las transformaciones del sistema huésped de células de mamíferos se han descrito en la patente de Estados Unidos N° 4.399.216. Típicamente las transformaciones en las levaduras se llevan a cabo según el procedimiento de van Solingen y col., *J. Bact.* 130 (2): 946 – 7 (1977) y Hsiao y col., *Proc. Natl. Acad. Sci. Estados Unidos* 76 (8): 3829 –

33 (1979). Sin embargo, se pueden usar otros procedimientos para introducir DNA en células, tales como microinyección nuclear, electroporación, fusión de protoplastos bacterianos con células intactas o policationes, por ejemplo polibreno o poliomitina. Para las diversas técnicas de transformación de células de mamíferos véase Keown y col., *Methods in Enzimology* 185: 527 – 37 (1990) y Mansour y col., *Nature* 336 (6197): 348 – 52 (1988).

Las células huésped adecuadas para clonar o expresar el ácido nucleico (por ejemplo DNA) en los vectores de esta memoria descriptiva incluyen células procariotas, levaduras o células eucariotas superiores. Los procariotas adecuados incluyen pero no se limitan a eubacterias, tales como los organismos Gram - negativos o Gram - positivos, por ejemplo, Enterobacteriáceas tales como *E. coli*. Están públicamente disponibles diversas cepas de *E. coli*, tales como *E. coli* K12 cepa MM294 (ATCC 3 1.446); *E. coli* X1 776 (ATCC 3 1.537); *E. coli* cepa W3 110 (ATCC 27.325); y K5 772 (ATCC 53.635). Otras células huésped procarióticas adecuadas incluyen

15 Enterobacteriáceas tales como *Escherichia*, por ejemplo, *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, por ejemplo *Salmonella typhimurium*; *Serratia*, por ejemplo, *Serratia marcescans* y *Shigella*, así como *Bacilli* tales como *B. subtilis* y *B. licheniformis* (por ejemplo, *B. licheniformis* 4 1 P descrito en el documento DD226, 7 10, publicado el 12 de abril de 1989), *Pseudomonas* tales como *P.*

20 *aeruginosa* y *Streptomyces*. Estos ejemplos son ilustrativos en lugar de limitantes. La cepa W3110 es una célula huésped o huésped parental particularmente preferida porque es una cepa huésped común para fermentaciones de productos de DNA recombinantes. Preferiblemente, la célula hospedadora secreta cantidades mínimas de enzimas proteolíticas. Por ejemplo, la cepa W3 110 se puede modificar para efectuar una

25 mutación genética en los genes que codifican proteínas endógenas para la huésped,

275
269

ejemplos de tales células huésped incluyen *E. coli* W3110 cepa 1A2, que tiene el genotipo completo de *ronA*; *E. coli* W3110 cepa 9E4, que tiene el genotipo completo de *ton4 ptr3*, *E. coli* W3110 cepa 27C7 (ATCC 55, 244), que tiene el genotipo completo de *tonA, ptr3 phoA E15 (argF - lac) l69 degP ompT /can'*; *E. coli* W3110 cepa 40B4, que es la cepa 37D6 con una mutación de supresión de *degP* no resistente a kanamicina; y una cepa de *E. coli* que tiene una proteasa periplásmica mutante descrita en la patente de Estados Unidos N° 4.946.783 expedida el 7 de agosto de 1990. Alternativamente, son adecuados los procedimientos *in vivo* de clonación, por ejemplo, PCR u otras reacciones con polimerasas de ácidos nucleicos.

Además de los procariotas, microbios eucariotas tales como hongos filamentosos o levaduras son adecuados huéspedes de clonación o expresión para vectores de proteínas de fusión. *Saccharomyces cerevisiae* es un microorganismo huésped eucariótico inferior comúnmente usado. Otros eucariotas incluyen *Schizosaccharomyces pombe* [Beach y Nurse, *Nature* 290: 140 - 3 (1981); documento EP 139.383 publicado el 2 de mayo de 1995]; huéspedes *Muyveromyces* [patente de Estados Unidos N° 4.943.529; Fleer y col., *Bio / Technology* 9 (10): 968 - 75 (1991)] tales como por ejemplo *K. lactis* (MW98 - 8C, CBS683, CBS4574) [de Louvencourt y col., *J. Bacteriol.* 154 (2) : 737 - 42 (1983)]; *K. fragilis* (ATCC 12, 424), *K. bulgaricus* (ATCC 16.045), *K. wickerhamii* (ATCC 24.178), *K. waltii* (ATCC 56.500), *K. drosophilae* (ATCC 36.906) [Van den Berg y col., *Bio / Technology* 8 (2); 135 - 9 (1990)]; *K. thermotolerans* y *K. marxianus*; *Yarrowia* (documento 402.226); *Pichia pastoris* (documento EP 183.070) [Sreekrishna y col., *J. Basic Microbiol.* 28 (4). 265 - 278 (1988)]; *Candida Trichoderma reesia* (documento EP 244.234); *Neurospora crassa* [Case y col., *Proc. Natl. Acad. Sci. Estados Unidos* 76 (10): 5259 - 63 (1979)]; *Schwanniomyces* tal como *Schwanniomyces occidentalis* (documento EP 394.538

publicado el 31 de octubre de 1990), y hongos filamentosos tales como por ejemplo, *Neurospora*, *Penicillium*, *Tolipocladium* (documento WO 91/00357 publicado el 10 de enero de 1991) y huéspedes *Aspergillus* tales como *A. nidulans* [Ballance y col., *Biochem. Biophys. Res. Comm.* 112 (1): 284 – 9 (1983)]; Tilburn y col., *Gene* 26 (2 – 3): 205 – 21 (1983); Yelton y col., *Proc. Natl. Acad. Sci. Estados Unidos* 81 (5): 1470 – 4 (1984)] y *A. Niger* [Kelly y Hynes, documento EMBO J. 4 (2): 475 – 9 (1985)]. Levaduras metilotrópicas se seleccionan de los géneros constituidos por *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis* y *Rhodotoruia*. Una lista de especies específicas que son ejemplares de esta clase de levaduras se puede encontrar en

10 C. Antony, *The Biochemistry of Methylotrophs* 269 (1982).

Células huésped adecuadas para la expresión de las proteínas de fusión de presente invención se derivan de organismos multicelulares. Los ejemplos de células de invertebrados incluyen células de insectos tales como *Drosophila S2* y *Spodoptera sp*, *Spodoptera high5* así como células de plantas. Los ejemplos de líneas de células

15 huésped de mamíferos útiles incluyen células de ovario de hámster chino (abreviadamente CHO) y COS. Los ejemplos más específicos incluyen la línea CV1 de riñón de mono transformada por SV40 (COS – 7, ATCC CRL 1651); la línea de riñón embrionaria humana [células 293 ó 293 subclonadas para crecimiento en cultivos en suspensión, Graham y col., *J. Gen Virol.*, 36 (1): 59 – 74 (1977)]; células de ovario de

20 hámster chino / -DHFR [CHO, Urlaub y Chasin, *Proc. Natl. Acad. Sci. Estados Unidos* 77 (7): 4216 – 20 (1980)]; células de Sertoli de ratón [TM4, Mather, *Biol. Reprod.* 23 (1): 243 – 52 (1980)]; células de pulmón humano (W138. ATCC CCL 75); células de hígado humano (Hep G2, HB 8065) y tumor de mama de ratón (MMT 060562, ATCC CCL 51). La selección de la célula huésped se considera que está dentro de la

25 experiencia en la técnica.

Las proteínas de fusión de la presente invención se pueden producir directamente por recombinación, o en forma de una proteína que tiene una secuencia de señal u otras secuencias adicionales que crean un sitio de escisión específico en el extremo N terminal de la proteína de fusión madura. En general, la secuencia de señal
5 puede ser un componente del vector, o puede ser parte del DNA que codifica la proteína de fusión que se inserta en el vector. La secuencia de señal puede ser una secuencia de señal procariótica seleccionada, por ejemplo, del grupo de la fosfatasa alcalina, penicilinasa, lpp, o las guías de enterotoxina II termoestables. Para la secreción de la levadura la secuencia de señal puede ser, por ejemplo, la guía de la invertasa de la
10 levadura, guía de factor alfa (que incluye las guías del factor cc de *Saccharomyces* y las guías del factor cc de *Kluyveromyces*, las últimas están descritas en la patente de Estados Unidos N° 5.010.182) o la guía de fosfatasa ácida, la guía de glucoamilasa de *C. albicans* (documento EP 362.179), o la señal descrita en el documento WO 90/13646. En la expresión de células de mamíferos, se pueden usar las secuencias de
15 señal de mamíferos para dirigir la secreción de la proteína, tales como las secuencias de señal de los polipéptidos secretados de la misma o especies relativas así como guías secretoras virales.

Ambos vectores de expresión y clonación contienen una secuencia de ácidos nucleicos que permite que el vector se replique en una o más células huésped
20 seleccionadas. Tales secuencias se conocen bien para una diversidad de bacterias, levaduras y virus. El origen de la replicación a partir del plásmido pBR322 es adecuado para la mayoría de las bacterias Gram – negativas, el origen a partir del plásmido 2u es adecuado para levaduras, y diversos orígenes virales (SV40, polioma, adenovirus, VSV o BPV) son útiles para vectores de clonación en células de mamíferos.

25 Los vectores de expresión y clonación típicamente contendrán un gen de

selección, también denominado marcador seleccionable. Los genes de selección típicos codifican proteínas que (a) confieren resistencia a antibióticos u otras toxinas, por ejemplo, ampicilina, neomicina, metotrexato o tetraciclina, (b) complementan deficiencias autotróficas o (c) suministran nutrientes críticos no disponibles a partir de medios complejos, por ejemplo, los genes que codifican la D – alaninarracemasa de Bacilos.

Un ejemplo de marcadores seleccionables adecuados de células de mamíferos son los que permiten la identificación de la células competentes para recoger el ácido nucleico que codifica las proteínas de fusión, tales como DHFR o timidinaquinasa.

10 Cuando se emplea DHFR de tipo natural una célula huésped adecuada es la línea celular CHO deficiente en actividad DHFR, preparada y propagada como se describe en [Urlaub y Chasin, *Proc. Natl. Acad. Sci. Estados Unidos*, 77 (7): 4216 – 20 (1980)]. Un gen de selección adecuado para uso en levaduras es el gen *trp1* presente en el plásmido YRp7 de levaduras [Stinchcomb y col., *Nature* 282 (5734): 39 – 43 (1979); Kingsman

15 y col., *Gene* 7 (2): 141 – 52 (1979); Tschumper y col., *Gene* 10 (2): 157 – 66 (1980)]. El gen *trp1* proporciona un marcador de selección para una cepa mutante de levaduras que carece de la capacidad de crecer en triptófano, por ejemplo, ATCC N° 44076 o PEPC1 [Jones *Genetics* 85: 23 – 33 (1977)].

Los vectores de expresión y clonación normalmente contienen un promotor operativamente unido a la secuencia de ácido nucleico que codifica la proteína de fusión para dirigir la síntesis de mRNA. Se conocen bien los promotores reconocidos por una diversidad de células huésped potenciales. Los promotores adecuados para uso con huésped procarióticos incluyen los sistemas promotores de la β -lactamasa y lactosa [Chang y col., *Nature* 275 (5681): 617 – 24 (1978); Goeddel y col., *Nature* 281 (5732):

25 544 – 8 (1979)], fosfatasa alcalina, un sistema promotor hacia arriba de triptófano

[Goeddel, *Nucleic Acids Res.* 8 (18): 4057 – 74 (1980); documento EP 36.776 publicado el 30 de septiembre de 1981] y promotores híbridos tales como el promotor tat [deBoer y col., *Proc. Natl. Acad. Sci. Estados Unidos* 80 (1): 21 – 5 (1983)]. Los promotores para uso en sistemas bacterianos también contendrán una secuencia Shine – Dalgarno
5 (abreviadamente S. D.) unida operativamente al DNA que codifica la proteína de fusión.

Los ejemplos de las secuencias promotoras adecuadas para uso con huésped de levaduras incluyen los promotores de 3 – fosfogliceratoquinasa [Hitzeman y col., *J. Biol. Chem.* 255 (24): 12073 – 80 (1980)] u otras enzimas glicolíticas [Hess, y col., *J. Adv Enzyme. Reg.* 7: 149 (1968); Holanda, *Biochemistry* 17 (23): 4900 – 7 (1978)], tales
10 como enolasa, gliceraldehído – 3 – fosfatodeshidrogenasa, hexoquinasa, piruvatodecarboxilasa, fosfofructoquinasa, glucosa – 6 – fosfatoisomerasa, 3– fosfogliceratomutasa, piruvatoquinasa, triosafofatoisomerasa, fosfoglucosaisomerasa y glucoquinasa.

Otros promotores de levaduras, que son promotores inducibles y que tienen la
15 ventaja adicional de transcripción controlada por las condiciones de crecimiento son las regiones de los promotores de la alcoholdehidrogenasa 2, isocitocromo C, fosfatasa ácida, enzimas catalíticas asociadas con el metabolismo del nitrógeno, metalotioneína, gliceraldehído – 3- fosfatodeshidrogenasa y enzimas responsables de la utilización de la maltosa y galactosa. Los vectores y promotores adecuados para el uso en la expresión
20 de levaduras son posteriormente descritos en el documento EP 73.657. La transcripción de mRNA que codifica proteínas de fusión a partir de vectores en células huésped de mamíferos se pueden controlar, por ejemplo, mediante promotores obtenidos a partir de los genomas de virus tales como virus de polioma, virus de la viruela aviar, adenovirus (tal como adenovirus 2), virus de paliloma bovino, virus del sarcoma aviar,
25 citomegalovirus, un retrovirus, virus de hepatitis B y virus 40 de los simios

(abreviadamente SV40), de promotores de mamíferos heterólogos, por ejemplo, el promotor de actina o un promotor de inmunoglobulina y a partir de promotores de choque térmico, siempre que dichos promotores sean compatibles con los sistemas de células huésped.

- 5 La transcripción de un polinucleótido que codifica una proteína de fusión mediante eucariotas superiores se puede incrementar insertando una secuencia de potenciación en el vector. Los potenciadores son elementos de DNA que actúan en cis, usualmente entre aproximadamente 10 y 300 bp (pares de bases), que actúan en un promotor para aumentar su transcripción. Muchas secuencias de potenciadores se
- 10 conocen ahora a partir de genes de mamíferos (globina, elastasa, albúmina, cetoproteína a e insulina). Sin embargo, típicamente, se usará un potenciador de un virus de células eucarióticas. Los ejemplos incluyen el potenciador SV40 en la zona tardía del origen de la replicación (bp 100 – 270), el potenciador del promotor temprano de citomegalovirus, el potenciador de polioma en la zona tardía del origen de la replicación y potenciadores
- 15 de adenovirus. El potenciador se puede cortar y empalmar con el vector en la posición 5' ó 3' a la secuencia que codifica la proteína de fusión pero se localiza preferiblemente en una posición 5' del promotor.

- Los vectores de expresión usados en las células huésped eucarióticas (levaduras, hongos, insectos, plantas, animales, seres humanos o células nucleadas de otros
- 20 organismos multicelulares) también contendrán secuencias necesarias para la terminación de la transcripción y para estabilizar el mRNA. Tales secuencias están comúnmente disponibles de las regiones 5' y ocasionalmente 3' no traducidas de DNA o cDNA eucarióticos o virales. Estas regiones contienen segmentos de nucleótidos transcritos como fragmentos poliadenilados en la porción no traducida del mRNA que
- 25 codifica la proteína de fusión.

Se pueden recuperar diversas formas de una proteína de fusión a partir del medio de cultivo o a partir de lisados de células huésped. Si está unida a la membrana, se puede liberar de la membrana utilizando una solución de detergente adecuada (por ejemplo, Triton - X 100) o mediante escisión enzimática. Las células empleadas en la
5 expresión de una proteína de fusión se pueden romper mediante diversos medios físicos o químicos, tales como en ciclo congelación - descongelación, sonicación, rupturamecánica o agentes de lisis celular.

Purificación de las proteínas de fusión heterólogas de la presente invención:

10 Una vez que las proteínas de fusión heterólogas de la presente invención se han expresado en la célula huésped adecuada, los análogos se pueden aislar y purificar. Los siguientes procedimientos son ejemplares de procedimientos de purificación adecuados: fraccionamiento sobre carboximetilcelulosa; filtración en gel como Sephadex G - 75; resina de intercambio aniónico tal como DEAE o mono - Q; intercambio catiónico tal
15 como CM o Mono - S; sefarosa proteína A para separar contaminantes tales como IgG; columnas de metal quelante para unir formas marcadas de epitopo del polipéptido; HPLC en fase inversa; cromatografía de enfoque; gel de sílice; precipitación con etanol; precipitación con sulfato amónico.

Se pueden emplear diversos procedimientos de purificación de proteínas y tales
20 procedimientos se conocen en la técnica y se describen, por ejemplo, en Deutscher, *Methods in Enzymology* 182: 83 - 9 (1990) y Scopes, *Protein Purification: Principles and Practice*, Springer - Verlag, NY (1982). La (s) etapa (s) de purificación seleccionada (s) dependerá (n) de la naturaleza del procedimiento de producción utilizados y la proteína de fusión producida. Por ejemplo, las proteínas de fusión que
25 comprenden un fragmento Fc se pueden purificar eficazmente usando una matriz de

afinidad de proteína A o proteína G. Los tampones de pH bajo o alto se pueden usar para eluir la proteína de fusión de la matriz de afinidad. Las condiciones de elución suaves ayudarán a prevenir la desnaturalización irreversible de la proteína de fusión. También se pueden usar los tampones que contienen imidazol. El ejemplo 3 describe
5 algunos protocolos de purificación exitosos para las proteínas de fusión de la presente invención.

Caracterización de las proteínas de fusión heterólogas de la presente invención:

Existen numerosos procedimientos para caracterizar las proteínas de fusión de la
10 presente invención. Algunos de estos procedimientos incluyen: electroforesis en dodecil sulfato sódico - gel de poliacrilamida (abreviadamente SDS - PAGE) unida a procedimientos de tinción de proteínas o inmunotransferencia que usa anticuerpos anti - IgG o anti - HSA. Otros procedimientos incluyen espectrometría de masas con ionización por desorción mediante láser asistida por matriz (abreviadamente
15 MALDI - MS), cromatografía líquida / espectrometría de masas, electroenfoque isoelectrico, intercambio aniónico analítico, cromatoenfoque y dicroísmo circular por nombrar unos pocos. Un número representativo de proteínas de fusión heterólogas se caracterizaron usando SDS - PAGE unido con inmunotransferencia así como espectrometría de masas (véanse los ejemplos 4 y 5 y las figuras 3 y 4).

20 Por ejemplo la tabla 3 (véase el ejemplo 5) ilustra la masa molecular calculada de un número representativo de proteínas de fusión así como la masa determinada por espectrometría de masas. Además, las figuras 3 y 4 ilustran los pesos moleculares de un número representativo de proteínas de fusión determinados por SDS PAGE. Todas las proteínas de fusión heterólogas ensayadas se expresaron y se secretaron
25 transitoriamente. Además, la secuencia de señal Igk se escindió para producir proteínas

con el extremo N terminal correcto.

Además, la tabla 3 muestra que en algunos casos la masa determinada por espectrometría de masas es mayor que la esperada. Esto es el resultado de la glicosilación de la porción Fc y la ampliación del extremo C terminal. La digestión
5 enzimática de las proteínas de fusión seguida de HPLC en fase inversa y espectrometría de masas pueden identificar las fracciones de péptidos que contienen restos de azúcar. Después estas fracciones de aminoácidos se pueden secuenciar por el aminoácido del extremo N terminal para identificar la posición de glicosilación potencial. Por ejemplo, la caracterización de exendina - 4 - Fc (SEC ID N° 29) muestra que la serina en la
10 posición 39 y la treonina en la posición 50 están glicosadas en el enlace en O y la asparagina en la posición 122 está glicosada en el enlace en N.

También se ensayaron un número representativo de proteínas de fusión de GLP - 1 para determinar su actividad. Existen numerosos procedimientos para detectar la actividad GLP - 1 *in vitro* e *in vivo* (véanse los ejemplos 6, 7, 8 y 9). La tabla 4
15 (ejemplo 6) ilustra la actividad del receptor de GLP - 1 asociada con diversas fusiones de GLP - 1. Los números son relativos a la actividad asociada con Val⁸ - GLP - 1 (7 - 37) OH. Todas las proteínas de fusión ensayadas tenían actividad de receptor de GLP - 1. Un nivel bajo de actividad *in vitro* no es necesariamente indicativo de un efecto débil *in vivo*. Debido al incremento sustancial de la semivida de estas proteínas de fusión, una
20 actividad débil *in vitro* generalmente no predice una débil actividad *in vivo*. La figura 7 y el ejemplo 7 ilustran la semivida prolongada asociada con las proteínas de fusión de la presente invención. Por ejemplo, Val⁸ - GLP-1 - Fc tiene una semivida de aproximadamente 45 horas en monos, Val⁸ - GLP-1 - HSA tenía una semivida de aproximadamente 87 horas en monos, Gly⁸ - Glu²² - GLP - 1 - CEx - engarce - IgG1
25 tenía una semivida después de la administración IV intravenosa de aproximadamente 55

horas en perros y Gly⁸ - Glu²² - GLP - 1 - CEx - engarce - IgG1 tenía una semivida después de administración SC (subcutánea) de aproximadamente 38 horas en perros.

Composiciones de la invención

5 La estabilidad física es también una característica esencial para la formulación de las proteínas terapéuticas. Los compuestos GLP - 1 han sido especialmente difíciles de fabricar y formular debido a los cambios estructurales que se presentan durante la elaboración. Por ejemplo, algunos compuestos GLP - 1 tienen una tendencia general a agregarse. Además, se ha demostrado que algunos compuestos GLP - 1 pasan de una
10 forma hélice α soluble y activa a una forma de hoja β insoluble y potencialmente inactiva. La fusión de los compuestos GLP - 1 a proteínas mayores tales como la región Fc de una IgG o albúmina no solamente actúa para aumentar la semivida del compuesto de GLP - 1, sino que también contribuye a la estabilidad física y conformacional de compuesto GLP - 1. Por ejemplo, Val⁸ - GLP - 1 - engarce - HSA en PBS es estable a
15 37°C hasta aproximadamente 30 días.

Las proteínas de fusión heterólogas de la presente invención se pueden formular con uno o más excipientes. Las proteínas de fusión activas de la presente invención se pueden combinar con un tampón farmacéuticamente aceptable y el pH se ajusta para proporcionar una estabilidad aceptable, y un pH aceptable para la administración tal
20 como administración parenteral.

Opcionalmente, se pueden añadir uno o más agentes antimicrobianos farmacéuticamente aceptables. Los agentes preferidos farmacéuticamente aceptables son metacresol y fenol. Se pueden añadir una o más sales farmacéuticamente aceptables para ajustar la intensidad iónica o tonicidad. Se pueden añadir uno o más excipientes
25 para ajustar adicionalmente la isotonicidad de la formulación. La glicerina es un

277
280

ejemplo de un excipiente de ajuste de isotonicidad. Farmacéuticamente aceptable significa adecuado para administración a un ser humano u otro animal y por lo tanto, no contiene elementos tóxicos o contaminantes indeseables y no interfiere con la actividad de los compuestos activos de esta memoria descriptiva.

- 5 Una forma de sal farmacéuticamente aceptable de las proteínas de fusión heterólogas de la presente invención se pueden usar en la presente invención. Los ácidos comúnmente empleados para formar sales de adición de ácido son ácidos inorgánicos tales como ácido clorhídrico, ácido bromhídrico, ácido yodhídrico, ácido sulfúrico, ácido fosfórico y similares, y ácidos orgánicos tales como ácido *p*-
10 toluensulfónico, ácido metanosulfónico, ácido oxálico, ácido *p*-bromofenilsulfónico, ácido carbónico, ácido succínico, ácido cítrico, ácido benzoico, ácido acético y similares.

Las sales de adición de ácido preferidas son las formadas con ácidos minerales tales como ácido clorhídrico y ácido bromhídrico.

- 15 Las sales de adición de base incluyen las derivadas de bases inorgánicas, tales como hidróxidos de amonio o de metal alcalino o alcalinotérreo, carbonatos, bicarbonatos y similares. Tales bases útiles para preparar las sales de esta invención de este modo incluyen hidróxido sódico, hidróxido potásico, hidróxido amónico, carbonato potásico y similares.

20

Administración de composiciones:

- La administración puede ser mediante cualquier vía que se sepa que es eficaz según el médico experto habitual. Uno de tales procedimientos es el periférico, vía parental. Comúnmente en la bibliografía médica se entiende la administración
25 parenteral como la inyección de una forma de dosificación en el cuerpo mediante una

jeringa estéril o algún otro dispositivo mecánico tal como una bomba de infusión. Las vías parenterales periféricas pueden incluir las vías de administración intravenosa, intramuscular, subcutánea e intraperitoneal.

Las proteínas de fusión heterólogas de la presente invención pueden ser también susceptibles de administración por vía oral, rectal, nasal o respiratoria inferior, que son vías no parenterales. De estas vías no parenterales, se prefieren la vía respiratoria inferior y la vía oral.

Las proteínas de fusión de la presente invención se pueden usar para tratar una amplia diversidad de enfermedades y afecciones. Las proteínas de fusión de la presente invención ejercen principalmente sus efectos biológicos actuando en un receptor denominado el "receptor de GLP - 1". De esta manera, los sujetos con enfermedades y/o afecciones que responden favorablemente a la estimulación del receptor de GLP - 1 o a la administración de los compuestos de GLP - 1 pueden tratarse con las proteínas de fusión de GLP - 1 de la presente invención. Se dice que estos sujetos "están en necesidad de tratamiento con compuestos de GLP - 1" o "en necesidad de estimulación del receptor de GLP - 1". Se incluyen los sujetos con diabetes no insulino dependiente, diabetes insulino dependiente, accidente cerebro vascular (véase el documento WO 00/16797), infarto de miocardio (véase documento WO 98/08531), obesidad (véase el documento WO 98/19698), cambios catabólicos después de cirugía (véase la patente de Estados Unidos N° 6.006.753), dispepsia funcional y síndrome de intestino irritable (véase el documento WO 99/64060). También se incluyen los sujetos que requieren tratamiento profiláctico con un compuesto GLP - 1, por ejemplo, sujetos con riesgo de desarrollar diabetes no insulino dependiente (véase el documento WO 00/07617). Los sujetos con tolerancia alterada a la glucosa o alteración de la glucosa en ayunas, sujetos cuyo peso corporal es aproximadamente un 25% superior al peso corporal normal para

la altura del sujeto y estructura corporal, los sujetos con pancreatectomía parcial, sujetos que tienen uno o los dos padres con diabetes no insulino dependiente, sujetos que han tenido diabetes gestacional y sujetos que han tenido pancreatitis aguda o crónica están en riesgo de desarrollar diabetes no insulino dependiente.

- 5 Una "cantidad eficaz" de un compuesto GLP – 1 es la cantidad que da como resultado un efecto terapéutico y/o profiláctico deseado sin causar efectos secundarios inaceptables cuando se administra a un sujeto en necesidad de estimulación del receptor de GLP – 1. Un "efecto terapéutico deseado" incluye uno o más de lo siguiente: 1) una
- 10 mejoría de (1) los síntoma (s) asociada con la enfermedad o afección; 2) un retraso en el comienzo de los síntomas asociado con la enfermedad o afección; 3) un incremento de longevidad comparado con la ausencia del tratamiento; y 4) mayor calidad de vida comparada con la ausencia del tratamiento. Por ejemplo, una "cantidad eficaz" de un compuesto GLP – 1 para el tratamiento de la diabetes es la cantidad daría como
- 15 resultado en un mayor control de la concentración de glucosa en sangre que en ausencia de tratamiento, con lo cual daría como resultado un retraso del comienzo de las complicaciones diabéticas tales como retinopatía, neuropatía o enfermedad de riñón. Una "cantidad eficaz" de un compuesto GLP – 1 para la prevención de diabetes es la
- 20 cantidad que demoraría, comparada con la ausencia de tratamiento, el comienzo de los niveles elevados de glucosa en sangre que requiere tratamiento con fármacos antihipoglucémicos tales como sulfonilureas, tiazolidindionas, insulina y/o bisguanidinas.

La dosis eficaz de proteína de fusión para normalizar los niveles de glucosa en sangre de un paciente dependerá de un número de factores, entre los cuales se incluyen, sin limitación, el sexo del sujeto, su peso y edad, la gravedad de la incapacidad para

25 regular los niveles de glucosa en sangre, la vía de administración y la biodisponibilidad,

el perfil farmacocinético de la proteína de fusión, la potencia (% de ingrediente activo) y la formulación.

La presente invención comprende los compuestos de GLP - 1 que han mejorado las propiedades bioquímicas y biofísicas por medio de la fusión a una proteína de albúmina, un fragmento de albúmina, un análogo de albúmina, una proteína Fc, un
5 fragmento Fc, o un análogo de Fc. Estas proteínas heterólogas se pueden expresar con éxito en células huésped, actividades de señalización mantenidas asociadas con la activación del receptor de GLP - 1, y tienen semividas prolongadas.

Los siguientes ejemplos se presentan para describir adicionalmente la invención.

10 El alcance de la presente invención no está construida de manera que esté solamente constituida por los siguientes ejemplos. Los expertos en la técnica reconocerán que los reactivos particulares, equipos y procedimientos descritos son meramente ilustrativos y no tienen la intención de limitar la presente invención de ninguna manera.

15 Ejemplo 1: Construcción de DNA que codifica proteínas de fusión heterólogas

Ejemplo 1a Construcción de DNA que codifica Val⁸ - GLP - 1 - (7 - 37) - Fc: Se aisló una porción de Fc de IgG1 humana de una genoteca de cDNA y contiene la región flexible completa y los dominios de CH₂ y CH₃. Un fragmento que contenía 696 pares de bases de esta porción Fc de IgG1 humana se subclonó en las posiciones *NheI* y
20 *Eco47III* del vector de expresión de mamíferos pJB02 para crear pJB02 / Fc (véase figura 5). Se generó DNA que codifica la secuencia de señal de secreción de Igκ condensada con Val⁸ - GLP - 1 (7 - 37) mediante hibridación *in vitro* de cuatro oligonucleótidos de superposición y complementarios:

5' - CTAGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
CCAGGTTCCACTGGTGACCAGTG - 3' [SEC ID N° 12]

5' - GAGGGCACCTTCACCTCCGACGTGTCTCTCTATCTGGAGGGCCAGGCCGCAAGGA
GTTTCATCGCCTGGCTGGTGAAGGGAAGAGGC - 3' SEC ID N° 13

5' - TGAAGGTGCCCTCCACGTGGTCAACAGTGGAACCTGGAACCCAGAGCAGCAGTA
CCCATAGCAGGAGTGTGTCTGTCTCCATGGTGG - 3' SEC ID N° 14

5' - GCCTCTTCCCTTCACCAGCCAGGCGATGAACTCCTTGGCGGCCTGGCCCTCCAGA
TAGGAGGACACGTCCGAGG - 3' SEC ID N° 15

La reacción de hibridación se llevó a cabo utilizando cantidades equivalentes de cada oligonucleótido (1 pm/μl de concentración final para cada oligo). La mezcla de oligonucleótidos se calentó durante 5 minutos a 100°C en tampón de ligación (Tris -
5 HCl 50 mM, pH 7,5, MgCl₂ 10 mM, DTT 10 mM, ATP 1mM , albúmina sérica bovina 25 μg/ml) y después se enfrió durante al menos 2 horas a 30°C.

El producto de hibridación resultante se ligó durante 2 horas a temperatura ambiente o durante una noche a 16°C a la estructura del vector pJB02 /Fc que se había digerido con *NheI* y *Eco47III*. Los productos de ligación se utilizaron para transformar
10 células XL 1 Blue competentes (Stratagene). Se analizaron los plásmidos recombinantes para detectar la presencia de inserciones que codifican péptidos mediante la digestión de clones con *NcoI* (que codifica la secuencia Kozak y la Met primera del péptido de señal) y se secuenció. El plásmido de expresión resultante usado para ensayos de transfección se denominó pJB02 - V8 - GLP -1 - Fc (figura 5).

15

Ejemplo 1 b Construcción de DNA que codifica Val⁸ - GLP - 1 (7 - 37) - HSA: Se

compró el plásmido HSA/pcDNA3.1GS en Invitrogen (n° de catálogo H – M12523M – pcDNA3.1/GS) y se utilizó como molde para aislar el cDNA que codifica albúmina sérica humana (abreviadamente HSA). El cDNA de HSA se preparó usando PCR en la que el DNA que codifica la secuencia guía así como el pro – péptido de seis aminoácidos se retiraron del extremo 5'. Además, se añadieron codones de parada directamente en el extremo 3' de la secuencia que codifica HSA. Finalmente, las posiciones de la enzima de restricción se modificaron genéticamente en los extremos 5' y 3' para facilitar la clonación. La secuencia DNA de HSA presente en el vector original comprada en Invitrogen contenía un cambio en una base en la región 3' del gen (posición 667) en comparación con la secuencia humana nativa. Este cambio daría como resultado un codón para Asn en lugar de Asp. De este modo, usando el procedimiento de mutagénesis mediante PCR de superposición de cadena descrito anteriormente, el codón se cambió para codificar para Asp en esta posición. El DNA resultante que codifica HSA se clonó en las posiciones *NheI* e *Hind III* de pJB02 para crear pJB02 – HSA (figura 6).

La secuencia guía de Igκ condensada con la secuencia Val⁸ – GLP – 1 (7 – 37) se generó como se describe en el ejemplo 1a. Este DNA se ligó en las posiciones *NheI* y *FspI* de pJB02 – HSA para crear pJB02 - Val⁸ – GLP – 1 – HSA.

Ejemplo 1 c Construcción de DNA que codifica Val⁸ – GLP – 1 (7 – 37) – engarce – HSA:

El vector pJB02 – HSA se preparó como se ha descrito en ejemplo 1b. El DNA que codifica la secuencia del engarce [GGGGS]₃ se ligó en fase al extremo 5' del DNA que codifica HSA para crear pJB02 – engarce – HSA (figura 7). El DNA que codifica la secuencia guía de Igκ condensada con la secuencia Val⁸ – GLP – 1 (7 – 37) y la parte 5' de la secuencia del engarce se generó como se describe en el ejemplo 1a. Este DNA se

ligó en las posiciones *NheI* y *BspEI* de pJB02 para crear pJB02 - Val⁸ - GLP - 1 -
engarce - HSA.

Ejemplo 1d Construcción de DNA que codifica exendina - 4 - Fc.

- 5 El plásmido pJB02 / Fc se preparó como se ha descrito en el ejemplo 1a. El DNA que codifica la secuencia de señal de Igκ condensada a la exendina - 4 se generó mediante hibridación *in vitro* de los siguientes oligonucleótidos de superposición y complementarios:

5' - CTAGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTG
GGTTCAGGTTCCACCGGTCAC - 3' [SEC ID N° 16]

5' - GGAGAGGGAACCTTCACCAGCGACCTGAGCAAGCAGATGGAGGAGGAGGCCGT
GAGACTG - 3' [SEC ID N° 17]

5' - TTCATCGAGTGGCTGAAGAACGGAGGACCAAGCAGCGGAGCCCCCTCCTCCT
AGC - 3' [SEC ID N° 18]

5' - GAACCTGGAACCCAGAGCAGCAGTACCCATAGCAGGAGTGTGTCTGTCTCCA
TGGTGG - 3' [SEC ID N° 19]

5' - CTCCTCCTCCATCTGCTTGCTCAGGTCGCTGGTGAAGGTTCCCTCTCCGTGA
CCGGTG - 3' [SEC ID N° 20]

10

5' - GCTAGGAGGAGGGGCTCCGCTGCTTGGTCCTCCGTTCTTCAGCCACTCGAT
GAACAGTCTCACGGC - 3' [SEC ID N° 21]

La reacción de hibridación se llevó a cabo como se ha descrito en el ejemplo 1a. El producto hibridado se ligó al vector pJB02 que se había digerido con *NheI* y *Eco47III* como se ha descrito en el ejemplo 1a para crear pJB02 – exendina – 4 – Fc.

5

Ejemplo 1e Construcción de DNA que codifica exendina – 4 – HSA:

El plásmido pJB02 / HSA se preparó como se ha descrito en el ejemplo 1b. El DNA que codifica la secuencia de señal de Igk condensada a la exendina – 4 se generó mediante hibridación *in vitro* de los mismos oligonucleótidos de superposición y complementarios descritos en el ejemplo 1d. Las reacciones de hibridación también se llevaron a cabo como se ha descrito anteriormente. El DNA se clonó en posiciones únicas *NheI* y *FspI* en pJB02 – HSA para crear pJB02 – exendina – 4 – HSA.

15 Ejemplo 1f Construcción de DNA que codifica exendina – 4 – engarce – HSA:

El plásmido pJB02 – engarce - HSA se construyó como se ha descrito en el ejemplo 1c. El DNA que codifica la secuencia de señal de Igk condensada con la exendina – 4 y la parte 5' de la secuencia del engarce se generó como en el ejemplo 1d. Este DNA se clonó en posiciones únicas *NheI* y *BspEI* en pJB02 – engarce - HSA para crear pJB02 – exendina – 4 – engarce - HSA.

20

Ejemplo 1g Construcción de DNA que codifica Val⁸ – GLP – 1 / C – Ex - Fc:

El plásmido pJB02 – exendina 4 - Fc se preparó como se ha descrito en el ejemplo 1d. El DNA que codifica exendina – 4 se escindió del vector con *AgeI* y *Eco47III*. El DNA que codifica Val⁸ – GLP – 1 / C – Ex se generó mediante hibridación

25

in vitro de los siguiente oligonucleótidos de superposición y complementarios:

5' -CCGGTCACGTGGAGGGCACCTTCACCTCCGACGTGTCCTCCTATCTGGA
GGGCCAGGCCGCCA - 3' [SEC ID N° 22]

5' - AGGAATTCATCGCCTGGCTGGTGAAGGGCCGGGGCAGCAGCGG
AGCCCCCTCCTCTAGC - 3' [SEC ID N° 23]

5' - CTCCAGATAGGAGGACACGTCGGAGGTGAAGGTGCCCTCCAC
GTGA - 3' [SEC ID N° 24]

5' - GCTAGGAGGAGGGGCTCCGCTGCTGCCCCGGCCCTTCACCAGCCAGGCGA
TGAATTCCTTGCGGCCTGGCC - 3' [SEC ID N° 25]

- 5 La reacción de hibridación se llevó a cabo como se ha descrito en el ejemplo 1a. El producto hibridado se ligó en lugar de exendina - 4 en el vector de expresión pJB02 - exendina - 4 - Fc para crear pJB02 - Val⁸ - GLP - 1 / C - Ex - Fc.

Ejemplo 1h Construcción de DNA que codifica Val⁸ - Glu²² - GLP - 1 - Fc:

- 10 El plásmido pJB02 - exendina 4 - Fc se preparó como se ha descrito en el ejemplo 1d. El DNA que codifica exendina - 4 se escindió del vector con *AgeI* y *Eco47III*. El DNA que codifica Val⁸ - Glu²² - GLP - 1 se generó mediante hibridación *in vitro* de los siguiente oligonucleótidos de superposición y complementarios:

5' -CCGGTCACGTGGAGGGCACCTTCACCTCCGACGTGTCCTCCTATCTCGA
GGAGCAGGCCGCCA - 3' [SEC ID N° 26]

5' - AGGAGTTCATCGCCTGGCTGGTGAAGGGCCGGGGC - 3' [SEC ID N° 27]

5' - GCCCCGGCCCTTCACCAGCCAGGCGATGAACTCCTTGGCGGCC
TGCTC - 3' [SEC ID N° 28]

5' - CTCGAGATAGGAGGACACGTCGGAGGTGAAGGTGCCCT
CCACGTGA - 3' [SEC ID N° 29]

La reacción de hibridación se llevó a cabo como se ha descrito en el ejemplo 1a. El producto hibridado se ligó en lugar de exendina - 4 en el vector de expresión pJB02 - exendina - 4 - Fc para crear pJB02 - Val⁸ - Glu²² - GLP - 1 - Fc.

5

Ejemplo 1i Construcción de DNA que codifica Val⁸ - Glu²² - GLP - 1 / C - Ex - Fc:

El plásmido pJB02 - exendina 4 - Fc se preparó como se ha descrito en el ejemplo 1d. El DNA que codifica exendina - 4 se escindió del vector con *AgeI* y *Eco47III*. El DNA que codifica Val⁸ - Glu²² - GLP - 1 / C - Ex se generó mediante
10 hibridación *in vitro* de los siguiente oligonucleótidos de superposición y complementarios:

5' - CCGGTCACGTGGAGGGCACCTTCACCTCCGACGTGTCCTCCTATCTCGA
GGAGCAGGCCGCCA - 3' [SEC ID N° 30]

5' - AGGAATTCATCGCCTGGCTGGTGAAGGGCCGGGGCAGCAGCGGA
GCCCCCTCCTCCTAGC - 3' [SEC ID N° 31]

5' - CTCGAGATAGGAGGACACGTCGGAGGTGAAGGTGCCC
TCCACGTGA - 3' [SEC ID N° 32]

5' - GCTAGGAGGAGGGGCTCCGCTGCTGCCCCGGCCCTTCACCAGCCAGGCCGA
TGAATTCCTTGGCGGCCTGCTC - 3' [SEC ID N° 33]

La reacción de hibridación se llevó a cabo como se ha descrito en el ejemplo 1a. El producto hibridado se ligó en lugar de exendina - 4 en el vector de expresión pJB02 - exendina - 4 - Fc para crear pJB02 - Val⁸ - Glu²² - GLP - 1 / C - Ex - Fc.

5

Ejemplo 1j Construcción de DNA que codifica Gly⁸ - GLP - 1 - Fc:

El plásmido pJB02 - exendina 4 - Fc se preparó como se ha descrito en el ejemplo 1d. El DNA que codifica exendina - 4 se escindió del vector con *AgeI* y *Eco47III*. El DNA que codifica Gly⁸ - GLP - 1 se generó mediante hibridación *in vitro*

10 de los siguiente oligonucleótidos de superposición y complementarios:

5' - CCGGTCACGGCGAGGGCACCTTCACTAGTGACGTGTCCTCCTATCTGGA
GGGCCAGGCCGCCA - 3' [SEC ID N° 34]

5' - AGGAGTTCATCGCCTGGCTGGTGAAGGGCCGGGGC - 3' [SEC ID N° 35]

5' - CTCCAGATAGGAGGACACGTCACCTAGTGAAGGTGCCCTC
GCCGTGA - 3' [SEC ID N° 36]

5' - GCCCCGGCCCTTCACCAGCCAGGCGATGAACTCCTTGGCGGC
CTGGCC - 3' [SEC ID N° 37]

La reacción de hibridación se llevó a cabo como se ha descrito en el ejemplo 1a.
El producto hibridado se ligó en lugar de exendina - 4 en el vector de expresión pJB02
5 - exendina - 4 - Fc para crear pJB02 - Gly³ - GLP -1 - Fc.

Ejemplo 2: Expresión de proteínas de fusión heterólogas

La expresión de las proteínas de fusión codificadas por las construcciones del
DNA del ejemplo 1 se llevó a cabo mediante transfección transitoria de células HEK
10 293EBNA (tanto adherentes como en suspensión). Las células se contaron y se
sembraron 24 horas antes de la transfección. El cóctel de transfección se preparó
mezclando el reactivo de transfección FuGene™6 (Roche Molecular Biochemicals, N°
de catálogo 1814443) con OptiMEM (Gibco/BRL) y se incubaron a temperatura
ambiente durante 5 minutos momento en el cual se añadió DNA y el cóctel se incubó
15 durante un tiempo adicional de 15 minutos. Inmediatamente antes de la transfección se
añadió medio de cultivo recién preparado a la placa. Las tablas 1 y 2 proporcionan
detalles de la transfección adicionales.

Tabla 1: Los reactivos usados en las transfecciones transitorias de células 293 EBNA.

Recipiente de cultivo de tejidos	Número de células sembradas	DNA (µg)	FuGene (µl)	Medio OptiMEM (ml)	Vol. De medio de cultivo (ml)
35 mm	5 x 10 ⁵	1,5	9	0,102	2
100 mm	2 x 10 ⁶	12	73	0,820	10
700 cm ² (BC)	2 x 10 ⁷	65	400	4,0	100

Tabla 2: Composición de los medios

MEDIO DE TRANSFECCION Y CULTIVO	MEDIO DE RECOLECCION
3:1 de DMEM F12	Base de Hybritech
FBS al 5%	Ca ²⁺ 1 mM
HEPES 20 mM	HEPES 20 mM
L – glutamina 2 mM	Nuselina (insulina humana) 1 µg/ml
Geneticina(G418 NEO) 50 µg/ml	Transferrina humana 1 µg/ml
Trobomicina 50 µg/ml	Trobomicina 50 µg/ml

5

Para transfecciones a pequeña escala (recipientes de 35 mm – 100 mm), se aclararon células con PBS y se cambiaron al medio de recogida 24 horas después de la transfección y se recogió el medio y se sustituyó cada 24 horas durante varios días. En el caso de transfecciones a gran escala (botellas cilíndricas de 700 cm²), las botellas

cilíndricas se aclararon con PBS 48 horas después de la transfección y se cambiaron al medio de recogida. Los medios se recogieron y se cambiaron cada 24 horas durante al menos 10 días consecutivos. Rutinariamente, solamente se usaron 10 recogidas para la purificación posterior de proteínas.

5 **Ejemplo 3: Purificación de proteínas de fusión heterólogas**

Ejemplo 3a purificación de Val⁸ – GLP – 1 – Fc

Se filtraron aproximadamente 4,5 litros de medio acondicionado (aproximadamente 20 µg/ml de nivel de expresión de proteínas de fusión) de transfecciones a gran escala utilizando un sistema de filtro CUNO y se concentraron hasta 250 ml usando un sistema de filtración de flujo tangencial Proflux con un filtro de membrana 10K. Se capturó Val⁸ – GLP – 1 – Fc con una columna de 5 ml de proteína A HiTrap en PBS 1 x, pH 7,4 a un caudal de 2 ml/min y se eluyó con ácido cítrico 50 mM pH 3,3. Las fracciones (1ml) se recogieron en tubos que contenían 4 ml PBS 1 x y 100 µl de Tris 1 M pH 8.

15 Las fracciones que contenían la proteína de fusión según se determinó mediante SDS – PAGE y HPLC de fase inversa en Zorbax C8, se combinaron y se aplicaron a una columna Superdex 75 60/60 en PBS 1 x, pH 7,4 y a un caudal de 10 ml/min. Las fracciones positivas (20 ml/tubo) se recogieron y se combinaron. Después las fracciones combinadas se sometieron a cromatografía en fase inversa C4 en TFA al 0,1% a un caudal de 3 ml/min. Se eluyó Val⁸ – GLP – 1 – Fc usando un gradiente desde 5% de B (TFA al 0,1% en acetonitrilo) hasta 100% de B en 70 minutos. Se recogieron las fracciones eluyentes (3 ml/tubo). Se retiró el acetonitrilo mediante secado a vacío y se añadió 1 ml de H₂O. La muestra purificada (aproximadamente 32 ml) se dializó dos veces frente a 4 litros de PBS 1x pH 7,4.

25 Después se filtró la muestra dializada usando un filtro de 0,22 µm MILLEX –

GV y se determinó la concentración usando absorción a 280 nm.

Ejemplo 3b purificación de Val⁸-GLP-1-HSA o Val⁸-GLP-1-~~engarce~~-HSA

5 Se filtraron aproximadamente 6,5 litros de medio acondicionado (aproximadamente 10 µg/ml de nivel de expresión de proteínas de fusión) utilizando un sistema de filtro CUNO y se concentraron hasta 380 ml usando un sistema de filtración de flujo tangencial Proflux con filtro de membrana 10K.

La proteína de fusión se capturó con una columna Q de 50 ml de flujo rápido
10 (Pharmacia) en Tris 20mM, pH 7,4 a un caudal de 5 ml/min. La proteína se eluyó usando un gradiente: desde 0% a 50% de Tris 20 mM, pH 7,4, NaCl 1M en 10 CV, después hasta 100% de B en 2CV

Las fracciones que contenían la proteína de fusión se combinaron y se sometieron a una cromatografía en fase inversa C4 en TFA al 0,1% en agua a un caudal
15 de 5 ml/min. La proteína de fusión se eluyó usando un gradiente de 20% de B (TFA al 0,1% en acetonitrilo) hasta 90% de B en 120 minutos. Se recogieron las fracciones (3,5 ml/tubo). Se retiró el acetonitrilo mediante secado a vacío.

Se diluyeron aproximadamente 9 ml de muestra combinada con PBS 1 x pH 7,4 hasta 40 ml y se dializó frente a 4 litros de PBS 1x pH 7,4 durante una noche. Se filtró
20 la muestra y se determinó la concentración usando absorción a 280 nm.

Ejemplo 3c purificación de Exendina-4-Fc:

Se filtraron aproximadamente 4 litros de medio acondicionado (aproximadamente 8 µg/ml de nivel de expresión de proteínas de fusión) utilizando un
25 sistema de filtro CUNO y se concentraron hasta 250 ml usando un sistema de filtración

de flujo tangencial Proflux con filtro de membrana 30K.

La exendina – 4 Fc se capturó con una columna de 5 ml de proteína A HiTrap en PBS 1 x, pH 7,4 a un caudal de 2 ml/min y se eluyó con ácido cítrico 50 mM pH 3,3. Las fracciones que contenían la proteína de fusión se combinaron, se filtraron y se
5 dializaron frente a 4 litros de PBS 1 x durante una noche. Después la muestra dializada se aplicó a una columna Superdex 75 60/60 en PBS 1 x, pH 7,4, NaCl 0,5 M a un caudal de 10 ml/min. Las fracciones (20 ml/tubo) que contenían la proteína de fusión se recogieron, se combinaron y se concentraron hasta aproximadamente 1 mg/ml.

Después se filtraron las muestras concentradas usando un filtro de 0,22 µm
10 MILLEX – GV.

Ejemplo 3d purificación de Exendina – 4 – HSA y exendina – 4 – engarce - HSA:

Se filtraron aproximadamente 1,1 litros de medio acondicionado (aproximadamente 6 µg/ml de nivel de expresión de proteínas de fusión) utilizando un
15 sistema de filtro CUNO y se concentraron hasta 175 ml usando un sistema de filtración de flujo tangencial Proflux con filtro de membrana 30K.

La proteína de fusión se capturó usando una columna de 5 ml de HiTrap Q – sefarosa (Pharmacia) en Tris 20 mM, pH 7,4 a un caudal de 2 ml/min. La proteína se eluyó usando un gradiente desde 0% hasta 50% de Tris 20 mM, pH 7,4, NaCl 1 M en 12
20 CV y después hasta 100% de B en 4 CV.

Las fracciones que contenían la proteína de fusión se combinaron y se sometieron a cromatografía en fase inversa C4 en TFA al 0,1% en agua a un caudal de 5 ml/min. Se eluyó la proteína de fusión usando un gradiente desde 10% de B (TFA al 0,1% en acetonitrilo) hasta 100% de B en 70 minutos. Se recogieron las fracciones (10
25 ml/tubo) que contenían la proteína de fusión. Se retiró el acetonitrilo usando un secador

a vacío.

Se dializaron aproximadamente 8 ml de muestra combinada frente a 4 litros de PBS 1 x, pH 7,4 durante una noche. Se filtró la muestra y se determinó la concentración mediante absorción a 280 nm. Después la muestra dializada se aplicó a una columna
5 Superdex 200 26/60 en PBS 1 x, pH 7,4, NaCl 0,5 M a un caudal de 2 ml/min. Las fracciones (3 ml/tubo) que contenían la proteína de fusión se recogieron, se combinaron, se concentraron y se filtraron.

Ejemplo 4: Caracterización de las proteínas de fusión mediante SDS PAGE:

SDS – PAGE seguida de inmunotransferencia se utilizó para analizar tanto la
10 proteína de fusión purificada como el medio acondicionado de las células transfectadas con diversos vectores de expresión de proteínas de fusión. La SDS - PAGE se realizó en un sistema Novex Powereasa 500 que usa geles Precast de Tris-Glicina al 16% Novex (EC6498), tampón de desarrollo (10x, LC2675) y tampón de muestra (L2676). Se redujeron las muestras con DTT 50 mM y se calentaron durante 3 – 5 minutos a 95°C
15 antes de la carga.

Después de desarrollar el gel SDS – PAGE, se utilizaron agua y tampón de transferencia (1x Tris – glicina Seprabuff (Owl Scientific N° de Cat. ER26 – S) con metanol al 20%) para lavar el SDS de los geles. Se utilizó un aparato de transferencia Novex con PVDF (BioRad, N° de Cat. 162 – 0174) y membranas de nitrocelulosa
20 (Biorad, N° de Cat. 1703965 ó 1703932). La transferencia se llevó a cabo a temperatura ambiente durante 90 minutos a 30 – 35 V. Se bloquearon las membranas en PBS 1X con Tween – 20 al 0,1% (Sigma, N° de Cat. P – 7949) y Leche 5% (BioRad, N° de Cat. 170 – 6404) durante 1 – 12 horas a 4°C. Se diluyen los anticuerpos en PBS 1X + Leche al 5% y se incuban las manchas detectadas en estas soluciones durante 1 – 2 horas a
25 4°C. Entre las incubaciones, se lavaron las manchas detectadas cuatro veces durante 5

minutos cada una con PBS 1X y Tween – 20 al 0,2% a temperatura ambiente. PBS se preparó o bien de PBS 10X GIBCO (Nº de Cat. 70011), para dar una composición final de fosfato monobásico de potasio 1 mM, fosfato dibásico de sodio 3 mM, cloruro sódico 153 mM, pH 7,4, o bien a partir de bolsas de PBS de Sigma (Nº de Cat. 1000 – 3), para dar NaCl 120 mM, KCl 2,7 mM y fosfato 10 mM, pH 7,4 a 25°C.

Los anticuerpos principales eran bien un anti – IgG1 de cabra monoclonal o anti – HSA de conejo. El anticuerpo secundario eran bien un anti IgG HRP de cabra o un anti IgG HRP de conejo. Se diluyó el anticuerpo secundario 1:5000. Se utilizó un sistema ECL (Amersham Pharmacia Biotech, Nº de Cat RN2108 y Nº de Cat RPN1674H) para revelar las manchas.

La figura 3A compara la proteína Fc purificada con medio acondicionado a partir de células transfectadas con pJB02 – Val⁸ – GLP -1 – Fc y pJB02 – exendina – 4 – Fc. La disminución en movilidad es coherente con el incremento de tamaño debido a la porción de GLP – 1 de la proteína de fusión. De manera similar la figura 3B compara HSA purificado con medio acondicionado a partir de células transfectadas con pJB02 – Val⁸ – GLP -2 – HSA, pJB02 – Val⁸ – GLP -1 – engarce – HSA, pJB02 – exendina – 4 – HSA, o pJB02 – exendina – 4 – engarce – HSA. La figura 4 identifica las preparaciones de proteínas de fusión purificadas.

20 Ejemplo 5: Caracterización de proteínas de fusión usando espectrometría de masas:

Todos los experimentos se realizaron en un espectrómetro de masas Micromass ToFSpec – 2E equipado con electrónica de enfoque retardado, un Reflectron (usado en análisis de péptidos de intervalo 0 – 8000 Da), un detector lineal (usado durante análisis de alta masa / buena señal) y detector de post aceleración (o P. A. D., usado para análisis de alta masa / señal extremadamente baja). La longitud eficaz de recorrido del

instrumento en modo lineal es 1,2 metros, en el modo Reflectron es 2,3 metros. Se ajustaron dos detectores de placa de microcanal dual para el modo de detección lineal y reflectron. El láser usado es un láser de nitrógeno Laser Science Inc. VSL – 337i que funcionaba a 337 nm a 5 disparos de láser por segundo. Todos los datos adquiridos usando un digitalizador interno de 2 GHz, 8 bit y hasta 50 disparos de láser se promediaron por espectro.

El instrumento funcionó en modo lineal durante el análisis de las proteínas de fusión de GLP – 1 en cuestión. El detector lineal es un dispositivo que detecta iones que viajan por debajo del tubo de fuga del instrumento MALDI – ToF – MS. Mide la abundancia de iones durante el tiempo y envía una señal al digitalizador para conversión. El digitalizador es un convertidor de analógico-digital que permite que la señal del espectrómetro de masas se transfiera al ordenador, donde se reconstruye en un espectro m/z utilizable.

Se utilizó una solución de ácido sinapínico saturado recristalizado (diluido en 50/50 Acn / H₂O y TFA al 0,1%) como matriz de ionización. El ácido sinapínico es una matriz adecuada para las proteínas superiores a 10 kDa. Las proteínas de referencia de masa apropiada se usaron para perfiles de calibración internos y externos con el fin de obtener determinaciones de masas precisas en las muestras analizadas. Todas las muestras se analizaron usando una dilución 1:2 muestra a matriz de dilución. Inicialmente el instrumento se puso en marcha a las siguientes condiciones del detector lineal:

Tensión de fuente: 20,0 keV	Tensión de pulso: 3,0 keV
Tensión de extracción : 20,0 keV	Grueso de láser: 50
Tensión de enfoque: 16,0 keV	Finura de láser: 50

Detector lineal: 3,7 keV
P. A. D.: (desconectado)

Estos ajustes se modificaron (según necesidad) para dar la mejor relación señal / ruido y la más alta resolución. La tabla 3 proporciona una caracterización de las diferentes proteínas de fusión de GLP – 1.

Tabla 3

PROTEÍNA DE FUSIÓN	MASA ESPERADA (KDa)	MASA DETERMINADA POR ESPEC. DE MASAS (KDa)
Val ⁸ – GLP -1 – IgG1	59,08	61,94
Val ⁸ – Glu ²² – GLP -1 – IgG1	59,23	63,61
Gly ³ – GLP -1 – IgG1	59,00	62,93
Val ⁸ – GLP -1 – CEx - IgG1	60,45	65,1 – 65,6
Val ⁸ – Glu ²² – GLP -1 – CEx - IgG1	60,69	65,86
Exendina – 4 – IgG1	60,69	65,86
Val ⁸ - GLP -1 – engarce– HSA	70,70	69,89 – 70,74
Exendina – 4 – HSA	70,56	70,62
Exendina – 4 – engarce - HSA	71,56	71,62

CEx se refiere a una extensión del extremo C terminal y comprende la secuencia de
10 Ser – Ser – Gly – Ala – Pro – Pro – Pro – Ser.
El engarce es Gly – Gly – Gly – Gly – Ser – Gly – Gly – Gly – Gly – Ser – Gly – Gly – Gly – Gly – Ser.

300
297

Ejemplo 6: Actividad de las proteínas de fusión heterólogas:

La capacidad de las proteínas de fusión de la presente invención para activar el receptor de GLP – 1 se evaluó utilizando ensayos *in vitro* tales como los descritos en el documento EP 619.322 de Gelfand y col., y la patente de Estados Unidos N° 5.120.712 respectivamente. La actividad de estos compuestos relativa a la actividad de Val⁸ – GLP -1 (7 – 37) OH se indica en la tabla 4. La figura 8 representa las curvas dosis respuesta *in vitro* de Val⁸ – GLP -1 y proteínas de fusión de Exendina – 4. Además, las tablas 5a y 5b proporcionan la actividad *in vitro* de una gran cantidad de análogos de GLP – 1 que se pueden condensar con una proteína de Fc o de albúmina para elaborar las proteínas de fusión biológicamente activas. Estas actividades se comparan con GLP – 1 (7 – 37) OH.

Tablas 4: Actividad *in vitro* de proteínas de fusión de GLP – 1.

Proteína de fusión	Actividad <i>in vitro</i> (% de Val ⁸ - GLP -1)
Val ⁸ - GLP -IgG1	1
exendina – 4 – IgG1	240
Val ⁸ - GLP -1 – engarce – HSA	0,2
exendina – 4 – HSA	20
Exendina – 4 – engarce – HSA	90
exendina – 4	500
Val ⁸ – Glu ²² – GLP – 1 – IgG1	3,7
Gly ³ – GLP – 1 – IgG1	3,3
Val ⁸ – GLP – 1 – CEx - IgG1	3,3

321
298

Val ⁸ - Glu ²² - GLP - 1 - CEx - IgG1	29
Gly ³ - Glu ²² - GLP - 1 - C2 - IgG1	75
Gly ³ - Glu ²² - GLP - 1 - CEx - engarce - IgG1	150
extendina - 4 - C2 - IgG1	250
extendina - 4 - engarce - IgG1	330
Gly ³ - Glu ²² - GLP - 1 - CEx - engarce - HSA	4
Gly ³ - Glu ²² - GLP - 1 - CEx - engarce - IgG4	80

CEx se refiere a una extensión del extremo C terminal y comprende la secuencia de Ser - Ser - Gly - Ala - Pro - Pro - Pro - Ser.

El engarce es Gly - Gly - Gly - Gly - Ser - Gly - Gly - Gly - Gly - Ser - Gly - Gly -

5 Gly - Gly - Ser.

C2 es Ser - Ser - Gly - Ala - Ser - Ser - Gly - Ala.

Las secuencias de aminoácidos de las proteínas de fusión descritas en las tablas 3 y 4 se representan en las SEC ID N° 13 a SEC ID N° 31:

La secuencia de aminoácidos de Val⁸ - GLP 1 - albúmina sérica humana se
10 representa en la SEC ID N° 13.

1 HVEGTFSTDV SSYLEGQAAK EFIAMLVGR GDAHKSEVAH RFKOLGEENF KALVLIAPAQ
61 YLOQCFFEDH VKLVNEVTET AKTCVADESA ENCDKSLHTL FGDKLCTVAT LRETYGEMAD
121 CCAKQEPERN ECFLOHKIDM PNLFRLVRFV VDVMTAFHD NEETPLKKYL YELARRHPYP
181 YAPELLFFAK RYKAAPTECC QAADKAACLL PKLDELDEG KASSAKQRLK CASLQKFGER
241 AFKAMAVARL SQRPFAEFA EVSKLVTDLT KVHTECCHGD LLECADDRAD LAKYICENQD
301 SISKLKECC EKPLLEKSHC LAEVENDEMP ADLPSLAADF VESKDVCKNY AEAKDVFLGM
361 FLYEYARRHP DYSVVLLRL AKTYETTLEK CCAAADPHDC YAKVFDEFKP LVEEPQGLIK
421 QNCELFBQLG EYKFMALLV RYTKKVPQVS TPTLVEVSRN LGKVGSKCKK HPEAKRMPCA
481 EGYLSVVLNQ LCVLHEKTFV SDRVTRCCTE SLVRRAPCFY ALEVDETYVP KEFNAETTFY
541 HADICTLSEK ERQIKQTAL VELVKHKPKA TKEQLKAVMD DFAAFVERCC KADDKETCPA
601 EGGKLVAAAS QALGL [SEC ID N° 13]

La secuencia de aminoácidos de Val⁸ - GLP-1- engarce - albúmina sérica humana se representa en la SEC ID N° 14.

```

1 HVEGTFTSDV SSYLEGQAAK EPIANLVKGR GGGGSGGGG SGGGSDAHK SEVAHRFKDL
61 GEENFKALVL IAFAYLQQC PFEDHVKLVN VTEFAKTCV ADESAENCDK SLHTLFGDKL
121 CTVATLRETY GENADCCAKQ EPERNECFLO HKDDNPNLPR LVRPEVDVMC TAFHDKETTF
181 LKCYLYELAR RHPFYAPEL LFFAKRYKAA FTECCQAADK AACLLPKLDE LRDEGKASSA
241 KQRLKASLQ KFGERAFAKAV AVARLSQRFK KAEFAEVSKL VTDLTKVHTE CCHGDLLECA
301 DDRADLAKYI CENQDSISSK LKCCCKPIL EKSHCIAEVE NDENPADLPS LAADFVESKD
361 VCKNYAEAKD VFLGMFLYEY ARRHFDYSVV LLLRLAKTYE TTLEKCCAAA DPHECYAKVP
421 DEFKPLVEEP QNLKQNCCL FEQLGEYKFO MALLVRYTKK VPQVSTPTLV EVERNLGKVG
481 SKCKHPEAK RMPCAEDYLS VVLNQLCVLH EKTPVSDRVT KCCTESLVNR RCPFALEVD
541 ETVVPKEFNA ETVTFHADIC TLSEKERQIK QTALVELVK HKFKATKEQL KAVMDFAAF
601 VEKCKADDK ETCFAEBGK LVAASQAALG L [SEC ID N° 14]

```

- 5 La secuencia de aminoácidos de Gly⁸ - Glu²² - GLP - 1- CEx - engarce - albúmina sérica humana se representa en la SEC ID N° 15.

```

1 HVEGTFTSDV SSYLEGQAAK EPIANLVKGR GSSGAPPPSG GGGGSGGGG GGGGSDAHK
61 EVHRFKDLG EENFKALVLI AFAQYLQQCP FEDHVKLVNE VTEFAKTCVA DESAENCDKS
121 LHTLFGDKLC TVATLRETYG ENADCCAKQE PERNECFLOH KDDNPNLPR LVRPEVDVMCT
181 AFHDKETTF LKCYLYELAR RHPFYAPEL LFFAKRYKAA FTECCQAADK AACLLPKLDEL
241 RDEGKASSAK QRLKASLQK FGERAFKANA VARLSQRFK AEFPAEVSKLV TDLTKVHTEC
301 CHGDLLECAD DRADLAKYIC ENQDSISSK LKCCCKPIL EKSHCIAEVEN NDENPADLPSL
361 AADFVESKD VCKNYAEAKD VFLGMFLYEY ARRHFDYSVV LLLRLAKTYET TLEKCCAAD
421 PHECYAKVFD EFKPLVEEPQ NLKQNCCLF FEQLGEYKFO MALLVRYTKK VPQVSTPTLV
481 VERNLGVGS KCKHPEAKR MPCAEDYLS VLNQLCVLH KTPVSDRVT KCCTESLVNRR
541 PCFALEVD ETVVPKEFNA ETVTFHADIC TLSEKERQIK QTALVELVKH KFKATKEQLK
601 AVMDFAAFV EKCKADDK ETCFAEBGK LVAASQAALG L [SEC ID N° 15]

```

- La secuencia de aminoácidos de extendina 4 - albúmina sérica humana se representa en la SEC ID N° 16.

1 HGEFTFTSDL SKQKEEAVR LPIEWLKNNG PSSGAPPPSD AHKSEVAHAF KDLGEMPKA
 61 LVLIAPAYL QOCPPEDHVK LVNVEVTEFAK TCVADESAEN CDKSLHTLFG DKLCVATLR
 121 ETYGEMADCC AKQEPERNEC FLQHKDDMPN LPRLVRPEVD VMCTAFHNE ETLKLYE
 181 IARRHFTYA PELLFFAKRY KAAFTCCQA ADKAACLLPK LDELDEGKA SSARQLKCA
 241 SLQKFGRAF KAVARLSQ RFFKAFAEV SKLVTDLTKV HTECHGDLL SCADRADLA
 301 KYICENQDSI SSKLKECCEK FLLEKSHCIA EVENDEMPAD LPSLAADFVE SKDVCIONYAE
 361 AKDVFLGFL YEYARRHFDY SVVLLLRALAK TYETTLERCC AAADPHECYA KVFDEPKPLV
 421 EEPQMLIKQ CELFEQLGEY KQKALLVKY TKKVPQVSTP TLVEVSRMLG KVGSKCKHP
 481 EAKMPCAEQ YLSVVLNQLC VLHKTQVSD KVTCCTESL VNRAPCFSAE EVDETVPEK
 541 FMAETFTFHA DICTLSEKER QIKQQTALVE LVKHKPKATK EQLKAVMDDF AAFVEKCKA
 601 DDKETCFAEK GKKLVAASQA ALGL [SEC ID N° 16]

La secuencia de aminoácidos de exendina 4 – engarce - albúmina sérica humana se representa en la SEC ID N° 17.

1 HGEFTFTSDL SKQKEEAVR LPIEWLKNNG PSSGAPPPSG GGGGSGGGGS GGGGSDAHKS
 61 EVAHRFIDL EENFKALVLI AFAQYLQOCP FEDHVKLVNE VTEFAKTCVA DESAENCDKS
 121 LHTLFGOKLC TVATLRETYG EMADCCAKQE PERNECFLQH KDDNPHLPRL VRPEVDVMCT
 181 AFHWEETTL KKYLYELARR HPYFYAPELL PFAKRYKAAF TECCQAADKA ACLLPKDEL
 241 RDEGKASEAK QRLKASLQK FGERAFKANA VARLSQRFPK AEFAEVSKLV TDLTKVHTEC
 301 CHGDLLECAD DRADLAKYIC ENQDSISSKL KECCEKPLLE KSHCIAEVEN DEMPADLPSL
 361 AADFVESKDV CKNYAEADV FLGMFLYEYA RHPDYSVVL LLRLAKTYET TLECCAAAD
 421 PHECYAKVFD EFKPLVKEPQ NLIKQNCLEF EQLGEYKQK ALLVRYTKV PQVSTPLVE

5

481 VSRMLGKVGK KCKHPEAKR MPCAEDYLEV VLNQLCVLHE KTFVSDRVTK CCTESLVNRR
 541 PCFSALEVDE TYVPKFNAE TTFHADICT LSEKERQIKQ QTALVELVCH KPKATKEQLK
 601 AVMDDFAAV EKCKADDEK TCFAEKGKL VAASQAALGL [SEC ID N° 17]

La secuencia de aminoácidos de Val¹ – GLP – 1 – IgG1 se representa en la SEC ID N°

10 18.

1 HVEGTFTSDV SSYLEGQAAK EFLAMLVKGR GAEPKSCDKT HTCPFCPAPE LLGGPSVFLF
 61 PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VIDAKTKPRE EQYNSTYKVV
 121 SVLTVLHQDW LMGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SREENTNQV
 181 SLTCLVKGFY PSDIAVENES NGQPENNYKT TTPVLDSGDS FFLYSKLTVD KSRWQQGNVF
 241 SCSVMHEALH NHYTQKLSLS SPGK [SEC ID N° 18]

La secuencia de aminoácidos de Val⁸ – GLP – 1 - CEx – IgG1 se representa en la SEC ID N° 19.

1 HVEGTFTSDV SSYLEGQAAK EFLAMLVKGR GSSGAPFPSA EPKSCDKTHT CPFCPAPELL
 61 GGPSVFLPPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NMYVDGVEVH NAKTKPREEQ
 121 YNSTYKVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIET ISKAKGQPRE PQVYTLPPSR
 181 EENTNQVSEL TCLVKGFYPS DIAVENESNG QPENNYKTP FVLDSGDSFF LYSKLTVDKS
 241 RWQQGNVESC SVMHEALHMH YTQKLSLSG GK [SEC ID N° 19]

5

La secuencia de aminoácidos de Val⁸ – Glu²² GLP – 1 – IgG1 se representa en la SEC ID N° 20.

1 HVEGTFTSDV SSYLEGQAAK EFLAMLVKGR GAEPKSCDKT HTCPFCPAPE LLGGPSVFLF
 61 PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VIDAKTKPRE EQYNSTYKVV
 121 SVLTVLHQDW LMGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SREENTNQV
 181 SLTCLVKGFY PSDIAVENES NGQPENNYKT TTPVLDSGDS FFLYSKLTVD KSRWQQGNVF
 241 SCSVMHEALH NHYTQKLSLS SPGK [SEC ID N° 20]

10

La secuencia de aminoácidos de Val⁸ – Glu²² GLP – 1 - CEx– IgG1 se representa en la SEC ID N° 21.

1 HVEGTFTSDV SSYLEEQAAK EPIANLVKGR GSSGAPPPSA EPKSCDKTHT CPFCPAPPELL
 61 GGPSVFLFFP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
 121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEXT ISKAGQPRE PQVYTLPPSR
 181 EENTKNQVSL TCLVKGFPYS DIAVENESNG QPENNYKTFP FVLDSGGSFF LYSKLTVDKS
 241 RNQQGNVFSC SVMHEALHNN YTKSLSLSP GK [SEC ID N° 21]

La secuencia de aminoácidos de Gly⁸ - Glu²² GLP - 1 - C2 - IgG1 se representa en la SEC ID N° 22.

1 HVEGTFTSDV SSYLEEQAAK EPIANLVKGR GSSGASSGAA EPKSCDKTHT CPFCPAPPELL
 61 GGPSVFLFFP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
 121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEXT ISKAGQPRE PQVYTLPPSR
 181 EENTKNQVSL TCLVKGFPYS DIAVENESNG QPENNYKTFP FVLDSGGSFF LYSKLTVDKS
 241 RNQQGNVFSC SVMHEALHNN YTKSLSLSP GK [SEC ID N° 22]

5

La secuencia de aminoácidos de Gly⁸ - Glu²² GLP - 1 - CEx - engarce - IgG1 se representa en la SEC ID N° 23.

1 HVEGTFTSDV SSYLEEQAAK EPIANLVKGR GSSGAPPPSG GGGSGGGGSG GGSAEPKSC
 61 DKHTTCPPCP APELLGGPSV FLFFPKPKDT LMISRTPEVT CVVDVSHED FEVQFNWYVD
 121 GVEVHDAKTK FREEQYNSTY RVVSVLTVLH QDWLNGKEYK CKVSNKALPA PIKTIISKAK
 181 GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPEDIAVE WEENGPENN YKTFPPVLDS
 241 DGSFFLYSKL TVDKSRNQQG NVFSCSVME ALHHYTKS LSLSPGK [SEC ID N° 23]

La secuencia de aminoácidos de Gly⁸ - Glu²² - GLP - 1 - CEx - engarce - IgG4 se representa en la SEC ID N° 24.

10

1 HVEGTFTSDV SSYLEEQAAK EPIANLVKGR GSSGAPPPSG GGGSGGGGSG GGSAESKYK
 61 PPCPSCPAPE FLGGPSVFLF PPKPKDTLMI SRTPEVTCVV VDVSQEDPEV QFNWYVDGVE
 121 VHDNAKTKPRE EQFNSTYRVV SVLTVLHQDW LNKKEYKCKV SNKGLPSSIE KTISKAGQK
 181 REPQVYTLPP SQEENTKNQV SLTCLVKGFI PSDIAVENES NGQPENNYKT TPFVLDSGGS
 241 PFLYSRLTVD KSRNQGNVVF SCSVMHEALH NHYTKSLSL SLGK [SEC ID N° 24]

La secuencia de aminoácidos de Gly⁸ - Glu²² GLP - 1 - CEx - 2 engarce - IgG1 se representa en la SEC ID N° 25.

1 HGEFTFTSDV SSYLEEQAAK EFLANLVKGR GSSGAPPPSG GGGSGGGGSG GGGSGGGGSG
61 GGGSGGGGSA EFKSCDKTHT CPPCPAPELL GGPSVFLPPP KPKDTLMISR TFEVTCVVVD
121 VSHEDPEVKF NNYVDGVEVH NAKTKPREEQ YNSTYRVVEV LTVLHQDMLN GKKEYCKVSN
181 KALPAPIEKT ISKANGQPRE PQVYTLPPSR EEMTKNQVSL TCLVKGFPYS DIAVENESNG
241 QPENNYKTTP PVLDSDGSFF LYSKLTVDKS RWQQGNVFC SVMHEALHNH YTKSLSLSP
301 GK [SEC ID N° 25]

5 La secuencia de aminoácidos de Gly⁸ - Glu²² GLP - 1 - 2 engarce - IgG1 se representa en la SEC ID N° 26.

1 HGEFTFTSDV SSYLEEQAAK EFLANLVKGR GGGGGSGGGG SGGGGSGGGG SGGGGSGGGG
61 SAEFKSCDKT HTCPFCPAPE LLGGPSVFLF PPKPKDTLMI SRTFEVTCVV VDVSHEDPEV
121 KFNWYVDGVE VHDARTKPRE EQYNSTYRVV SVLTVLHQDW LNKKEYCKKV SNKALPAPIE
181 KTISKANGQP REPQVYTLPP SREEMTKNQV SLTCLVKGFPY PSDIAVENES NGQPENNYKT
241 TFPVLDSDGS FFLYSKLTVD KSRWQQGNV SCSVMHEALH NHYTKSLSL SPGK
[SEC ID N° 26]

La secuencia de aminoácidos de Gly⁸ - Glu²² GLP - 1 - 2 CEx - IgG1 se representa
10 en la SEC ID N° 27.

1 HGEFTFTSDV SSYLEEQAAK EFLANLVKGR GSSGAPPPSS SGAPPPSAEP KSCDKTHTCP
61 PCFAPELLGG PSVFLFPPKP KDTLMISRTP EVTCVVVDVS HEDPEVKFNW YVDGVEVHNA
121 KTKPREEQYN STYRVVSVLT VLNQDMLNGK EYKCKVSNKA LPAPIEKTIS KANGQPREPQ
181 VYTLPPSRER NTKNQVSLTC LVKGFYPEDI AVENESNGQP ENNYKTTPFV LDSDGSFFLY
241 SKLTVDKSRW QQGNVFCSV MHEALHNHYT QKSLSLSPGK [SEC ID N° 27]

La secuencia de aminoácidos de Gly⁸ - Glu²² - Val²⁵ - Ile³³ GLP - 1 - CEx - engarce -
IgG1 se representa en la SEC ID N° 28.

1 HGEGTFTSDV SSYLEEQAVK EFLANLINGR GSSGAPPPSG GGSGGGGSG GGGEAEPKSC
61 DKTHTCPPCP APELLGGPSV FLPPPKPKDT LMISRTPEVT CUVVDVSHED PEVKFNWYVD
121 GVEVHNAKTK PREEQYNSTY RVVSVLTVLH QDMLNGKEYK CKVSNKALPA FIEKTISKAK
181 GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESNQQPEMN YKTTTPVLDS
241 DGSFFLYSKL TVDKSRWQQG NVFSCSVME ALHHYTKQS LSLSPGK [SEC ID N° 28]

La secuencia de aminoácidos de exendina - 4 - IgG1 se representa en la SEC ID N° 29.

1 HGEGTFTSDL SKQMEEEAVR LFIEMLNKGG PSSGAPPPSA EPKSCDKTHT CPFCPAPELL
61 GGPSVFLPPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
121 YNSTYRVVSV LTVLHQDMLN GKEYKCKVSN KALPAPIEKT ISKAEQPRE PQVYTLPPSR
181 KENTKNQVSL TCLVKGFPYS DIAVENESNG QPENNYKTTT PVLDSGGSFF LYSKLTVDKS
241 RWQQGNVFSC SVMHEALHHH YTKSLSLSP GK [SEC ID N° 29]

5

La secuencia de aminoácidos de exendina - 4 - C2 - IgG1 se representa en la SEC ID N° 30.

1 HGEGTFTSDL SKQMEEEAVR LFIEMLNKGG PSSGASSGAA EPKSCDKTHT CPFCPAPELL
61 GGPSVFLPPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
121 YNSTYRVVSV LTVLHQDMLN GKEYKCKVSN KALPAPIEKT ISKAEQPRE PQVYTLPPSR
181 KENTKNQVSL TCLVKGFPYS DIAVENESNG QPENNYKTTT PVLDSGGSFF LYSKLTVDKS
241 RWQQGNVFSC SVMHEALHHH YTKSLSLSP GK [SEC ID N° 30]

La secuencia de aminoácidos de exendina - 4 - engarce - IgG1 se representa en la SEC

10 ID N° 31.

1 HGEGTFTSDL SKQMEEEAVR LFIEMLNKGG PSSGAPPPSG GGSGGGGSG GGSAAEPKSC
61 DKTHTCPPCP APELLGGPSV FLPPPKPKDT LMISRTPEVT CUVVDVSHED PEVKFNWYVD
121 GVEVHNAKTK PREEQYNSTY RVVSVLTVLH QDMLNGKEYK CKVSNKALPA FIEKTISKAK
181 GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPEMN YKTTTPVLDS
241 DGSFFLYSKL TVDKSRWQQG NVFSCSVME ALHHYTKQS LSLSPGK [SEC ID N° 31]

Tabla 5a: Actividad de los análogos de GLP-1 *in vitro*

<u>Compuesto GLP - 1</u>	<u>Activación del receptor de GLP - 1</u>
GLP-1 (7-37)OH	1.0
Val ⁸ -GLP-1 (7-37)OH	0.47 (n = 6)
Gly ⁸ -His ¹¹ -GLP-1 (7-37)OH	0.282
Val ⁸ -Ala ¹¹ -GLP-1 (7-37)OH	0.021
Val ⁸ -Lys ¹¹ -GLP-1 (7-37)OH	0.001
Val ⁸ -Tyr ¹² -GLP-1 (7-37)OH	0.81
Val ⁸ -Glu ¹⁶ -GLP-1 (7-37)OH	0.047

Val ⁸ -Ala ¹⁶ -GLP-1(7-37)OH	0.112
Val ⁸ -Tyr ¹⁶ -GLP-1(7-37)OH	1.175
Val ⁸ -Lys ²⁰ -GLP-1(7-37)OH	0.33
Gln ²² -GLP-1(7-37)OH	0.42
Val ⁸ -Ala ²² -GLP-1(7-37)OH	0.56
Val ⁸ -Ser ²² -GLP-1(7-37)OH	0.50
Val ⁸ -Asp ²² -GLP-1(7-37)OH	0.40
Val ⁸ -Glu ²² -GLP-1(7-37)OH	1.29
Val ⁸ -Lys ²² -GLP-1(7-37)OH	0.58
Val ⁸ -Pro ²² -GLP-1(7-37)OH	0.01
Val ⁸ -His ²² -GLP-1(7-37)OH	0.14
Val ⁸ -Lys ²² -GLP-1(7-36)NH ₂	0.53
Val ⁸ -Glu ²² -GLP-1(7-36)NH ₂	1.0
Gly ⁸ -Glu ²² -GLP-1(7-37)OH	1.07
Val ⁸ -Lys ²³ -GLP-1(7-37)OH	0.18
Val ⁸ -His ²⁴ -GLP-1(7-37)OH	0.007

Val ⁸ -Lys ²⁴ -GLP-1(7-37)OH	0.02
Val ⁸ -His ²⁶ -GLP-1(7-37)OH	1.6
Val ⁸ -Glu ²⁶ -GLP-1(7-37)OH	1.5
Val ⁸ -His ²⁷ -GLP-1(7-37)OH	0.37
Val ⁸ -Ala ²⁷ -GLP-1(7-37)OH	0.47
Gly ⁸ -Glu ³⁰ -GLP-1(7-37)OH	0.29
Val ⁸ -Glu ³⁰ -GLP-1(7-37)OH	0.29
Val ⁸ -Asp ³⁰ -GLP-1(7-37)OH	0.15
Val ⁸ -Ser ³⁰ -GLP-1(7-37)OH	0.19
Val ⁸ -His ³⁰ -GLP-1(7-37)OH	0.19
Val ⁸ -Glu ³³ -GLP-1(7-37)OH	0.039
Val ⁸ -Ala ³³ -GLP-1(7-37)OH	0.1
Val ⁸ -Gly ³³ -GLP-1(7-37)OH	0.01
Val ⁸ -Glu ³⁴ -GLP-1(7-37)OH	0.17
Val ⁸ -Pro ³⁵ -GLP-1(7-37)OH	0.094
Val ⁸ -His ³⁵ -GLP-1(7-37)OH	0.41
Val ⁸ -Glu ³⁵ -GLP-1(7-37)OH	0.15

Val ⁸ -Glu ³⁶ -GLP-1(7-37)OH	0.11
Val ⁸ -His ³⁶ -GLP-1(7-37)OH	0.22
Val ⁸ -His ³⁷ -GLP-1(7-37)OH	0.33
Val ⁸ -Leu ¹⁶ -Glu ²⁶ -GLP-1(7-37)OH	0.23
Val ⁸ -Lys ²² -Glu ³⁰ -GLP-1(7-37)OH	0.37
Val ⁸ -Lys ²² -Glu ²³ -GLP-1(7-37)OH	0.35
Val ⁸ -Glu ²² -Ala ²⁷ -GLP-1(7-37)OH	1.02
Val ⁸ -Glu ²² -Lys ²³ -GLP-1(7-37)OH	1.43
Val ⁸ -Lys ³³ -Val ³⁴ -GLP-1(7-37)OH	0.08
Val ⁸ -Lys ³³ -Asn ³⁴ -GLP-1(7-37)OH	0.09
Val ⁸ -Gly ³⁴ -Lys ³³ -GLP-1(7-37)OH	0.34
Val ⁸ -Gly ³⁶ -Pro ³⁷ -GLP-1(7-37)NH ₂	0.53

Tabla 5b. Actividad de análogos de GLP-1 *in vivo* cti.

Compuesto d GLP-1	Activación del receptor de GLP-1
GLP-1(7-37)OH	1.0
Val ⁸ -GLP-1(7-37)OH	0.47
Gly ⁸ -GLP-1(7-37)OH	0.80
Val ⁸ -Tyr ¹² -GLP-1(7-37)OH	0.80
Val ⁸ -Tyr ¹² -GLP-1(7-36)NH ₂	0.52
Val ⁸ -Trp ¹² -GLP-1(7-37)OH	0.52
Val ⁸ -Leu ¹⁸ -GLP-1(7-37)OH	0.52
Val ⁸ -Val ¹⁸ -GLP-1(7-37)OH	0.52
Val ⁸ -Tyr ¹⁸ -GLP-1(7-37)OH	1.18
Gly ⁸ -Glu ²² -GLP-1(7-37)OH	1.03
Val ⁸ -Leu ²³ -GLP-1(7-37)OH	0.24
Val ⁸ -Tyr ¹² -Tyr ¹⁸ -GLP-1(7-37)OH	0.70
Val ⁸ -Trp ¹² -Glu ²² -GLP-1(7-37)OH	0.80
Val ⁸ -Tyr ¹² -Glu ²² -GLP-1(7-37)OH	1.27
Val ⁸ -Tyr ¹⁸ -Phe ¹⁹ -GLP-1(7-37)OH	1.32
Val ⁸ -Tyr ¹⁸ -Glu ²² -GLP-1(7-37)OH	1.69, 1.79
Val ⁸ -Trp ¹⁸ -Glu ²² -GLP-1(7-37)OH	2.30, 2.16
Val ⁸ -Leu ¹⁸ -Glu ²² -GLP-1(7-37)OH	2.02
Val ⁸ -Ile ¹⁸ -Glu ²² -GLP-1(7-37)OH	1.55
Val ⁸ -Phe ¹⁸ -Glu ²² -GLP-1(7-37)OH	1.08
Val ⁸ -Trp ¹⁸ -Glu ²² -GLP-1(7-37)OH	1.50, 3.10
Val ⁸ -Tyr ¹⁸ -Glu ²² -GLP-1(7-37)OH	2.40, 2.77
Val ⁸ -Phe ¹⁸ -Glu ²² -GLP-1(7-37)OH	0.94
Val ⁸ -Ile ¹⁸ -Glu ²² -GLP-1(7-37)OH	1.88
Val ⁸ -Lys ¹⁸ -Glu ²² -GLP-1(7-37)OH	1.18
Val ⁸ -Trp ¹⁹ -Glu ²² -GLP-1(7-37)OH	1.50
Val ⁸ -Phe ¹⁹ -Glu ²² -GLP-1(7-37)OH	0.70
Val ⁸ -Phe ²⁰ -Glu ²² -GLP-1(7-37)OH	1.27

Val ⁸ -Glu ²² -Leu ²⁵ -GLP-1(7-37)OH	1.32
Val ⁸ -Glu ²² -Ile ²⁵ -GLP-1(7-37)OH	1.46
Val ⁸ -Glu ²² -Val ²⁵ -GLP-1(7-37)OH	2.21, 1.36
Val ⁸ -Glu ²² -Ile ²⁷ -GLP-1(7-37)OH	0.94
Val ⁸ -Glu ²² -Ala ²⁷ -GLP-1(7-37)OH	1.03
Val ⁸ -Glu ²² -Ile ³³ -GLP-1(7-37)OH	2.21, 1.79, 1.60
Val ⁸ -Asp ⁹ -Ile ¹¹ -Tyr ¹⁶ -Glu ²² - GLP-1(7-37)OH	2.02
Val ⁸ -Tyr ¹⁶ -Trp ¹⁹ -Glu ²² -GLP-1(7- 37)OH	1.64
Val ⁸ -Trp ¹⁶ -Glu ²² -Val ²⁵ -Ile ³³ - GLP-1(7-37)OH	2.35
Val ⁸ -Trp ¹⁶ -Glu ²² -Ile ³³ -GLP-1(7- 37)OH	1.93
Val ⁸ -Glu ²² -Val ²⁵ -Ile ³³ -GLP-1(7- 37)OH	2.30, 2.73, 3.15
Val ⁸ -Trp ¹⁶ -Glu ²² -Val ²⁵ -GLP-1(7- 37)OH	2.07
Val ⁸ -Cys ¹⁶ -Lys ²⁶ -GLP-1(7-37)OH	1.97
Val ⁸ -Cys ¹⁶ -Lys ²⁶ -Arg ³⁸ -GLP-1(7- 37)OH	2.4, 1.9

Tabla 6. Actividad *in vitro* de análogos de GLP/exendina

Secuencia de péptidos	Actividad <i>in vitro</i> (% de Val – GLP – 1 (7 – 37) OH)
HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGP-NH2	6.21
HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPP PS-NH2	6.75, 3.25
HVEGTFTSDLSKQMEEEAVRLFIANLVKGRG	2.86
HVEGTFTSDVSSYLEEEAVRLFIANLVKGRG	1.47
HVEGTFTSDLSKQMEGQAAKEFIANLVKGRG	0.11

HVEGTFTSDVSKQMEGQAAKEFIANLVKGRG	0.04
HGEGTFTSDLSKQMEGQAAKEFIEWLKNGGP-NH2	1.44
HGEGTFTSDLSKQMEEEAAKEFIEWLKNGGP-NH2	2.80
HGEGTFTSDVSSYLEEEAVRLFIEWLKNGGP-NH2	5.40
HGEGTFTSDLSSYLEEEAVRLFIEWLKNGGP-NH2	5.07
HAEGTFTSDVSSYLEGQAAKEFIANLVKGRPSSGAPPPS -NH2	3.30
HAEGTFTSDVSKQLEEEAAKEFIANLVKGRG	2.15
HVEGTFTSDVSSYLEGQAAKEFIEWLKNGGP-NH2	2.36
HGEGTFTSDLSKQMEEEAVRLFIANLVKGRG	3.25
HVEGTFTSDVSSYLEEEAAKEFIANLVKGRG	1.00
HVEGTFTSDVSSYLEGQAAKEFIANLVKGRG	0.20
HVEGTFTSDVSSYLEGQAAKEFIANLVKGRG	1.00
HAEGTFTSDVSSYLEGQAAKEFIANLVKGRG	2.12

Ejemplo 7: Farmacocinética *in vivo* de Val⁸ – GLP – 1 – IgG1 y Val⁸ – GLP – 1- HSA:

5 Se realizó en monos cinomólogos un estudio farmacocinético de Val⁸ – GLP – 1 – IgG1 y Val⁸ – GLP – 1 – HSA. Se administró una dosis a los monos de 5,6 nmoles/kg

- tanto con Val⁸ - GLP - 1 - IgG1 o con Val⁸ - GLP - 1 - HSA purificadas. Se administraron los compuestos en forma de administración intravenosa de un bolo. Se recogió sangre antes de la administración de la dosis y a 0,083, 0,25, 0,5, 1, 4, 8, 12, 24, 48, 72, 96, 120, 144, 168 y 216 horas después de la dosis en tubos que contenían EDTA.
- 5 Se determinaron las concentraciones en plasma de Val⁸ - GLP - 1 inmunorreactiva utilizando un radioinmunoensayo que utiliza un antisuero policlonal cuya especificidad principal es para la región (7 - 16) del extremo N terminal de Val⁸ - GLP - 1 (7 - 37). La figura 9 representa la concentración en plasma de Val⁸ - GLP - 1 - Fc y Val⁸ - GLP - 1 - engarce - HSA seguido de una dosis única intravenosa a dos monos cinomólogos.
- 10 La proteína de fusión de Fc tenía una semivida de aproximadamente 45 horas y la proteína de fusión de albúmina tenía una semivida de aproximadamente 87 horas.

Ejemplo 8: Farmacodinámica *in vivo* de exendina - 4 - IgG1:

- Se estudiaron dos perros sabuesos macho normales canulados crónicamente después de una noche de ayuno. Se accedió a puertos de acceso vascular venoso y
- 15 arterial y se insertó percutáneamente un catéter en una vena cefálica y se aseguró. Los animales se colocaron en jaulas y sus catéteres se acoplaron a un sistema giratorio / de correas. Una solución que contenía la proteína de fusión exendina - 4 - IgG1 (11, 8 M) a vena cefálica.

- Después se limpió el catéter con 10 ml de solución salina. Dos horas después, se inició
- 20 una pinza hiperglucémica (150 mg/dl) y se continuó durante tres horas. Se tomaron muestras de sangre arterial durante este período de 5 horas para determinar las concentraciones en plasma de la proteína de fusión, glucosa e insulina.

- Los resultados de este estudio se compararon con los de un estudio similar previo en el que los dos animales habían recibido un bolo de solución salina., s.
- 25 c. y tres horas más tarde se estudiaron utilizando una pinza hiperglucémica durante 3

horas (150 mg/dl).

En ambos conjuntos de estudios, las concentraciones de glucosa en plasma se determinaron usando un analizador de glucosa Beckman. Las concentraciones de insulina en plasma se determinaron mediante personal empleado de Linco Research, Inc. Usando un equipo RIA desarrollado en sus laboratorios. Los datos se ilustran en las figuras 10 y 11.

Ejemplo 9: farmacocinética *in vivo* de Gly⁸ - Glu²² - GLP - 1 - CEx - engarce - IgG1:

10 Dos grupos de tres perros sabuesos machos normales recibieron 0,1 mg/Kg de Gly⁸ - Glu²² - GLP - 1 - CEx - engarce - IgG1 mediante administración subcutánea (abreviadamente SC) o intravenosa (abreviadamente IV). Las concentraciones en plasma de la inmunorreactividad de Gly⁸ - Glu²² - GLP - 1 - CEx - engarce - IgG1 se determinaron mediante radioinmunoensayo en muestras recogidas desde 30 minutos
15 antes de la dosis hasta 216 horas después de la dosis para ambos grupos IV y SC. Estas concentraciones se usaron posteriormente para determinar los parámetros farmacocinéticos indicados. La semivida de eliminación media de Gly⁸ - Glu²² - GLP - 1 - CEx - engarce - IgG1 administrada IV fue aproximadamente de 55 horas y la eliminación total en el cuerpo era de 1,5 ml/h/kg. La semivida de eliminación media de
20 la Gly⁸ - Glu²² - GLP - 1 - CEx - engarce - IgG1 administrada SC fue aproximadamente de 38 horas.

Habiendo descrito la invención como antecede, se declara como propiedad lo contenido en las siguientes

REIVINDICACIONES

1. Una proteína de fusión heteróloga que comprende un primer polipéptido con
5 un extremo N terminal y un extremo C terminal condensados con un
segundo polipéptido con un extremo N terminal y un extremo C terminal en
la que el primer polipéptido es un compuesto GLP – 1 y el segundo
polipéptido se selecciona del grupo constituido por
- a) albúmina humana;
 - 10 b) análogos de albúmina humana; y
 - b) fragmentos de albúmina humana,

y en la que el extremo C terminal del primer polipéptido se fusiona al extremo N
terminal del segundo polipéptido.

- 15 2. Una proteína de fusión heteróloga que comprende un primer polipéptido con
un extremo N terminal y un extremo C terminal condensados con un
segundo polipéptido con un extremo N terminal y un extremo C terminal en
la que el primer polipéptido es un compuesto de GLP – 1 y el segundo
polipéptido se selecciona del grupo constituido por
- 20 a) albúmina humana;
 - b) análogos de albúmina humana; y
 - c) fragmentos de albúmina humana,

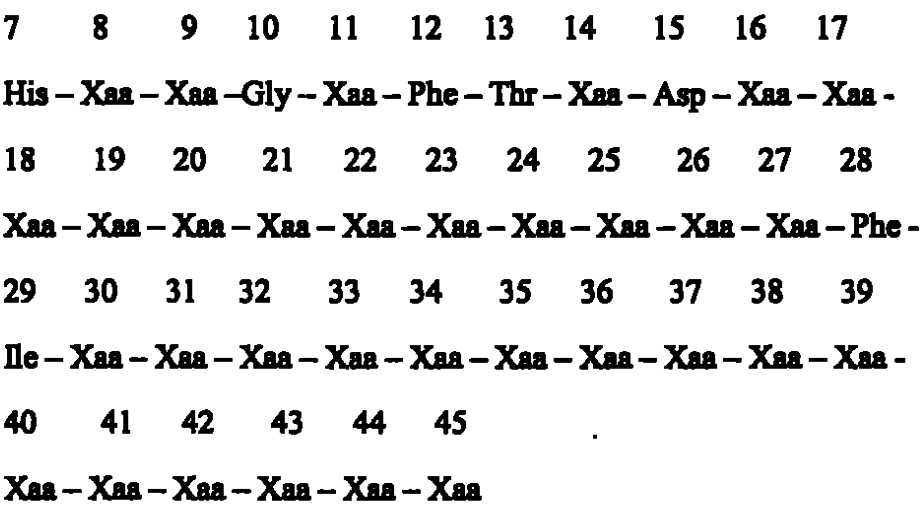
y en la que el extremo C terminal del primer polipéptido se fusiona al extremo N
terminal del segundo polipéptido mediante un enlazador de péptidos.

25

3. La proteína de fusión heteróloga de la reivindicación 2 en la que el engarce de péptidos se selecciona del grupo constituido por:
- a) un péptido rico en glicina;
 - b) un péptido que tiene la secuencia [Gly – Gly – Gly – Gly - Ser]_n en el que n es 1, 2, 3, 4, 5, ó 6; y
 - c) un péptido que tiene la secuencia [Gly – Gly – Gly – Gly - Ser]₃.

4. La proteína de fusión heteróloga de las reivindicaciones 1, 2, ó 3 en la que el compuesto de GLP – 1 comprende la secuencia de fórmula I [SEC ID N° 2]

10



15

Fórmula I (SEC ID N° 2)

20

en la que:

Xaa en la posición 8 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

Xaa en la posición 9 es Glu, Asp, o Lys;

Xaa en la posición 11 es Thr, Ala, Gly, Ser, Leu, Ile, Val, Glu, Asp o Lys;

Xaa en la posición 14 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp o Lys;

25 Xaa en la posición 16 es Val, Ala, Gly, Ser, Thr, Leu, Ile, Tyr, Glu, Asp, Trp o

Lys;

Xaa en la posición 17 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp o Lys;

Xaa en la posición 18 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, Trp, Tyr o Lys;

5 Xaa en la posición 19 es Tyr, Phe, Trp, Glu, Asp, Gln, o Lys;

Xaa en la posición 20 es Leu, Ala, Gly, Ser, Thr, Ile, Val, Glu, Asp, Met, Trp, Tyr o Lys;

Xaa en la posición 21 es Glu, Asp, o Lys;

Xaa en la posición 22 es Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, o Lys

10 Xaa en la posición 23 es Gln, Asn, Arg, Glu, Asp, o Lys;

Xaa en la posición 24 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Arg, Glu, Asp o Lys;

Xaa en la posición 25 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

Xaa en la posición 26 es Lys, Arg, Gln, Glu, Asp, o His;

Xaa en la posición 27 es Leu, Glu, Asp o Lys;

15 Xaa en la posición 30 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

Xaa en la posición 31 es Trp, Phe, Tyr, Glu, Asp o Lys;

Xaa en la posición 32 es Leu, Gly, Ala, Ser, Thr, Ile, Val, Glu, Asp o Lys;

Xaa en la posición 33 es Val, Gly, Ala, Ser, Thr, Leu, Ile, Glu, Asp o Lys;

Xaa en la posición 34 es Asn, Lys, Arg, Glu, Asp o His;

20 Xaa en la posición 35 es Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

Xaa en la posición 36 es Gly, Arg, Lys, Glu, Asp o His;

Xaa en la posición 37 es Pro, Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys o

se ha suprimido;

Xaa en la posición 38 es Ser, Arg, Lys, Glu, Asp o His o se ha suprimido;

25 Xaa en la posición 39 es Ser, Arg, Lys, Glu, Asp o His, o se ha suprimido;

Xaa en la posición 40 es Gly, Asp, Glu o Lys o se ha suprimido;
Xaa en la posición 41 es Ala, Phe, Trp, Tyr, Glu, Asp o Lys, o se ha suprimido;
Xaa en la posición 42 es Ser, Pro, Lys, Glu o Asp, o se ha suprimido;
Xaa en la posición 43 es Ser, Pro, Glu, Asp o Lys, o se ha suprimido;
5 Xaa en la posición 44 es Gly, Pro, Glu, Asp o Lys, o se ha suprimido;
y
Xaa en la posición 45 es Ala, Ser, Val, Glu, Asp o Lys, o se ha suprimido;
con la condición de que cuando el aminoácido en la posición 37, 38, 39, 40, 41, 42, 43 ó
44 se ha suprimido, entonces cada aminoácido cadena abajo de ese aminoácido también
10 se ha suprimido.

5. La proteína de fusión heteróloga de las reivindicaciones 1, 2 ó 3 en la que el
compuesto de GLP – 1 comprende la secuencia de fórmula II (SEC ID N° 3)

15 7 8 9 10 11 12 13 14 15 16 17
Xaa – Xaa – Xaa – Gly – Xaa – Xaa – Thr – Ser – Asp – Xaa – Ser -
18 19 20 21 22 23 24 25 26 27 28
Xaa – Tyr – Leu – Glu – Xaa – Xaa – Xaa – Ala – Xaa – Xaa – Phe -
29 30 31 32 33 34 35 36 37
20 Ile – Xaa – Xaa – Leu – Xaa – Xaa – Xaa – Xaa – Xaa

Fórmula II (SEC ID N° 3)

en la que:
Xaa en la posición 7 es: L - histidina, D – histidina, desaminohistidina, 2 -
aminohistidina, β – hidroxihistidina, homohistidina, α – fluorometilhistidina o α –
25 metilhistidina;

- Xaa en la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser o Thr;
Xaa en la posición 9 es: Thr, Ser, Arg, Lys, Trp, Phe, Tyr, Glu o His;
Xaa en la posición 11 es: Asp, Glu, Arg, Thr, Ala, Lys o His;
Xaa en la posición 12 es: His, Trp, Phe o Tyr;
5 Xaa en la posición 16 es: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Tyr, Glu o Ala;
Xaa en la posición 18 es: His, Pro, Asp, Glu, Arg, Ser, Ala o Lys;
Xaa en la posición 19 es: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;
Xaa en la posición 23 es: His, Asp, Lys, Glu, Gln o Arg;
Xaa en la posición 24 es: Glu, Arg, Ala o Lys;
10 Xaa en la posición 26 es: Trp, Tyr, Phe, Asp, Lys, Glu o His;
Xaa en la posición 27 es: Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;
Xaa en la posición 30 es: Ala, Glu, Asp, Ser o His;
Xaa en la posición 31 es: Asp, Glu, Ser, Thr, Arg, Trp o Lys;
Xaa en la posición 33 es: Asp, Arg, Val, Lys, Ala, Gly o Glu;
15 Xaa en la posición 34 es: Glu, Lys o Asp;
Xaa en la posición 35 es: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o
Glu;
Xaa en la posición 36 es: Thr, Ser, Asp, Trp, Tyr, Phe, Arg, Glu o His;
Xaa en la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly,
20 Gly – Pro o se ha suprimido.

6. La proteína de fusión heteróloga de las reivindicaciones 1, 2 ó 3 en la que el
compuesto GLP – 1 comprende la secuencia de fórmula III (SEC ID N° 4)
- | | | | | | | | | | | | | | | | | | | | | | | |
|----|-----|---|-----|----|-----|----|-----|----|-----|----|-----|---|-----|---|-----|---|-----|---|-----|---|-----|---|
| 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | | | | | | | | | | | | |
| 25 | Xaa | – | Xaa | – | Glu | – | Gly | – | Xaa | – | Xaa | – | Thr | – | Ser | – | Asp | – | Xaa | – | Ser | – |

18 19 20 21 22 23 24 25 26 27 28
Ser – Tyr – Leu – Glu – Xaa – Xaa – Xaa – Ala – Xaa – Xaa – Phe -
29 30 31 32 33 34 35 36 37
Ile – Ala – Xaa – Leu – Xaa – Xaa – Xaa – Xaa – Xaa

5 **Fórmula III (SEC ID Nº 4)**

en la que:

Xaa en la posición 7 es: L - histidina, D - histidina, desaminohistidina, 2 - aminohistidina, β - hidroxihistidina, homohistidina, α - fluorometilhistidina o α - metilhistidina;

- 10 Xaa en la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser o Thr;
 Xaa en la posición 11 es: Asp, Glu, Arg, Thr, Ala, Lys o His;
 Xaa en la posición 12 es: His, Trp, Phe o Tyr;
 Xaa en la posición 16 es: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu o Ala;
 Xaa en la posición 22 es: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;
- 15 Xaa en la posición 23 es: His, Asp, Lys, Glu o Gln;
 Xaa en la posición 24 es: Glu, His, Ala o Lys;
 Xaa en la posición 25 es: Asp, Lys, Glu o His;
 Xaa en la posición 27 es: Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;
 Xaa en la posición 30 es: Ala, Glu, Asp, Ser o His;
- 20 Xaa en la posición 33 es: Asp, Arg, Val, Lys, Ala, Gly o Glu;
 Xaa en la posición 34 es: Glu, Lys o Asp;
 Xaa en la posición 35 es: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o Glu;
- Xaa en la posición 36 es: Asp, Glu o His;
- 25 Xaa en la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly,

Gly – Pro o se ha suprimido.

7. La proteína de fusión heteróloga de las reivindicaciones 1, 2 ó 3 en la que el compuesto GLP – 1 comprende la secuencia de fórmula IV (SEC ID Nº 5)

5

7 8 9 10 11 12 13 14 15 16 17
Xaa – Xaa – Glu – Gly – Thr – Xaa – Thr – Ser – Asp – Xaa – Ser –
18 19 20 21 22 23 24 25 26 27 28
Ser – Tyr – Leu – Glu – Xaa – Xaa – Ala – Ala – Xaa – Glu – Phe –
10 29 30 31 32 33 34 35 36 37
Ile – Xaa – Trp – Leu – Val – Lys – Xaa – Arg – Xaa

Fórmula IV (SEC ID Nº 5)

en la que:

Xaa en la posición 7 es: L - histidina, D - histidina, desaminohistidina, 2 -
15 aminohistidina, β – hidroxihistidina, homohistidina, α – fluorometilhistidina o α –
metilhistidina;

Xaa en la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser, o Thr;

Xaa en la posición 12 es: His, Trp, Phe o Tyr;

Xaa en la posición 16 es: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu o Ala;

20 Xaa en la posición 22 es: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;

Xaa en la posición 23 es: His, Asp, Lys, Glu o Gln;

Xaa en la posición 26 es: Asp, Lys, Glu o His;

Xaa en la posición 30 es: Ala, Glu, Asp, Ser o His;

Xaa en la posición 35 es: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o
25 Glu;

Xaa en la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, -NH₂, Gly, Gly – Pro o Gly – Pro - NH₂, o se ha suprimido.

8. La proteína de fusión heteróloga de las reivindicaciones 1, 2 ó 3 en la que el compuesto GLP – 1 comprende la secuencia de fórmula V (SEC ID N° 6)

5 7 8 9 10 11 12 13 14 15 16 17
Xaa – Xaa – Glu – Gly – Thr – Phe – Thr – Ser – Asp – Val – Ser –
18 19 20 21 22 23 24 25 26 27 28
Ser – Tyr – Leu – Glu – Xaa – Xaa – Ala – Ala – Lys – Xaa – Phe –
29 30 31 32 33 34 35 36 37
10 Ile – Xaa – Trp – Leu – Val – Lys – Gly – Arg – Xaa
Fórmula V (SEC ID N° 6)

en la que:

Xaa en la posición 7 es: L - histidina, D - histidina, desaminohistidina, 2 - aminohistidina, β – hidroxihistidina, homohistidina, α – fluorometilhistidina o α –
15 metilhistidina;

Xaa en la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser o Thr;
Xaa en la posición 22 es: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;
Xaa en la posición 23 es: His, Asp, Lys, Glu o Gln;
Xaa en la posición 24 es: Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;
20 Xaa en la posición 30 es: Ala, Glu, Asp, Ser o His;

Xaa en la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly – Pro o se ha suprimido.

9. La proteína de fusión heteróloga de las reivindicaciones 1, 2 ó 3 en la que el
25 compuesto GLP – 1 comprende la secuencia de fórmula VIII (SEC ID N° 11)

5 7 8 9 10 11 12 13 14 15 16 17
Xaa – Xaa – Glu – Gly – Thr – Phe – Thr – Ser – Asp – Xaa – Ser –
18 19 20 21 22 23 24 25 26 27 28
Xaa – Xaa – Xaa – Glu – Xaa – Xaa – Ala – Xaa – Xaa – Xaa – Phe –
29 30 31 32 33 34 35 36 37
Ile – Xaa – Trp – Leu – Xaa – Xaa – Gly – Xaa – Xaa

Fórmula VIII (SEC ID N° 11)

10 en la que:

Xaa en la posición 7 es: L - histidina, D - histidina, desaminohistidina, 2 - aminohistidina, β - hidroxihistidina, homohistidina, α - fluorometilhistidina o α - metilhistidina;

Xaa en la posición 8 es: Gly, Ala o Val;

15 Xaa en la posición 16 es: Leu o Val;

Xaa en la posición 18 es: Lys o Ser;

Xaa en la posición 19 es: Gln o Tyr;

Xaa en la posición 20 es: Met o Leu;

Xaa en la posición 22 es: Glu o Gln;

20 Xaa en la posición 23 es: Glu o Gln;

Xaa en la posición 25 es: Val o Ala;

Xaa en la posición 26 es: Arg o Lys;

Xaa en la posición 27 es: Leu o Glu;

Xaa en la posición 30 es: Glu o Ala;

25 Xaa en la posición 33 es: Val o Lys;

Xaa en la posición 34 es: Asn o Lys;

Xaa en la posición 36 es: Gly o Arg; y

Xaa en la posición 37 es: Gly, Pro, Pro – Ser – Ser – Gly – Ala – Pro – Pro – Pro – Ser, o se está ausente.

5

10. La proteína de fusión heteróloga de las reivindicaciones 1 a 9 en la que el compuesto GLP – 1 no tiene más de 6 aminoácidos que son diferentes del correspondiente aminoácido en GLP – 1 (7 – 37) OH, GLP – 1 (7 – 36) OH o exendina – 4.

10

11. La proteína de fusión heteróloga de la reivindicación 10 en la que el compuesto GLP – 1 no tiene más de 5 aminoácidos que son diferentes del correspondiente aminoácido en GLP – 1 (7 – 37) OH, GLP – 1 (7 – 36) OH o exendina – 4.

15

12. La proteína de fusión heteróloga de la reivindicación 11 en la que el compuesto GLP – 1 no tiene más de 4 aminoácidos que son diferentes del correspondiente aminoácido en GLP – 1 (7 – 37) OH, GLP – 1 (7 – 36) OH o exendina – 4.

20

13. La proteína de fusión heteróloga de la reivindicación 12 en la que el compuesto GLP – 1 no tiene más de 3 aminoácidos que son diferentes del correspondiente aminoácido en GLP – 1 (7 – 37) OH, GLP – 1 (7 – 36) OH o exendina – 4.

25

14. La proteína de fusión heteróloga de la reivindicación 13 en la que el compuesto GLP – 1 no tiene más de 2 aminoácidos que son diferentes del correspondiente aminoácido en GLP – 1 (7 – 37) OH, GLP – 1 (7 – 36) OH o exendina – 4.
- 5
15. La proteína de fusión heteróloga de una cualquiera de las reivindicaciones de 1 a 14 en la que Xaa en la posición 8 es glicina o valina.
- 10
16. La proteína de fusión heteróloga de la reivindicación 8 en la que el compuesto GLP – 1 no tiene más de 2 aminoácidos que son diferentes del correspondiente aminoácido en GLP – 1 (7 – 37) OH o GLP – 1 (7 – 36)OH y Xaa en la posición 8 es glicina o valina y Xaa en la posición 30 es alanina, ácido glutámico, ácido aspártico, serina o histidina.
- 15
17. La proteína de fusión heteróloga de la reivindicación 16 en la que Xaa en la posición 30 es ácido glutámico.
- 20
18. La proteína de fusión heteróloga de la reivindicación 8 en la que el compuesto GLP – 1 no tiene más de 2 aminoácidos que son diferentes del correspondiente aminoácido en GLP – 1 (7 – 37) OH o GLP – 1 (7 – 36)OH y Xaa en la posición 8 es glicina o valina y Xaa en la posición 37 es histidina, fenilalanina, tirosina, triptófano, ácido aspártico, ácido glutámico, serina, treonina, arginina o lisina.
- 25
19. La proteína de fusión heteróloga de la reivindicación 18 en la que Xaa en la

posición 37 es histidina.

- 5 20. La proteína de fusión heteróloga de la reivindicación 8 en la que el compuesto GLP – 1 no tiene más de 2 aminoácidos que son diferentes del correspondiente aminoácido en GLP – 1 (7 – 37) OH o GLP – 1 (7 – 36)OH y Xaa en la posición 8 es glicina, valina, leucina, isoleucina, serina, treonina o metionina y Xaa en la posición 22 es ácido aspártico, ácido glutámico, lisina, arginina, asparagina, glutamina o histidina.
- 10 21. La proteína de fusión heteróloga de la reivindicación 20 en la que Xaa en la posición 22 es lisina o ácido glutámico.
- 15 22. La proteína de fusión heteróloga de la reivindicación 8 en la que el compuesto GLP – 1 es Val⁸ – GLP – 1 (7 – 37).
- 20 23. La proteína de fusión heteróloga de la reivindicación 8 en la que el compuesto GLP – 1 es Gly⁸ – GLP – 1 (7 – 37).
- 20 24. La proteína de fusión heteróloga de una cualquiera de las reivindicaciones de 1 a 23 en la que el segundo polipéptido es albúmina humana.
- 25 25. La proteína de fusión heteróloga de la reivindicación 24 en la que el segundo polipéptido tiene la secuencia de SEC ID N° 34.
- 25 26. La proteína de fusión heteróloga de una cualquiera de las reivindicaciones de

1 a 23 en la que el segundo polipéptido es un fragmento del extremo N terminal de albúmina.

27. Una proteína de fusión heteróloga que comprende un primer polipéptido con un extremo N terminal y un extremo C terminal fusionado a un segundo polipéptido con un extremo N terminal y un extremo C terminal en la que el primer polipéptido es un compuesto GLP – 1 y el segundo polipéptido se selecciona del grupo constituido por
- a) la porción Fc de una inmunoglobulina;
 - b) un análogo de la porción Fc de una inmunoglobulina;
- 10 y
- c) fragmentos de la porción Fc de una inmunoglobulina, y en la que el extremo C terminal del primer polipéptido está condensado con extremo N terminal del segundo polipéptido.
- 15 28. Una proteína de fusión heteróloga que comprende un primer polipéptido con un extremo N terminal y un extremo C terminal condensado a un segundo polipéptido con un extremo N terminal y un extremo C terminal en la que el primer polipéptido es un compuesto GLP – 1 y el segundo polipéptido se selecciona del grupo constituido por
- 20 a) la porción Fc de una inmunoglobulina;
 - b) un análogo de la porción Fc de una inmunoglobulina;
- y
- c) fragmentos de la porción Fc de una inmunoglobulina, y en la que el extremo C terminal del primer polipéptido está condensado con el extremo N terminal del segundo polipéptido mediante un engarce de
- 25

péptidos.

29. La proteína de fusión heteróloga de la reivindicación 28 en la que el engarce de péptidos se selecciona del grupo constituido por
- 5

a) un péptido rico en glicina;

b) un péptido que tiene la secuencia [Gly – Gly – Gly – Gly - Ser]_n en la que n es 1, 2, 3, 4, 5 ó 6; y

c) un péptido que tiene la secuencia [Gly – Gly – Gly – Gly - Ser]₃.
- 10

30. La proteína de fusión heteróloga de una cualquiera de las reivindicaciones 27, 28 ó 29 en la que el compuesto GLP – 1 comprende la secuencia de fórmula 1 [SEC ID N° 2]

7891011121314151617

His – Xaa – Xaa – Gly – Xaa – Phe – Thr – Xaa – Asp – Xaa – Xaa -

15

1819202122232425262728

Xaa – Xaa – Xaa – Xaa – Xaa – Xaa – Xaa – Xaa – Xaa – Xaa – Phe -

2930313233343536373839

Ile – Xaa – Xaa – Xaa – Xaa – Xaa – Xaa – Xaa – Xaa – Xaa – Xaa -

404142434445

20Xaa – Xaa – Xaa – Xaa – Xaa – Xaa

Fórmula I (SEC ID N° 2)

en la que:

Xaa en la posición 8 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

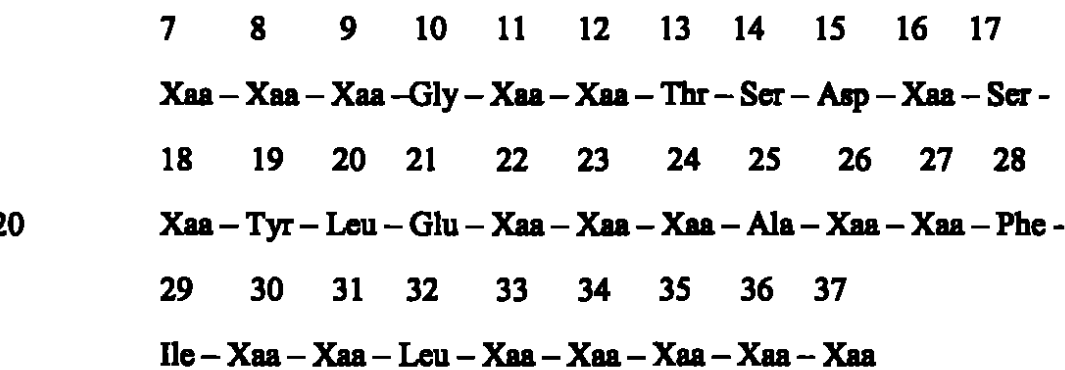
Xaa en la posición 9 es Glu, Asp, o Lys;

25Xaa en la posición 11 es Thr, Ala, Gly, Ser, Leu, Ile, Val, Glu, Asp o Lys;

- Xaa en la posición 14 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp o Lys;
- Xaa en la posición 16 es Val, Ala, Gly, Ser, Thr, Leu, Ile, Tyr, Glu, Asp, Trp o Lys;
- Xaa en la posición 17 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp o Lys;
- 5 Xaa en la posición 18 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, Trp, Tyr o Lys;
- Xaa en la posición 19 es Tyr, Phe, Trp, Glu, Asp, Gln, o Lys;
- Xaa en la posición 20 es Leu, Ala, Gly, Ser, Thr, Ile, Val, Glu, Asp, Met, Trp, Tyr o Lys;
- 10 Xaa en la posición 21 es Glu, Asp, o Lys;
- Xaa en la posición 22 es Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, o Lys
- Xaa en la posición 23 es Gln, Asn, Arg, Glu, Asp, o Lys;
- Xaa en la posición 24 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Arg, Glu, Asp o Lys;
- Xaa en la posición 25 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;
- 15 Xaa en la posición 26 es Lys, Arg, Gln, Glu, Asp, o His;
- Xaa en la posición 27 es Leu, Glu, Asp o Lys;
- Xaa en la posición 30 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;
- Xaa en la posición 31 es Trp, Phe, Tyr, Glu, Asp o Lys;
- Xaa en la posición 32 es Leu, Gly, Ala, Ser, Thr, Ile, Val, Glu, Asp o Lys;
- 20 Xaa en la posición 33 es Val, Gly, Ala, Ser, Thr, Leu, Ile, Glu, Asp o Lys;
- Xaa en la posición 34 es Asn, Lys, Arg, Glu, Asp o His;
- Xaa en la posición 35 es Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;
- Xaa en la posición 36 es Gly, Arg, Lys, Glu, Asp o His;
- Xaa en la posición 37 es Pro, Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys o
- 25 se ha suprimido;

Xaa en la posición 38 es Ser, Arg, Lys, Glu, Asp o His o se ha suprimido;
Xaa en la posición 39 es Ser, Arg, Lys, Glu, Asp o His, o se ha suprimido;
Xaa en la posición 40 es Gly, Asp, Glu o Lys o se ha suprimido;
Xaa en la posición 41 es Ala, Phe, Trp, Tyr, Glu, Asp o Lys, o se ha suprimido;
5 Xaa en la posición 42 es Ser, Pro, Lys, Glu o Asp, o se ha suprimido;
Xaa en la posición 43 es Ser, Pro, Glu, Asp o Lys, o se ha suprimido;
Xaa en la posición 44 es Gly, Pro, Glu, Asp o Lys, o se ha suprimido;
y
Xaa en la posición 45 es Ala, Ser, Val, Glu, Asp o Lys, o se ha suprimido;
10 con la condición de que cuando el aminoácido en la posición 37, 38, 39, 40, 41,
42, 43 ó 44 se ha suprimido, después cada aminoácido cadena abajo de este aminoácido
también se ha suprimido.

31. La proteína de fusión heteróloga de las reivindicaciones 27, 28 ó 29 en la
15 que el compuesto GLP – 1 comprende la secuencia de fórmula II (SEC ID N° 3):



Fórmula II (SEC ID N° 3)

en la que:

25 Xaa en la posición 7 es: L - histidina, D – histidina, desaminohistidina, 2 -

aminohistidina, β – hidroxihistidina, homohistidina, α – fluorometilhistidina o α – metilhistidina;

- Xaa en la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser o Thr;
- Xaa en la posición 9 es: Thr, Ser, Arg, Lys, Trp, Phe, Tyr, Glu o His;
- 5 Xaa en la posición 11 es: Asp, Glu, Arg, Thr, Ala, Lys o His;
- Xaa en la posición 12 es: His, Trp, Phe o Tyr;
- Xaa en la posición 16 es: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Tyr, Glu o Ala;
- Xaa en la posición 18 es: His, Pro, Asp, Glu, Arg, Ser, Ala o Lys;
- Xaa en la posición 19 es: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;
- 10 Xaa en la posición 23 es: His, Asp, Lys, Glu, Gln o Arg;
- Xaa en la posición 24 es: Glu, Arg, Ala o Lys;
- Xaa en la posición 26 es: Trp, Tyr, Phe, Asp, Lys, Glu o His;
- Xaa en la posición 27 es: Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;
- Xaa en la posición 30 es: Ala, Glu, Asp, Ser o His;
- 15 Xaa en la posición 31 es: Asp, Glu, Ser, Thr, Arg, Trp o Lys;
- Xaa en la posición 33 es: Asp, Arg, Val, Lys, Ala, Gly o Glu;
- Xaa en la posición 34 es: Glu, Lys o Asp;
- Xaa en la posición 35 es: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o Glu;
- 20 Xaa en la posición 36 es: Thr, Ser, Asp, Trp, Tyr, Phe, Arg, Glu o His;
- Xaa en la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly – Pro o se ha suprimido.

32. La proteína de fusión heteróloga de las reivindicaciones 27, 28 ó 29 en la
- 25 que el compuesto GLP – 1 comprende la secuencia de fórmula III (SEC ID N° 4):

7 8 9 10 11 12 13 14 15 16 17
 Xaa – Xaa – Glu – Gly – Xaa – Xaa – Thr – Ser – Asp – Xaa – Ser –
 18 19 20 21 22 23 24 25 26 27 28
 5 Ser – Tyr – Leu – Glu – Xaa – Xaa – Xaa – Ala – Xaa – Xaa – Phe –
 29 30 31 32 33 34 35 36 37
 Ile – Ala – Xaa – Leu – Xaa – Xaa – Xaa – Xaa – R

Fórmula III (SEC ID N° 4)

en la que:

- 10 Xaa en la posición 7 es: L - histidina, D - histidina, desaminohistidina, 2 - aminohistidina, β - hidroxihistidina, homohistidina, α - fluorometilhistidina o α - metilhistidina;

Xaa en la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser o Thr;

Xaa en la posición 11 es: Asp, Glu, Arg, Thr, Ala, Lys o His;

- 15 Xaa en la posición 12 es: His, Trp, Phe o Tyr;

Xaa en la posición 16 es: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu o Ala;

Xaa en la posición 22 es: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;

Xaa en la posición 23 es: His, Asp, Lys, Glu o Gln;

Xaa en la posición 24 es: Glu, His, Ala o Lys;

- 20 Xaa en la posición 25 es: Asp, Lys; Glu o His;

Xaa en la posición 27 es: Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;

Xaa en la posición 30 es: Ala, Glu, Asp, Ser o His;

Xaa en la posición 33 es: Asp, Arg, Val, Lys, Ala, Gly o Glu;

Xaa en la posición 34 es: Glu, Lys o Asp;

- 25 Xaa en la posición 35 es: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o

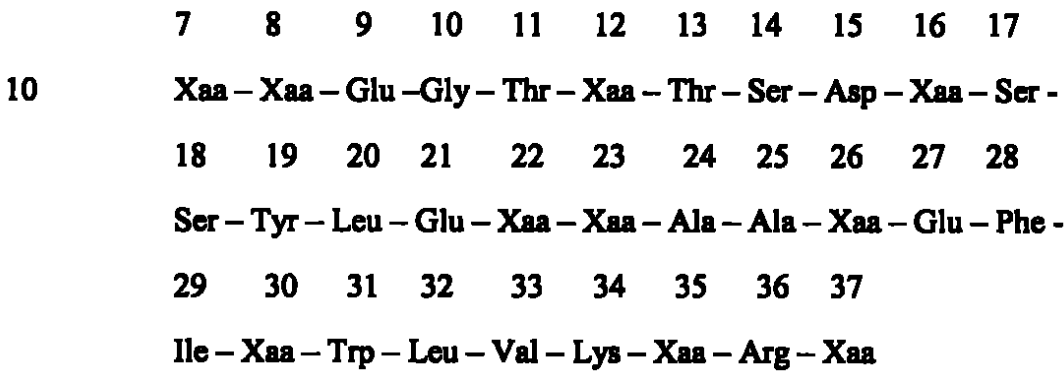
Glu;

Xaa en la posición 36 es: Asp, Glu o His;

Xaa en la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly,
Gly – Pro o se ha suprimido.

5

33. La proteína de fusión heteróloga de las reivindicaciones 27, 28 ó 29 en la
que el compuesto GLP – 1 comprende la secuencia de fórmula IV (SEC ID N° 5):



15 Fórmula IV (SEC ID N° 5)
en la que:

Xaa en la posición 7 es: L - histidina, D - histidina, desaminohistidina, 2 -
aminohistidina, β – hidroxihistidina, homohistidina, α – fluorometilhistidina o α –
metilhistidina;

- 20 Xaa en la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser, o Thr;
Xaa en la posición 12 es: His, Trp, Phe o Tyr;
Xaa en la posición 16 es: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu o Ala;
Xaa en la posición 22 es: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;
Xaa en la posición 23 es: His, Asp, Lys, Glu o Gln;
- 25 Xaa en la posición 26 es: Asp, Lys, Glu o His;

Xaa en la posición 30 es: Ala, Glu, Asp, Ser o His;

Xaa en la posición 35 es: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o Glu;

Xaa en la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly,

5 Gly – Pro o se ha suprimido.

34. La proteína de fusión heteróloga de las reivindicaciones 27, 28 ó 29 en la que el compuesto GLP – 1 comprende la secuencia de V (SEC ID N° 6):

	7	8	9	10	11	12	13	14	15	16	17
	Xaa	Xaa	Glu	Gly	Thr	Phe	Thr	Ser	Asp	Val	Ser
10	18	19	20	21	22	23	24	25	26	27	28
	Ser	Tyr	Leu	Glu	Xaa	Xaa	Ala	Ala	Lys	Xaa	Phe
	29	30	31	32	33	34	35	36	37		
	Ile	Xaa	Trp	Leu	Val	Lys	Gly	Arg	Xaa		

Fórmula V (SEC ID N° 6)

15 en la que:

Xaa en la posición 7 es: L - histidina, D - histidina, desaminohistidina, 2 - aminohistidina, β - hidroxihistidina, homohistidina, α - fluorometilhistidina o α - metilhistidina;

Xaa en la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser o Thr;

20 Xaa en la posición 22 es: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;

Xaa en la posición 23 es: His, Asp, Lys, Glu o Gln;

Xaa en la posición 24 es: Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;

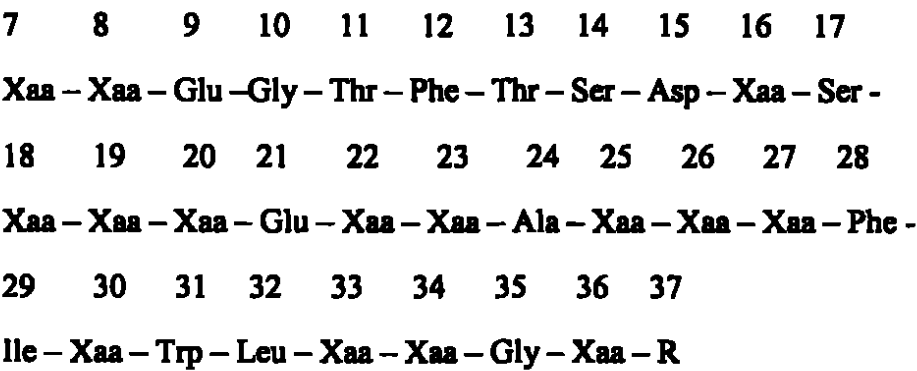
Xaa en la posición 30 es: Ala, Glu, Asp, Ser o His;

Xaa en la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly,

25 Gly – Pro o se ha suprimido.

35. La proteína de fusión heteróloga de las reivindicaciones 27, 28 ó 29 en la que el compuesto GLP – 1 comprende la secuencia de fórmula VIII [SEC ID N° 11]

5



10

Fórmula VIII (SEC ID N° 11)
en la que:

Xaa en la posición 7 es: L - histidina, D - histidina, desaminohistidina, 2 -
15 aminohistidina, β – hidroxihistidina, homohistidina, α – fluorometilhistidina o α –
metilhistidina;

- Xaa en la posición 8 es: Gly, Ala o Val;
- Xaa en la posición 16 es: Leu o Val;
- Xaa en la posición 18 es: Lys o Ser;
- 20 Xaa en la posición 19 es: Gln o Tyr;
- Xaa en la posición 20 es: Met o Leu;
- Xaa en la posición 22 es: Glu o Gln;
- Xaa en la posición 23 es: Glu o Gln;
- Xaa en la posición 25 es: Val o Ala;
- 25 Xaa en la posición 26 es: Arg o Lys;

20

25

Xaa en la posición 27 es: Leu o Glu;

Xaa en la posición 30 es: Glu o Ala;

Xaa en la posición 33 es: Val o Lys;

Xaa en la posición 34 es: Asn o Lys;

5 Xaa en la posición 36 es: Gly o Arg; y

Xaa en la posición 37 es: Gly, Pro, Pro – Ser – Ser – Gly – Ala – Pro – Pro – Pro
– Ser o está ausente.

10 36. La proteína de fusión heteróloga de las reivindicaciones 27 a 35 en la que el compuesto GLP – 1 no tiene más de 6 aminoácidos que son diferentes del correspondiente aminoácido en GLP – 1 (7 – 37) OH, GLP – 1 (7 – 36) OH o exendina – 4.

15 37. La proteína de fusión heteróloga de la reivindicación 36 en la que el compuesto GLP – 1 no tiene más de 5 aminoácidos que son diferentes del correspondiente aminoácido en GLP – 1 (7 – 37) OH, GLP – 1 (7 – 36) OH o exendina – 4.

20 38. La proteína de fusión heteróloga de la reivindicación 37 en la que el compuesto GLP – 1 no tiene más de 4 aminoácidos que son diferentes del correspondiente aminoácido en GLP – 1 (7 – 37) OH, GLP – 1 (7 – 36) OH o exendina – 4.

25 39. La proteína de fusión heteróloga de la reivindicación 38 en la que el compuesto GLP – 1 no tiene más de 3 aminoácidos que son diferentes del correspondiente aminoácido en GLP – 1 (7 – 37) OH, GLP – 1 (7 – 36) OH o exendina – 4.

- 5 40. La proteína de fusión heteróloga de la reivindicación 39 en la que el compuesto GLP – 1 no tiene más de 2 aminoácidos que son diferentes del correspondiente aminoácido en GLP – 1 (7 – 37) OH, GLP – 1 (7 – 36) OH o exendina – 4.
41. La proteína de fusión heteróloga de una cualquiera de la reivindicaciones 27 a 40 en la que Xaa en la posición 8 es glicina o valina.
- 10 42. La proteína de fusión heteróloga de la reivindicación 34 en la que el compuesto GLP – 1 no tiene más de 2 aminoácidos que son diferentes del correspondiente aminoácido en GLP – 1 (7 – 37) OH o GLP – 1 (7 – 36)OH y Xaa en la posición 8 es glicina o valina y Xaa en la posición 30 es alanina, ácido glutámico, ácido aspártico, serina o histidina.
- 15 43. La proteína de fusión heteróloga de la reivindicación 42 en la que Xaa en la posición 30 es ácido glutámico.
- 20 44. La proteína de fusión heteróloga de la reivindicación 34 en la que el compuesto GLP – 1 no tiene más de 2 aminoácidos que son diferentes del correspondiente aminoácido en GLP – 1 (7 – 37) OH o GLP – 1 (7 – 36)OH y Xaa en la posición 8 es glicina o valina y Xaa en la posición 37 es histidina, fenilalanina, tirosina, triptófano, ácido aspártico, ácido glutámico, serina, treonina, arginina o lisina.
- 25 45. La proteína de fusión heteróloga de la reivindicación 44 en la que Xaa en la posición 37 es histidina.

46. La proteína de fusión heteróloga de la reivindicación 34 en la que el compuesto GLP - 1 no tiene más de 2 aminoácidos que son diferentes del correspondiente aminoácido en GLP - 1 (7 - 37) OH o GLP - 1 (7 - 36) y Xaa en la posición 8 es glicina, valina, leucina, isoleucina, serina, treonina o metionina y Xaa en la posición 22 es ácido aspártico, ácido glutámico, lisina, arginina, asparagina, glutamina o histidina.
47. La proteína de fusión heteróloga de la reivindicación 46 en la que Xaa en la posición 22 es lisina o ácido glutámico.
48. La proteína de fusión heteróloga de la reivindicación 34 en la que el compuesto GLP - 1 es Val - 8 - GLP - 1 (7 - 37).
49. La proteína de fusión heteróloga de la reivindicación 34 en la que el compuesto GLP - 1 es Gly - 8 - GLP - 1 (7 - 37).
50. La proteína de fusión heteróloga de una cualquiera de las reivindicaciones 27 a 49 en la que el segundo polipéptido es la porción Fc de una Ig seleccionada del grupo constituido por: IgG1, IgG2, IgG3, IgG4, IgE, IgA, IgD o IgM.
51. La proteína de fusión heteróloga de una cualquiera de las reivindicaciones 27 a 50 en la que el segundo polipéptido es la porción Fc de una Ig seleccionada del grupo constituido por: IgG1, IgG2, IgG3 e IgG4.

52. La proteína de fusión heteróloga de la reivindicación 51 en la que el segundo polipéptido es la porción Fc de una inmunoglobulina IgG1.
53. La proteína de fusión heteróloga de la reivindicación 51 en la que el segundo polipéptido es la porción Fc de una inmunoglobulina IgG4.
54. La proteína de fusión heteróloga de las reivindicaciones 50 a 53 en la que la IgG es humana.
55. La proteína de fusión heteróloga de una cualquiera de las reivindicaciones 27 a 54 en la que la porción Fc comprende los dominios de la región flexible, CH₂ y CH₃.
56. La proteína de fusión heteróloga de la reivindicación 52 en la que el segundo polipéptido tiene la secuencia de SEC ID NO 32.
57. Un polinucleótido que codifica una proteína de fusión heteróloga de una cualquiera de las reivindicaciones 1 a 56.
58. Un vector que comprende el polinucleótido de la reivindicación 57.
59. Una célula huésped que comprende el vector de la reivindicación 58.
60. Una célula huésped que expresa al menos una proteína de fusión heteróloga de una cualquiera de las reivindicaciones 1 a 56.

61. La célula huésped de la reivindicación 60 en la que dicha célula huésped es una célula del ovario de un hámster chino (CHO).
- 5 62. Un procedimiento para producir una proteína de fusión heteróloga que comprende las etapas de transcripción y traducción de un polinucleótido de la reivindicación 57 en condiciones en las que la proteína de fusión heteróloga se expresa en cantidades detectables.
- 10 63. Un procedimiento para normalizar los niveles de glucosa en sangre en un mamífero en necesidad del mismo que comprende la administración de una cantidad terapéuticamente eficaz de una proteína de fusión heteróloga de una cualquiera de las reivindicaciones 1 a 56.
- 15 64. Un procedimiento para tratar un paciente con diabetes mellitus no insulino dependiente que comprende la administración de una cantidad terapéuticamente eficaz de una proteína de fusión heteróloga de una cualquiera de las reivindicaciones 1 a 56.
- 20 65. Un procedimiento para tratar la obesidad que comprende la administración de una cantidad terapéuticamente eficaz de la proteína de fusión heteróloga de una cualquiera de las reivindicaciones 1 a 56.
- 25 66. El uso de una proteína de fusión heteróloga como se reivindica en una cualquiera de las reivindicaciones 1 a 56 para la elaboración de un medicamento con el fin de tratar a pacientes con diabetes mellitus no insulino dependiente.

67. El uso de una proteína de fusión heteróloga como se reivindica en una cualquiera de las reivindicaciones 1 a 56 para la elaboración de un medicamento con el fin de tratar a pacientes con obesidad.

5

68. Una formulación farmacéutica adaptada para el tratamiento de pacientes con diabetes no insulino dependiente que comprende una proteína de fusión heteróloga de una cualquiera de las reivindicaciones 1 a 56.

- 10 69. La proteína de fusión heteróloga de las reivindicaciones 1, 2 ó 3 en la que el compuesto de GLP – 1 comprende la secuencia de aminoácidos de fórmula IX (SEC ID N° 12)

**Xaa₇-Xaa₈-Glu-Gly-Thr-Xaa₁₂-Thr-Ser-Asp-Xaa₁₆-Ser-
Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Gln-Ala-Xaa₂₅-Lys-Xaa₂₇-
Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Lys-Gly-Arg-Xaa₃₇**

- 15 **Fórmula IX (SEC ID N° 12)**

en la que:

Xaa₇ es: L - histidina, D - histidina, desaminohistidina, 2 - aminohistidina, β - hidroxihistidina, homohistidina, α - fluorometilhistidina o α - metilhistidina;

- 20 Xaa₈ es: Ala, Gly, Val, Leu, Ile, Ser o Thr;

Xaa₁₂ es: Phe, Trp o Tyr;

Xaa₁₆ es: Val, Trp, Ile, Leu, Phe o Tyr;

- Xaa₁₈ es: Ser, Trp, Tyr, Phe, Lys, Ile, Leu, Val;**
- Xaa₁₉ es: Tyr, Trp o Phe;**
- Xaa₂₀ es: Leu, Phe, Tyr o Trp;**
- 5 Xaa₂₂ es: Gly, Glu, Asp o Lys;**
- Xaa₂₅ es: Ala, Val, Ile, o Leu;**
- Xaa₂₇ es: Glu, Ile o Ala;**
- Xaa₃₀ es: Ala o Glu;**
- Xaa₃₃ es: Val o Ile; y**
- 10 Xaa₃₇ es: Gly, His, NH₂ ó está ausente.**
- 70. La proteína de fusión heteróloga de la reivindicación 69 en la que el segundo polipéptido es albúmina humana.**
- 15 71. La proteína de fusión heteróloga de la reivindicación 70 en la que el segundo polipéptido tiene la secuencia SEC ID N° 34.**
- 72. La proteína de fusión heteróloga de la reivindicación 69 en la que el segundo polipéptido es un fragmento del extremo N terminal de albúmina.**
- 20 73. La proteína de fusión heteróloga de una cualquiera de las reivindicaciones 27, 28 ó 29 en la que el compuesto GLP – 1 comprende la secuencia de fórmula IX (SEC ID N° 12)**

**Xaa₇-Xaa₈-Glu-Gly-Thr-Xaa₁₂-Thr-Ser-Asp-Xaa₁₆-Ser-
Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Gln-Ala-Xaa₂₅-Lys-Xaa₂₇-
Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Lys-Gly-Arg-Xaa₃₇**

Fórmula IX (SEC ID Nº 12)

en la que:

- 5 Xaa₇ es: L - histidina, D - histidina, desaminohistidina, 2 - aminohistidina, β - hidroxihistidina, homohistidina, α - fluorometilhistidina o α - metilhistidina;**
- Xaa₈ es: Ala, Gly, Val, Leu, Ile, Ser o Thr;**
- Xaa₁₂ es: Phe, Trp o Tyr;**
- Xaa₁₆ es: Val, Trp, Ile, Leu, Phe o Tyr;**
- 10 Xaa₁₈ es: Ser, Trp, Tyr, Phe, Lys, Ile, Leu, Val;**
- Xaa₁₉ es: Tyr, Trp o Phe;**
- Xaa₂₀ es: Leu, Phe, Tyr o Trp;**
- Xaa₂₂ es: Gly, Glu, Asp o Lys;**
- Xaa₂₅ es: Ala, Val, Ile, o Leu;**
- 15 Xaa₂₇ es: Glu, Ile o Ala;**
- Xaa₃₀ es: Ala o Glu;**
- Xaa₃₃ es: Val o Ile; y**
- Xaa₃₇ es: Gly, His, NH₂ ó está ausente.**
- 20 74. La proteína de fusión heteróloga de la reivindicación 73 en la que el segundo polipéptido es la porción Fc de una Ig seleccionada del grupo constituido por IgG1, IgG2, IgG3, IgG4, IgE, IgA, IgD o IgM.**

75. La proteína de fusión heteróloga de las reivindicaciones 73 ó 74 en la que el segundo polipéptido es la porción Fc de una Ig seleccionada del grupo constituido por: IgG1, IgG2, IgG3 e IgG4.

5

76. La proteína de fusión heteróloga de la reivindicación 75 en la que el segundo polipéptido es la porción Fc de una inmunoglobulina IgG1.

77. La proteína de fusión heteróloga de la reivindicación 75 en la que el
10 segundo polipéptido es la porción Fc de una inmunoglobulina IgG4.

78. La proteína de fusión heteróloga de las reivindicaciones 73 a 78 en la que la IgG es humana.

15 79. La proteína de fusión heteróloga de la reivindicación 75 en la que el segundo polipéptido tiene la secuencia de SEC ID N° 32.

80. La proteína de fusión heteróloga de una cualquier de las reivindicaciones 69 a 79 en la que el compuesto GLP – 1 se selecciona del grupo constituido por

				5					10					15
Ala	Glu	Pro	Lys	Ser	Cys	Asp	Lys	Thr	His	Thr	Cys	Pro	Pro	Cys
				20					25					30
Pro	Ala	Pro	Glu	Lys	Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro
				35					40					45
Lys	Pro	Lys	Asp	Thr	Lys	Met	Ile	Ser	Arg	Thr	Pro	Glu	Val	Thr
				50					55					60
Cys	Val	Val	Val	Asp	Val	Ser	His	Glu	Asp	Pro	Glu	Val	Lys	Phe
				65					70					75
Asn	Trp	Tyr	Val	Asp	Gly	Val	Glu	Val	His	Asn	Ala	Lys	Thr	Lys
				80					85					90
Pro	Arg	Glu	Glu	Gln	Tyr	Asn	Ser	Thr	Tyr	Arg	Val	Val	Ser	Val
				95					100					105
Leu	Thr	Val	Leu	His	Gln	Asp	Trp	Leu	Asn	Gly	Lys	Glu	Tyr	Lys
				110					115					120
Cys	Lys	Val	Ser	Asn	Lys	Ala	Leu	Pro	Ala	Pro	Ile	Glu	Lys	Thr
				125					130					135
Ile	Ser	Lys	Ala	Lys	Gly	Gln	Pro	Arg	Glu	Pro	Gln	Val	Tyr	Thr
				140					145					150
Leu	Pro	Pro	Ser	Arg	Glu	Glu	Met	Thr	Lys	Asn	Gln	Val	Ser	Leu
				155					160					165
Thr	Cys	Leu	Val	Lys	Gly	Phe	Tyr	Pro	Ser	Asp	Ile	Ala	Val	Glu
				170					175					180
Trp	Glu	Ser	Asn	Gly	Gln	Pro	Glu	Asn	Asn	Tyr	Lys	Thr	Thr	Pro
				185					190					195
Pro	Val	Leu	Asp	Ser	Asp	Gly	Ser	Phe	Phe	Leu	Tyr	Ser	Lys	Leu
				200					205					210
Thr	Val	Asp	Lys	Ser	Arg	Trp	Gln	Gln	Gly	Asn	Val	Phe	Ser	Cys
				215					220					225
Ser	Val	Met	His	Glu	Ala	Leu	His	Asn	His	Tyr	Thr	Gln	Lys	Ser
				230										
Leu	Ser	Leu	Ser	Pro	Gly	Lys			[SEC ID N° 32]					

Fig. 2

5					10					15							
Asp	Ala	His	Lys	Ser	Glu	Val	Ala	His	Arg	Phe	Lys	Asp	Leu	Gly	Glu	Glu	Asn
20					25					30					35		
Phe	Lys	Ala	Leu	Val	Leu	Ile	Ala	Phe	Ala	Gln	Tyr	Leu	Gln	Gln	Cys	Pro	Phe
	40						45					50					
Glu	Asp	His	Val	Lys	Leu	Val	Asn	Glu	Val	Thr	Glu	Phe	Ala	Lys	Thr	Cys	Val
55					60				65					70			
Ala	Asp	Glu	Ser	Ala	Glu	Asn	Cys	Asp	Lys	Ser	Leu	His	Thr	Leu	Phe	Gly	Asp
	75					80					85					90	
Lys	Leu	Cys	Thr	Val	Ala	Thr	Leu	Arg	Glu	Thr	Tyr	Gly	Glu	Met	Ala	Asp	Cys
		95						100					105				
Cys	Ala	Lys	Gln	Glu	Pro	Glu	Arg	Asn	Glu	Cys	Phe	Leu	Gln	His	Lys	Asp	Asp
110					115					120						125	
Asn	Pro	Asn	Leu	Pro	Arg	Leu	Val	Arg	Pro	Glu	Val	Asp	Val	Met	Cys	Thr	Ala
	130						135					140					
Phe	His	Asp	Asn	Glu	Glu	Thr	Phe	Leu	Lys	Lys	Tyr	Leu	Tyr	Glu	Ile	Ala	Arg
145					150					155					160		
Arg	His	Pro	Tyr	Phe	Tyr	Ala	Pro	Glu	Leu	Leu	Phe	Phe	Ala	Lys	Arg	Tyr	Lys
	165					170					175					180	
Ala	Ala	Phe	Thr	Glu	Cys	Cys	Gln	Ala	Ala	Asp	Lys	Ala	Ala	Cys	Leu	Leu	Pro
		185						190					195				
Lys	Leu	Asp	Glu	Leu	Arg	Asp	Glu	Gly	Lys	Ala	Ser	Ser	Ala	Lys	Gln	Arg	Leu
200					205					210						215	
Lys	Cys	Ala	Ser	Leu	Gln	Lys	Phe	Gly	Glu	Arg	Ala	Phe	Lys	Ala	Trp	Ala	Val
	220						225					230					
Ala	Arg	Leu	Ser	Gln	Arg	Phe	Pro	Lys	Ala	Glu	Phe	Ala	Glu	Val	Ser	Lys	Leu
235					240					245					250		
Val	Thr	Asp	Leu	Thr	Lys	Val	His	Thr	Glu	Cys	Cys	His	Gly	Asp	Leu	Leu	Glu
	255					260						265					270
Cys	Ala	Asp	Asp	Arg	Ala	Asp	Leu	Ala	Lys	Tyr	Ile	Cys	Glu	Asn	Gln	Asp	Ser
		275						280					285				
Ile	Ser	Ser	Lys	Leu	Lys	Glu	Cys	Cys	Glu	Lys	Pro	Leu	Leu	Glu	Lys	Ser	His
290					295					300						305	
Cys	Ile	Ala	Glu	Val	Glu	Asn	Asp	Glu	Met	Pro	Ala	Asp	Leu	Pro	Ser	Leu	Ala
	310						315					320					
Ala	Asp	Phe	Val	Glu	Ser	Lys	Asp	Val	Cys	Lys	Asn	Tyr	Ala	Glu	Ala	Lys	Asp
325					330					335					340		
Val	Phe	Leu	Gly	Met	Phe	Leu	Tyr	Glu	Tyr	Ala	Arg	Arg	His	Pro	Asp	Tyr	Ser
	345					350					355					360	
Val	Val	Leu	Leu	Leu	Arg	Leu	Ala	Lys	Thr	Tyr	Glu	Thr	Thr	Leu	Glu	Lys	Cys
		365						370					375				
Cys	Ala	Ala	Ala	Asp	Pro	His	Glu	Cys	Tyr	Ala	Lys	Val	Phe	Asp	Glu	Phe	Lys
380					385					390						395	
Pro	Leu	Val	Glu	Glu	Pro	Gln	Asn	Leu	Ile	Lys	Gln	Asn	Cys	Glu	Leu	Phe	Glu
	400						405					410					
Asn	Leu	Gly	Glu	Tyr	Lys	Phe	Gln	Asn	Ala	Leu	Leu	Val	Arg	Tyr	Thr	Lys	Lys
415					420					425					430		
Val	Pro	Gln	Val	Ser	Thr	Pro	Thr	Leu	Val	Glu	Val	Ser	Arg	Asn	Leu	Gly	Lys
	435					440						445				450	
Val	Gly	Ser	Lys	Cys	Cys	Lys	His	Pro	Glu	Ala	Lys	Arg	Met	Pro	Cys	Ala	Glu
		455						460					465				
Asp	Tyr	Leu	Ser	Val	Val	Leu	Asn	Gln	Leu	Cys	Val	Leu	His	Glu	Lys	Thr	Pro
470					475					480						485	
Val	Ser	Asp	Arg	Val	Thr	Lys	Cys	Cys	Thr	Glu	Ser	Leu	Val	Asn	Arg	Arg	Pro

Fig. 2 continuación

										490				495				500			
Cys	Phe	Ser	Ala	Leu	Glu	Val	Asp	Glu	Thr	Tyr	Val	Pro	Lys	Glu	Phe	Asn	Ala				
505					510					515					520						
Glu	Thr	Phe	Thr	Phe	His	Ala	Asp	Ile	Cys	Thr	Leu	Ser	Glu	Lys	Glu	Arg	Gln				
		525					530					535					540				
Ile	Lys	Lys	Gln	Thr	Ala	Leu	Val	Glu	Leu	Val	Lys	His	Lys	Pro	Lys	Ala	Thr				
				545					550					555							
Lys	Glu	Gln	Leu	Lys	Ala	Val	Met	Asp	Asp	Phe	Ala	Ala	Phe	Val	Glu	Lys	Cys				
	560					565				570					575						
Cys	Lys	Ala	Asp	Asp	Lys	Glu	Thr	Cys	Phe	Ala	Glu	Glu	Gly	Lys	Lys	Leu	Val				
			580					585													
Ala	Ala	Ser	Gln	Ala	Ala	Leu	Gly	Leu													

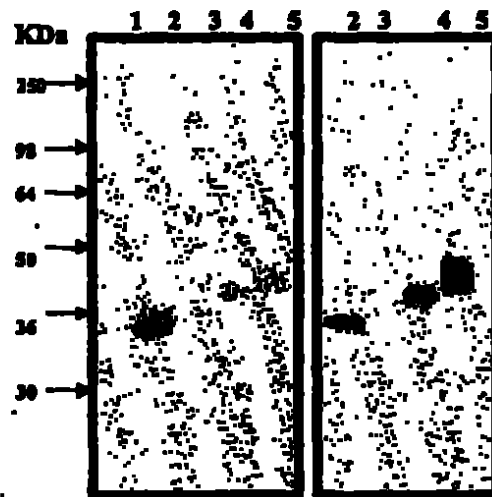
[SEC ID N° 34]

250
347

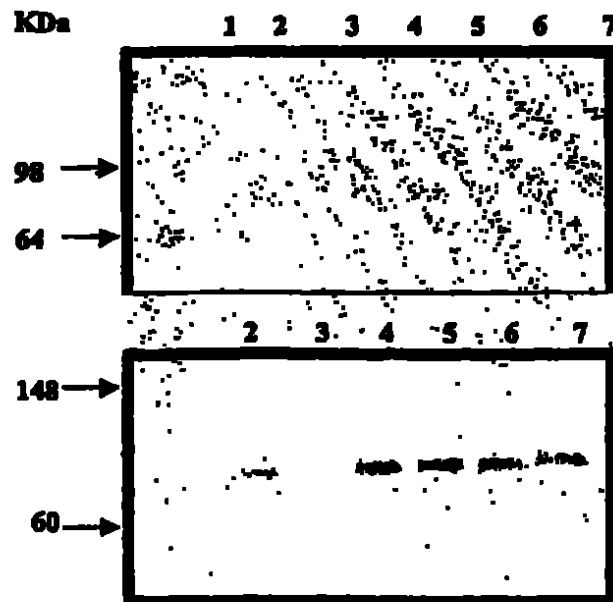
4/17

Fig. 3

A.



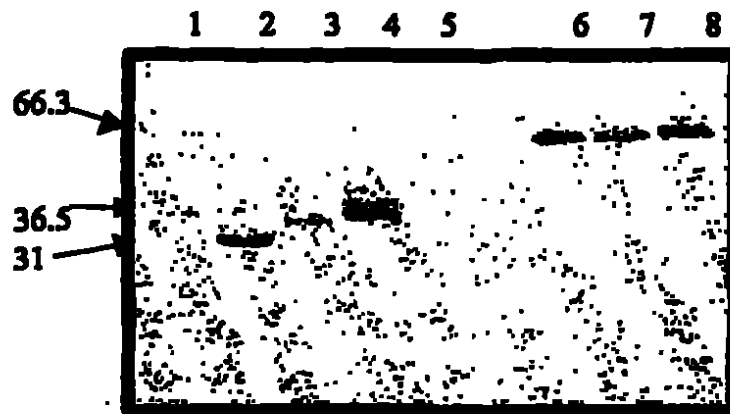
B.



35T
348

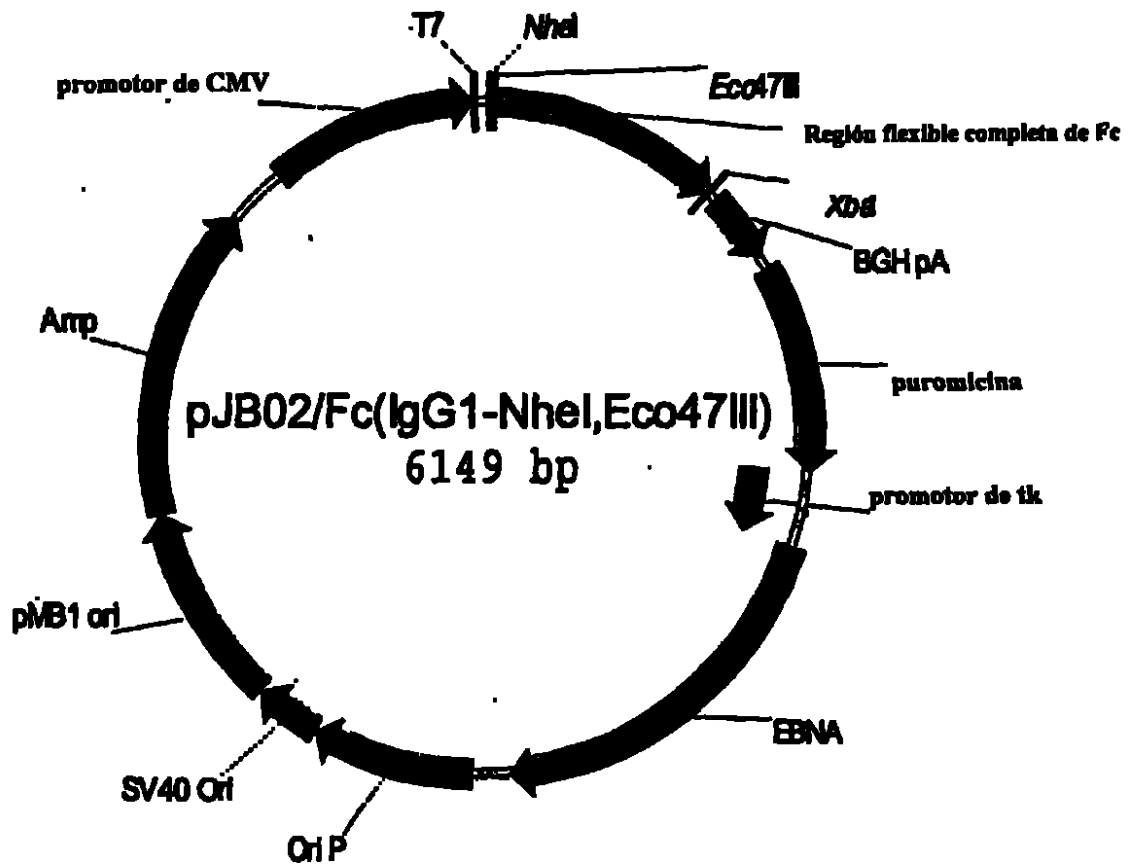
5/17

Fig. 4



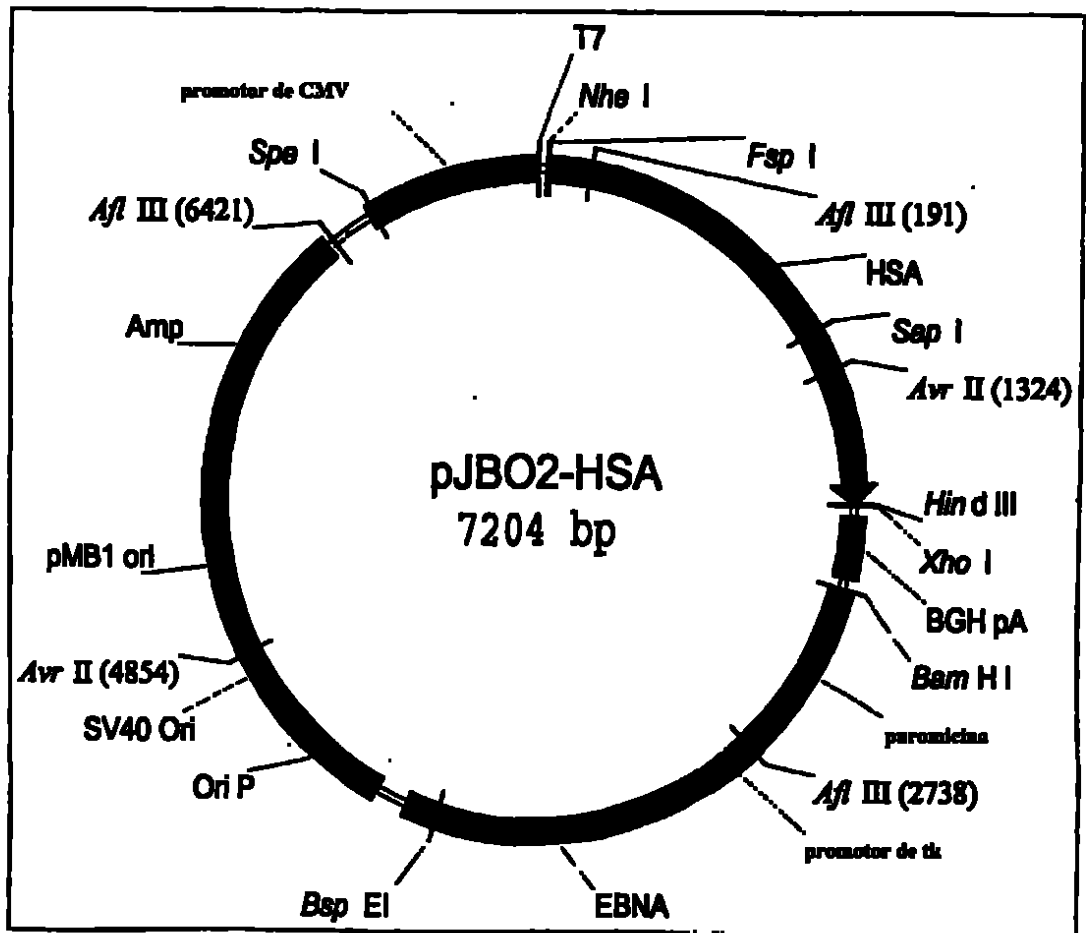
6/17

Fig. 5



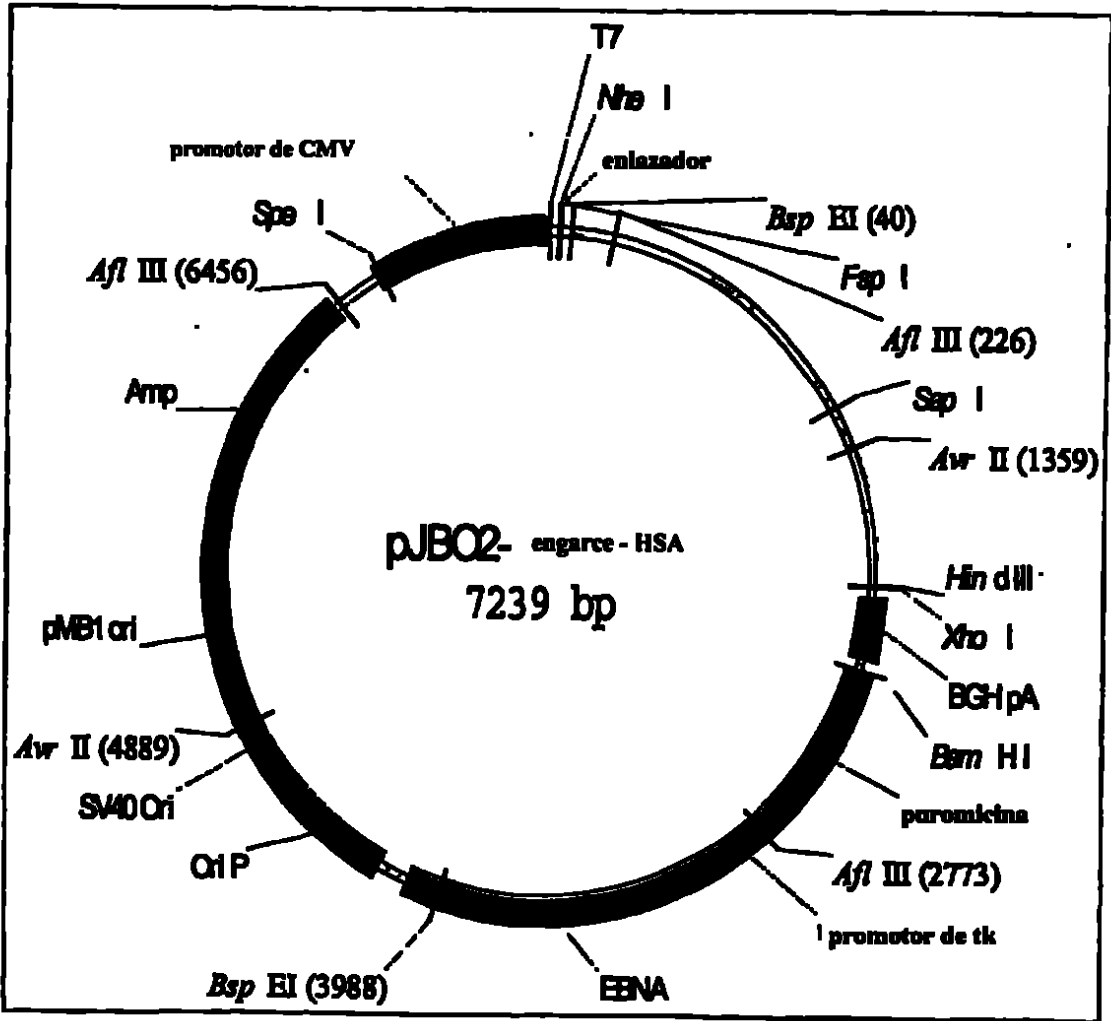
7/17

Fig. 6



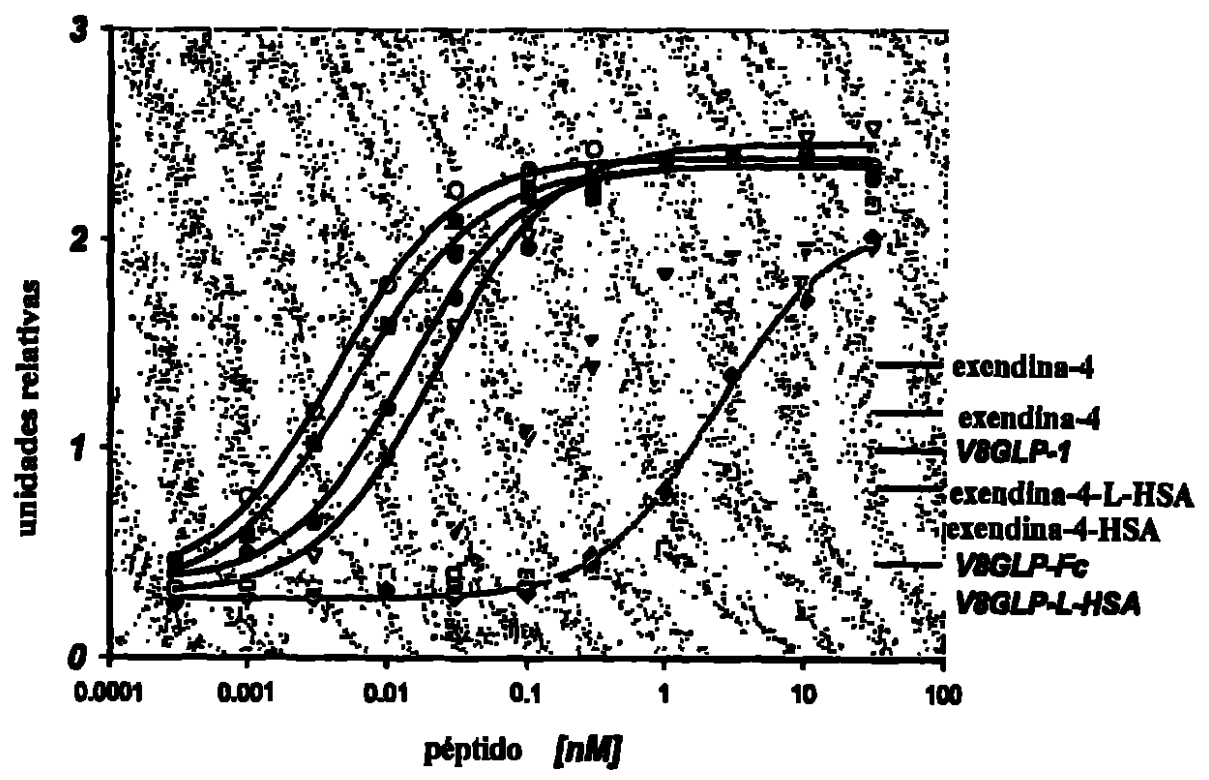
8/17

Fig. 7



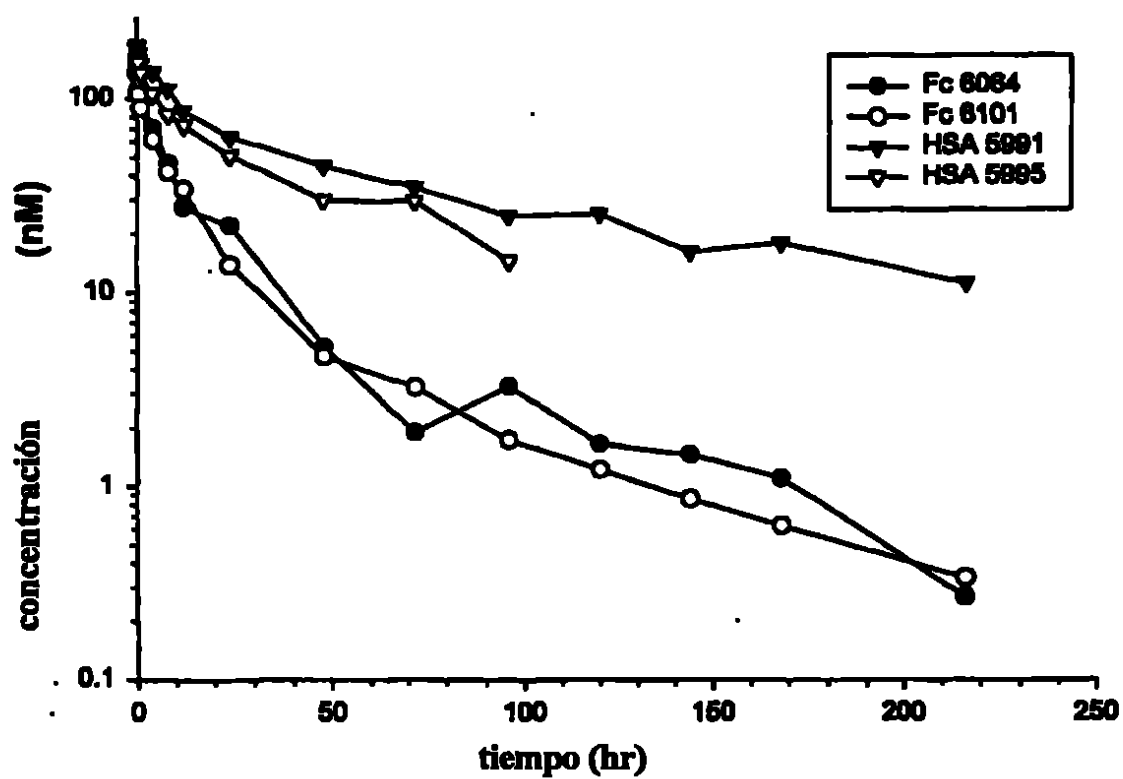
9/17

Fig. 8



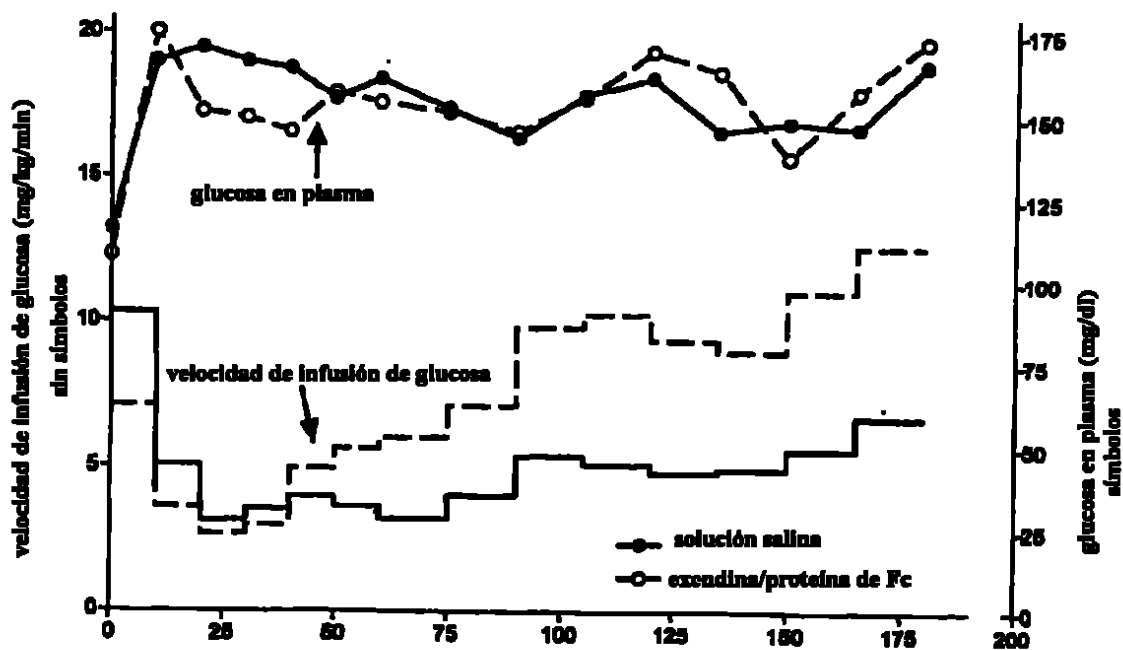
10/17

Fig. 9



11/17

A.

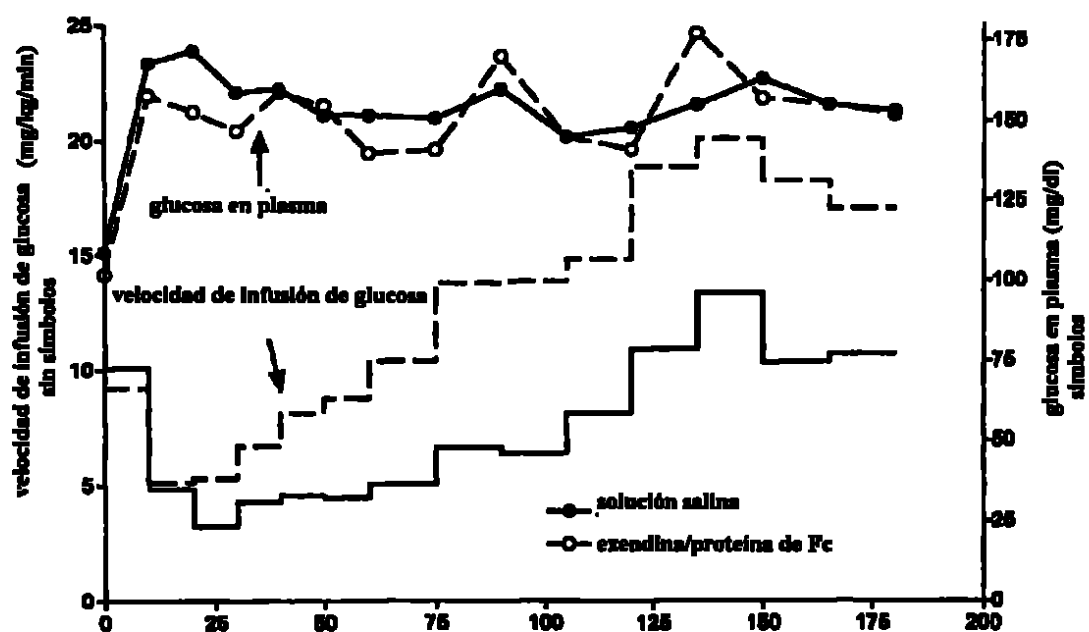


Tiempo desde que comienza la pinza hiperglucémica (min);
inyección de proteína de fusión 2 horas antes del comienzo de la
pinza

12/17

Fig. 10

B.

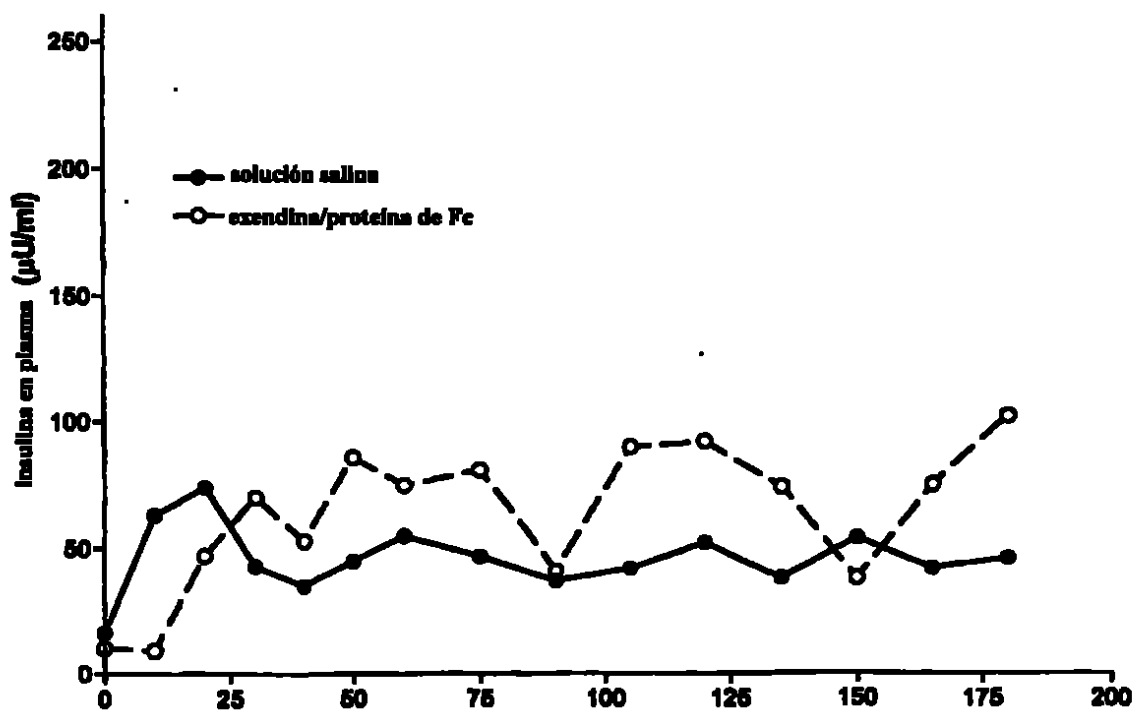


Tiempo desde que comienza la pinza hiperglucémica (min); inyección de proteína de fusión 2 horas antes del comienzo de la pinza

13/17

Fig. 11

A.

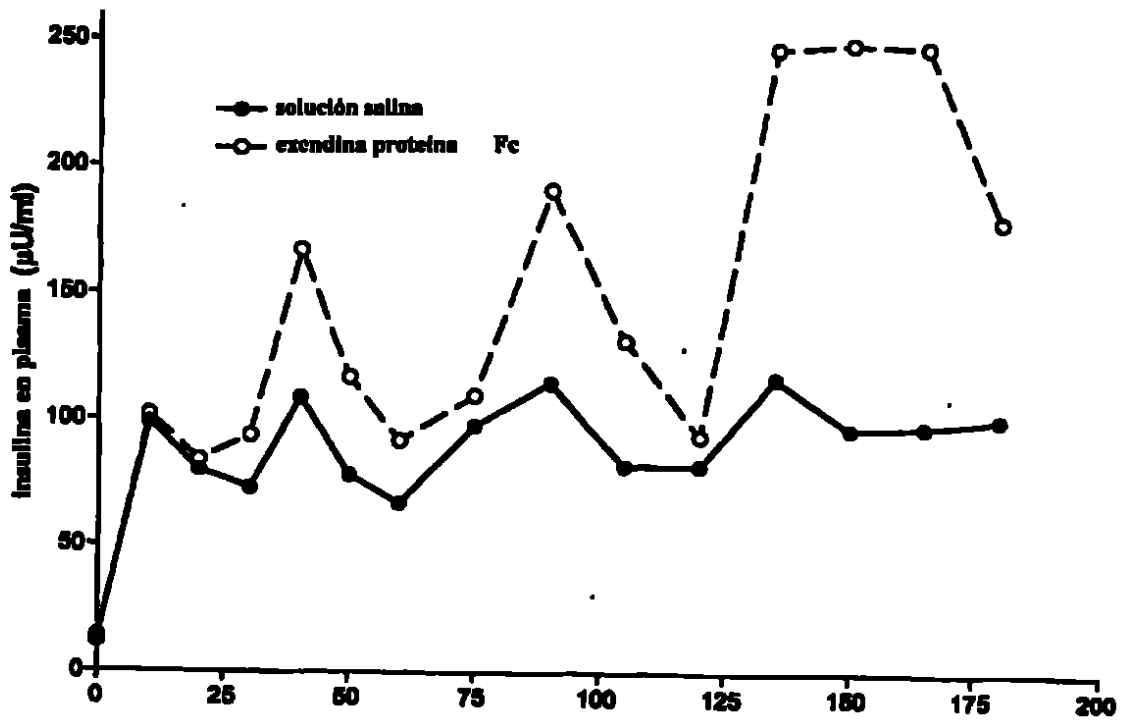


Tiempo desde que comienza la pinza hiperglucémica (min); inyección de proteína de fusión 2 horas antes del comienzo de la pinza

14/17

Fig. 11

B.



Tiempo desde que comienza la pinza hiperglucémica (min); inyección de proteína de fusión 2 horas antes del comienzo de la pinza

Fig. 12

1
GAGCCCAAATCTTGTGACAAAACCTCACACATGCCACCGTGCCAGCACC
50
TGAACCTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCAAACCCAAGG
100
ACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGAC
150
GTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGT
200
GGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCA
250
CGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAAT
300
GGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCAGCCCCCAT
350
CGAGAAAACCATCTCCAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGT
400
ACACCCTGCCCCCATCCCGGGAGGAGATGACCAAGAACCAGGTGAGCCTG
450
ACCTGCCTGGTCAAAGGCTTCTATCCAGCGACATCGCCGTGGAGTGGGA
500
GAGCAATGGGCAGCCGGAGAACAACCTACAAGACCACGCCTCCCGTGCTGG
550
ACTCCGACGGCTCCTTCTTCTCTATAGCAAGCTCACCGTGGACAAGAGC
600
AGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCT
650
GCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAT
700
AGT [SEC ID N° 33]

Fig. 13

1
GATGCGCACAAGAGTGAGGTTGCTCATCGGTTTAAAGATTGGGAGAAGA
50
AAATTTCAAAGCCITGGTGTGATTGCCTTTGCTCAGTATCTTCAGCAQT
100
GTCCATTTGAAGATCATGTAAAATTAGTGAATGAAGTAACTGAATTTGCA
150
AAAACATGTGTTGCTGATGAGTCAGCTGAAAATTGTGACAAATCACTTCA
200
TACCCTTTTTGGAGACAAATTATGCACAGTTGCAACTCTTCGTGAAACCT
250
ATGGTGAAATGGCTGACTGCTGTGCAAAACAAGAACCTGAGAGAAATGAA
300
TGCTTCTTGCAACACAAAGATGACAACCCAAACCTCCCCGATTGGTGAG
350
ACCAGAGGTTGATGTGATGTGCACTGCTTTTCATGACAATGAAGAGACAT
400
TTTTGAAAAAATACTTATATGAAATTGCCAGAAGACATCCTTACTTTTAT
450
GCCCCGGAATCCTTTTCTTTGCTAAAAGGTATAAAGCTGCTTTTACAGA
500
ATGTTGCCAAGCTGCTGTATAAAGCTGCCTGCCTGTTGCCAAGCTCGATG
550
AACTTCGGGATGAAGGGAAGGCTTCGTCTGCCAAACAGAGACTCAAGTGT
600
GCCAGTCTCCAAAATTTGGAGAAAGAGCTTTCAAAGCATGGGCAGTAGC
650
TCGCCTGAGCCAGAGATTTCCCAAAGCTGAGTTTGCAGAAGTTTCCAAGT
700
TAGTGACAGATCTTACCAAAGTCCACACGGAATGCTGCCATGGAGATCTG
750
CTTGAATGTGCTGATGACAGGGCGGACCTTGCCAAGTATATCTGTGAAAA
800
TCAAGATTGATCTCCAGTAACTGAAGGAATGCTGTGAAAAACCTCTGT
850
TGAAAAAATCCCACATGCAATTGCCGAAGTGAAAAATGATGAGATGCCTGCT
900
GACTTGCCCTTCATTAGCTGCTGATTTTGTGAAAGTAAGGATGTTTGCAA
950
AACTATGCTGAGGCAAAGGATGTCTTCCTGGGCATGTTTTTGTATGAAT
1000
ATGCAAGAAGGCATCCTGATTACTCTGTGCTGCTGCTGCTGAGACTTGCC

Fig. 13 continuación

1050
AAGACATATGAAACCACTCTAGAGAAGTGCTGTGCCGCTGCAGATCCTCA
1100
TGAATGCTATGCCAAAGTGTTGATGAATTTAAACCTCTTGTGGAAGAGC
1150
CTCAGAATTTAATCAAACAAAATTGTGAGCTTTTGTGAGCAGCTTGGAGAG
1200
TACAAATTCCAGAATGCGCTATTAGTTCGTTACACCAAGAAAGTACCCCA
1250
AGTGTCAACTCCAACCTCTTGTAGAGGTCTCAAGAAACCTAGGAAAAGTGG
1300
GCAGCAAATGTTGTAAACATCCTGAAGCAAAAAGAATGCCCTGTGCAGAA
1350
GACTATCTATCCGTGGTCCTGAACCAGTTATGTGTGTTGCATGAGAAAAC
1400
GCCAGTAAGTGACAGAGTCACCAAATGCTGCACAGAATCCTTGGTGAACA
1450
GGCGACCATGCTTTTCAGCTCTGGAAGTCGATGAAACATACGTTCCCAA
1500
GAGTTTAATGCTGAAACATTACCTTCCATGCAGATATATGCACACTTTC
1550
TGAGAAGGAGAGACAAATCAAGAAACAACTGCACCTTGTTGAGCTCGTGA
1600
AACACAAGCCCAAGGCAACAAAAGAGCAACTGAAAGCTGTTATGGATGAT
1650
TTCGCAGCTTTTGTAGAGAAGTGCTGCAAGGCTGACGATAAGGAGACCTG
1700
CTTTGCCGAGGAGGGTAAAAAAGCTTGTGCTGCAAGTCAAGCTGCCTTAG
1750
GCTTATAATGAC [SEC ID N° 35]

LISTA DE SECUENCIAS

- <110>

Eli Lilly y Compañía
- <120>

proteínas de fusión de GLP - 1
- <130>

x-13991
- <150>

US 60/251,954
- <151>

2000-06-12
- <160>

35
- <170>

patente en versión 3.1
- <210>

1
- <211>

31
- <212>

PRT
- <213>

Homo sapiens
- <400>

1
- His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly

1 5 10 15
- Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly

20 25 30
- <210>

2
- <211>

39
- <212>

PRT
- <213>

Secuencia artificial
- <220>
- <223>

construcción sintética
- <220>
- <221>

MISC_característica

- <222> (2) . . (2)
<223> Xaa en la posición 2 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (3) . . (3)
<223> Xaa en la posición 3 es Glu, Asp o Lys;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (5) . . (5)
<223> Xaa en la posición 5 es Thr, Ala, Gly, Ser, Leu, Ile, Val, Glu, Asp o Lys;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (8) . . (8)
<223> Xaa en la posición 8 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp o Lys;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (10) . . (10)
<223> Xaa en la posición 10 es Val, Ala, Gly, Ser, Thr, Leu, Ile, Tyr, Glu, Asp o Lys;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (11) . . (11)
<223> Xaa en la posición 11 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp o Lys;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (12) . . (12)

<223> Xaa en la posición 12 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, Lys, Trp o Tyr;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (13) . .(13)

<223> Xaa en la posición 13 es Tyr, Phe, Trp, Glu, Asp, Gln o Lys;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (14) . .(14)

<223> Xaa en la posición 14 es Leu, Ala, Gly, Ser, Thr, Ile, Val, Glu, Asp, Met, Lys, Trp o Tyr;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (15) . .(15)

<223> Xaa en la posición 15 es Glu, Asp o Lys;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (16) . .(16)

<223> Xaa en la posición 16 es Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp, Trp o Lys;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (17) . .(17)

<223> Xaa en la posición 17 es Gln, Asn, Arg, Glu, Asp o Lys;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (18) . .(18)

<223> Xaa en la posición 18 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Arg, Glu, Asp, o Lys;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (19) . .(19)

<223> Xaa en la posición 19 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (20) . .(20)

<223> Xaa en la posición 20 es Lys, Arg, Gln, Glu, Asp o His;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (21) . .(21)

<223> Xaa en la posición 21 es Leu, Glu, Asp o Lys;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (24) . .(24)

<223> Xaa en la posición 24 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp, o Lys;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (25) . .(25)

<223> Xaa en la posición 25 es Trp, Phe, Tyr, Glu, Asp o Lys;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (26) . .(26)

<223> Xaa en la posición 26 es Leu, Gly, Ala, Ser, Thr, Ile, Val, Glu, Asp, o Lys;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (27) . .(27)

<223> Xaa en la posición 27 es Val, Gly, Ala, Ser, Thr, Leu, Ile, Glu, Asp o Lys;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (28) . .(28)

<223> Xaa en la posición 28 es Asn, Lys, Arg, Glu, Asp o His;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (29) . .(29)

<223> Xaa en la posición 29 es Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (30) . .(30)

<223> Xaa en la posición 30 es Gly, Arg, Lys, Glu, Asp o His;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (31) . .(31)

<223> Xaa en la posición 31 es Pro, Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys o se ha suprimido;

<220>

<221> MISC_CHARACTERÍSTICA

- <222> (32) . .(32)
- <223> Xaa en la posición 32 es Ser, Arg, Lys, Glu, Asp o His o se ha suprimido;
- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (33) . .(33)
- <223> Xaa en la posición 33 es Ser, Arg, Lys, Glu, Asp o His o se ha suprimido;
- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (34) . .(34)
- <223> Xaa en la posición 34 es Gly, Asp, Glu, o Lys o se ha suprimido;
- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (35) . .(35)
- <223> Xaa en la posición 35 es Ala, Phe, Trp, Tyr, Glu, Asp o Lys o se ha suprimido;
- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (36) . .(36)
- <223> Xaa en la posición 36 es Ser, Pro, Lys, Glu o Asp o se ha suprimido;
- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (37) . .(37)
- <223> Xaa en la posición 37 es Ser, Pro, Glu, Asp o Lys o se ha suprimido;
- <220>
- <221> MISC_CHARACTERÍSTICA

- <222> (38) . .(38)
- <223> Xaa en la posición 38 es Gly, Pro, Glu, Asp o Lys o se ha suprimido;
- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (39) . .(39)
- <223> Xaa en la posición 39 es Ala, Ser, Val, Glu, Asp o Lys o se ha suprimido;

<400> 2
His Xaa Xaa Gly Xaa Phe Thr Xaa Asp Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15
Xaa Xaa Xaa Xaa Xaa Phe Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30
Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35

- <210> 3
- <211> 32
- <212> PRT
- <213> Secuencia artificial

- <220>
- <223> construcción sintética
- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (1) . . (1)
- <223> Xaa en la posición 1 es L – histidina, D – histidina o se ha suprimido;

- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (2) . .(2)
- <223> Xaa en la posición 2 es Gly, Ala, Val, Leu, Ile, Ser o Thr;
- <220>

- <221> MISC_CHARACTERÍSTICA
<222> (3) . .(3)
<223> Xaa en la posición 3 es Thr, Ser, Arg, Lys, Trp, Phe, Tyr, Glu o His;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (5) . .(5)
<223> Xaa en la posición 5 es Asp, Glu, Arg, Thr, Ala, Lys o His;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (6) . .(6)
<223> Xaa en la posición 6 es His, Trp, Phe, o Tyr;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (10) . .(10)
<223> Xaa en la posición 10 es Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Tyr, Glu o Ala;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (12) . .(12)
<223> Xaa en la posición 12 es His, Pro, Asp, Glu, Arg, Ser, Ala o Lys;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (13) . .(13)
<223> Xaa en la posición 13 es Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;
- <220>

- <221> MISC_CHARACTERÍSTICA
<222> (17) . .(17)
<223> Xaa en la posición 17 es His, Asp, Lys, Glu, Gln o Arg;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (18) . .(18)
<223> Xaa en la posición 18 es Glu, Arg, Ala o Lys;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (20) . .(20)
<223> Xaa en la posición 20 es Trp, Tyr, Phe, Asp, Lys, Glu o His;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (21) . .(21)
<223> Xaa en la posición 21 es Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (24) . .(24)
<223> Xaa en la posición 24 es Ala, Glu, Asp, Ser o His;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (25) . .(25)
<223> Xaa en la posición 25 es Asp, Glu, Ser, Thr, Arg, Trp o Lys;
- <220>
<221> MISC_CHARACTERÍSTICA

375
370

- <222> (27) . .(27)
<223> Xaa en la posición 27 es Asp, Arg, Val, Lys, Ala, Gly o Glu;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (28) . .(28)
<223> Xaa en la posición 28 es Glu, Lys o Asp;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (29) . .(29)
<223> Xaa en la posición 29 es Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o Glu;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (30) . .(30)
<223> Xaa en la posición 30 es Thr, Ser, Asp, Trp, Tyr, Phe, Arg, Glu o His;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (31) . .(31)
<223> Xaa en la posición 31 es Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly o se ha suprimido;
- <220>
<221> MISC_CHARACTERÍSTICA
<222> (32) . .(32)
<223> Xaa en la posición 32 es Pro o se ha suprimido;

<400> 3

Xaa	Xaa	Xaa	Gly	Xaa	Xaa	Thr	Ser	Asp	Xaa	Ser	Xaa	Xaa	Leu	Glu	Gly
1				5					10					15	
Xaa	Xaa	Ala	Xaa	Xaa	Phe	Ile	Xaa	Xaa	Leu	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa
				20					25					30	

<210> 4
<211> 32
<212> PRT
<213> Secuencia artificial

<220>
<223> construcción sintética

<220>
<221> MISC_CHARACTERÍSTICA
<222> (1) . . (1)
<223> Xaa en la posición 1 es L – histidina, D – histidina o se ha suprimido;

<220>
<221> MISC_CHARACTERÍSTICA
<222> (2) . . (2)
<223> Xaa en la posición 2 es Gly, Ala, Val, Leu, Ile, Ser o Thr;

<220>
<221> MISC_CHARACTERÍSTICA
<222> (5) . . (5)
<223> Xaa en la posición 5 es Asp, Glu, Arg, Thr, Ala, Lys o His;

<220>
<221> MISC_CHARACTERÍSTICA
<222> (6) . . (6)
<223> Xaa en la posición 6 es His, Trp, Phe, o Tyr;

<220>
<221> MISC_CHARACTERÍSTICA
<222> (10) . . (10)
<223> Xaa en la posición 10 es Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu o Ala;

375
322

- <220>

<221> MISC_CHARACTERÍSTICA

<222> (16) . . (16)

<223> Xaa en la posición 16 es Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;
- <220>

<221> MISC_CHARACTERÍSTICA

<222> (17) . . (17)

<223> Xaa en la posición 17 es His, Asp, Lys, Glu o Gln;
- <220>

<221> MISC_CHARACTERÍSTICA

<222> (18) . . (18)

<223> Xaa en la posición 18 es Glu, His, Ala o Lys;
- <220>

<221> MISC_CHARACTERÍSTICA

<222> (19) . .(19)

<223> Xaa en la posición 19 es Asp, Lys, Glu o His;
- <220>

<221> MISC_CHARACTERÍSTICA

<222> (21) . .(21)

<223> Xaa en la posición 21 es Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;
- <220>

<221> MISC_CHARACTERÍSTICA

<222> (24) . .(24)

<223> Xaa en la posición 24 es Ala, Glu, Asp, Ser o His;

37(373

- <220>

<221> MISC_CHARACTERÍSTICA

<222> (27) . . (27)

<223> Xaa en la posición 27 es Asp, Arg, Val, Lys, Ala, Gly o Glu;
- <220>

<221> MISC_CHARACTERÍSTICA

<222> (28) . . (28)

<223> Xaa en la posición 28 es Glu, Lys o Asp;
- <220>

<221> MISC_CHARACTERÍSTICA

<222> (29) . . (29)

<223> Xaa en la posición 29 es Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o Glu;
- <220>

<221> MISC_CHARACTERÍSTICA

<222> (30) . . (30)

<223> Xaa en la posición 30 es Arg, Glu o His;
- <220>

<221> MISC_CHARACTERÍSTICA

<222> (31) . . (31)

<223> Xaa en la posición 31 es Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, o se ha suprimido;
- <220>

<221> MISC_CHARACTERÍSTICA

<222> (32) . . (32)

<223> Xaa en la posición 32 es Pro o se ha suprimido;

<400> 4

Xaa Xaa Glu Gly Xaa Xaa Thr Ser Asp Xaa Ser Ser Tyr Leu Glu Xaa

1 5 10 15

Xaa Xaa Xaa Lys Xaa Phe Ile Xaa Trp Leu Xaa Xaa Xaa Xaa Xaa Xaa

20 25 30

<210> 5

<211> 32

<212> PRT

<213> Secuencia artificial

<220>

<223> construcción sintética

<220>

<221> MISC_CHARACTERÍSTICA

<222> (1) . . (1)

<223> Xaa en la posición 1 es L – histidina, D – histidina o se ha suprimido;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (2) . . (2)

<223> Xaa en la posición 2 es Gly, Ala, Val, Leu, Ile, Ser, Met o Thr;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (6) . . (6)

<223> Xaa en la posición 6 es His, Trp, Phe, o Tyr;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (10) . . (10)

<223> Xaa en la posición 10 es Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu o Ala;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (16) . . (16)

<223> Xaa en la posición 16 es Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (17) . .(17)

<223> Xaa en la posición 17 es His, Asp, Lys, Glu o Gln;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (20) . .(20)

<223> Xaa en la posición 20 es Asp, Lys, Glu o His;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (24) . .(24)

<223> Xaa en la posición 24 es Ala, Glu, Asp, Ser o His;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (29) . .(29)

<223> Xaa en la posición 29 es Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly Pro, His o Glu;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (31) . .(31)

<223> Xaa en la posición 31 es Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly o se ha suprimido;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (32) . .(32)

<223> Xaa en la posición 32 es Pro o se ha suprimido;

<400> 5

Xaa Xaa Glu Gly Thr Xaa Thr Ser Asp Xaa Ser Ser Tyr Leu Glu Xaa
1 5 10 15

Xaa Ala Ala Xaa Glu Phe Ile Xaa Trp Leu Val Lys Xaa Arg Xaa Xaa
20 25 30

$\langle 210 \rangle$

6

 $\langle 211 \rangle$

32

<212>

PRT

◁213> Secuencia artificial

<220>

<223>

construcción sintética

<220>

<221>

MISC_CARACTERÍSTICA

<222>

(1) . . (1)

<223>

Xaa en la posición 1 es L – histidina, D – histidina o se ha suprimido;

<220>

<221>

MISC_CARACTERÍSTICA

<222>

(2) . . (2)

<223>

Xaa en la posición 2 es Gly, Ala, Val, Leu, Ile, Ser o Thr;

<220>

<221>

MISC_CARACTERÍSTICA

<222>

(16) . . (16)

<223>

Xaa en la posición 16 es Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys:

<220>

<221>

MISC_CARACTERÍSTICA

<222>

(17) . . (17)

<223>

Xaa en la posición 17 es His, Asp, Lys, Glu o Gln;

<220>

<221>

MISC_CARACTERÍSTICA

<222> (18) . .(18)
 <223> Xaa en la posición 18 es Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;

<220>
 <221> MISC_CHARACTERÍSTICA
 <222> (24) . .(24)
 <223> Xaa en la posición 24 es Ala, Glu, Asp, Ser o His;

<220>
 <221> MISC_CHARACTERÍSTICA
 <222> (31) . .(31)
 <223> Xaa en la posición 31 es Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly – Pro, o se ha suprimido.

<220>
 <221> MISC_CHARACTERÍSTICA
 <222> (32) . .(32)
 <223> Xaa en la posición 32 es Pro o se ha suprimido;

<400> 6
 Xaa Xaa Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Xaa
 1 5 10 15

Xaa Xaa Ala Lys Glu Phe Ile Xaa Trp Leu Val Lys Gly Arg Xaa Xaa
 20 25 30

<210> 7
 <211> 31
 <212> PRT
 <213> Secuencia artificial
 <220>
 <223> construcción sintética
 <220>
 <221> MISC_CHARACTERÍSTICA

<222> (1) . . (1)
<223> Xaa en la posición 1 es L – histidina, D – histidina o se ha suprimido;

<220>
<221> MISC_CHARACTERÍSTICA
<222> (2) . . (2)
<223> Xaa en la posición 2 es Ala, Gly, Val, Thr, Ile y alfa metil – Ala;

<220>
<221> MISC_CHARACTERÍSTICA
<222> (15) . . (15)
<223> Xaa en la posición 15 es Glu, Gln, Ala, Thr, Ser y Gly;

<220>
<221> MISC_CHARACTERÍSTICA
<222> (21) . . (21)
<223> Xaa en la posición 21 es Glu, Gln, Ala, Thr, Ser y Gly;

<400> 7
Xaa Xaa Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Xaa Gly
1 5 10 15

Gln Ala Ala Lys Xaa Phe Ile Ala Trp Leu Val Lys Gly Arg Gly
20 25 30

<210> 8
<211> 30
<212> PRT
<213> Secuencia artificial
<220>
<223> construcción sintética
<220>
<221> MISC_CHARACTERÍSTICA
<222> (19) . . (19)

<223> Xaa en la posición 19 es Lys o Arg;
<220>
<221> MOD_RES
<222> (27) . .(27)
<223> ACETILACIÓN;
<220>
<221> MISC_CARACTERÍSTICA
<222> (30) . .(30)
<223> Xaa en la posición 30 es Gly;
<220>
<221> MOD_RES
<222> (30) . .(30)
<223> AMIDACIÓN
<400> 8
Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Gly Gly Glu
1 5 10 15

Ala Ala Xaa Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Xaa
20 25 30

<210> 9
<211> 39
<212> PRT
<213> Secuencia artificial

<220>
<223> construcción sintética

<400> 9
His Ser Asp Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Gln
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser
35

<210> 10
<211> 39
<212> PRT
<213> Secuencia artificial
<220>
<223> construcción sintética
<400> 10

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser
35

<210> 11
<211> 39
<212> PRT
<213> Secuencia artificial
<220>
<223> construcción sintética
<220>

<221> MISC_CHARACTERÍSTICA
<222> (1) . (1)
<223> Xaa en la posición 1 es L – histidina, D – histidina o se ha suprimido;
<220>
<221> MISC_CHARACTERÍSTICA
<222> (2) . (2)
<223> Xaa en la posición 2 es Gly, Ala o Val;
<220>

- <221>

MISC_CHARACTERÍSTICA
- <222>

(10) . .(10)
- <223>

Xaa en la posición 10 es Leu o Val;
- <220>
- <221>

MISC_CHARACTERÍSTICA
- <222>

(12) . .(12)
- <223>

Xaa en la posición 12 es Lys o Ser;
- <220>
- <221>

MISC_CHARACTERÍSTICA
- <222>

(13) . .(13)
- <223>

Xaa en la posición 13 es Gln o Tyr;
- <220>
- <221>

MISC_CHARACTERÍSTICA
- <222>

(14) . .(14)
- <223>

Xaa en la posición 14 es Met o Leu;
- <220>
- <221>

MISC_CHARACTERÍSTICA
- <222>

(16) . .(16)
- <223>

Xaa en la posición 16 es Glu o Gln;
- <220>
- <221>

MISC_CHARACTERÍSTICA
- <222>

(17) . .(17)
- <223>

Xaa en la posición 17 es Glu o Gln;
- <220>
- <221>

MISC_CHARACTERÍSTICA
- <222>

(19) . .(19)
- <223>

Xaa en la posición 19 es Val o Ala;

- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (20) . .(20)
- <223> Xaa en la posición 20 es Arg o Lys;

- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (21) . .(21)
- <223> Xaa en la posición 21 es Leu o Glu;

- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (24) . .(24)
- <223> Xaa en la posición 24 es Glu o Ala;

- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (27) . .(27)
- <223> Xaa en la posición 27 es Val o Lys;

- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (28) . .(28)
- <223> Xaa en la posición 28 es Asn o Lys;

- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (30) . .(30)
- <223> Xaa en la posición 30 es Gly o Arg; y

- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (31) . .(31)
- <223> Xaa en la posición 31 es Gly o Pro;

- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (32) . .(32)
- <223> Xaa en la posición 32 es Ser o está ausente;

- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (33) . .(33)
- <223> Xaa en la posición 33 es Ser o está ausente;

- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (34) . .(34)
- <223> Xaa en la posición 34 es Gly o está ausente;

- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (35) . .(35)
- <223> Xaa en la posición 35 es Ala o está ausente;

- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (36) . .(36)
- <223> Xaa en la posición 36 es Pro o está ausente;

- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (37) . .(37)
- <223> Xaa en la posición 37 es Pro o está ausente;

- <220>
- <221> MISC_CHARACTERÍSTICA
- <222> (38) . .(38)

<223> Xaa en la posición 38 es Pro o está ausente;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (39) . .(39)

<223> Xaa en la posición 39 es Pro o está ausente;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (39) . .(39)

<223> Xaa en la posición 39 es Ser o está ausente;

<400> 11

Xaa	Xaa	Glu	Glu	Thr	Phe	Thr	Ser	Asp	Xaa	Ser	Xaa	Xaa	Xaa	Glu	Xaa
1				5				10						15	
Xaa	Ala	Xaa	Xaa	Xaa	Phe	Ile	Xaa	Trp	Leu	Xaa	Xaa	Gly	Xaa	Xaa	Xaa
				20				25				30			
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa								
				35											

<210> 12

<211> 31

<212> PRT

<213> Secuencia artificial

<220>

<223> construcción sintética

<220>

<221> MISC_CHARACTERÍSTICA

<222> (1) . .(1)

<223> Xaa en la posición 1 es L – histidina, D – histidina o se ha suprimido;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (2) . .(2)

<223> Xaa en la posición 2 es Ala, Gly, Val, Leu, Ile, Ser o Thr;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (6) . (6)

<223> Xaa en la posición 6 es Phe, Trp o Tyr;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (10) . (10)

<223> Xaa en la posición 10 es Val, Trp, Ile, Leu, Phe o Tyr;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (12) . (12)

<223> Xaa en la posición 12 es Ser, Trp, Tyr, Phe, Lys, Ile, Leu, Val;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (13) . (13)

<223> Xaa en la posición 13 es Tyr, Trp o Phe;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (14) . (14)

<223> Xaa en la posición 14 es Leu, Phe, Tyr o Trp;

<220>

<221> MISC_CHARACTERÍSTICA

<222> (16) . (16)

<223> Xaa en la posición 16 es Gly, Glu, Asp o Lys;

<220>

<221> MISC_CHARACTERÍSTICA

<223> Xaa en la posición 19 es Ala, Val, Ile o Leu:

<222> (21) . .(21)

<223> Xaa en la posición 21 es Glu, Ile o Ala;

<220>

<222> (24) ..(24)

<223> Xaa en la posición 24 es Ala o Glu;

<220>

<222> (27) . .(27)

<223> Xaa en la posición 27 es Val o Ile; y

<220>

<221> MOD RES

<222> (30) ..(30)

<223> AMIDACIÓN

<220>

<222> (31) ..(31)

<223> Xaa en la posición 31 es Gly, His o está ausente;

<400> 12

Xaa Xaa Glu Gly Thr Xaa Thr Ser Asp Xaa Ser Xaa Xaa Xaa Glu Xaa

1 5 10 15

Gln Ala Xaa Lys Xaa Phe Ile Xaa Trp Leu Xaa Lys Gly Arg Xaa

20 25 30

◁210▷ 13

<211> 616

<212> PRT

<213> secuencia artificial

<220>

<223> construcción sintética

<400> 13

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Asp
 20 25 30

Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu
 35 40 45

Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln
 50 55 60

Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe
 65 70 75 80

Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser
 85 90 95

Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg
 100 105 110

Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu
 115 120 125

Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro
 130 135 140

Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp
 145 150 155 160

Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg
 165 170 175

His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr
 180 185 190

Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys
 195 200 205

Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser Ser
 210 215 220

Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg
 225 230 235 240
 Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys
 245 250 255
 Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val
 260 265 270
 His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg
 275 280 285
 Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser
 290 295 300
 Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys
 305 310 315 320
 Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu
 325 330 335
 Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu
 340 345 350
 Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg
 355 360 365
 His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr
 370 375 380
 Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys
 385 390 395 400
 Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln
 405 410 415
 Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr
 420 425 430
 Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln
 435 440 445
 Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val
 450 455 460
 Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala
 465 470 475 480
 Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu
 485 490 495

Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu
500 505 510

Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr
515 520 525

Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile
530 535 540

Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu
545 550 555 560

Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys
565 570 575

Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala
580 585 590

Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala
595 600 605

Ala Ser Gln Ala Ala Leu Gly Leu
610 615

<210> 14

<211> 631

<212> PRT

<213> secuencia artificial

<220>

<223> construcción sintética

<400> 14

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Gly
20 25 30

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ala
35 40 45

His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn
50 55 60

Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Cys
65 70 75 80

Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala
 85 90 95
 Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu
 100 105 110
 His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu
 115 120 125
 Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg
 130 135 140
 Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg
 145 150 155 160
 Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp Asn
 165 170 175
 Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His
 180 185 190
 Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys
 195 200 205
 Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu
 210 215 220
 Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser Ser Ala
 225 230 235 240
 Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala
 245 250 255
 Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala
 260 265 270
 Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val His
 275 280 285
 Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala
 290 295 300
 Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys
 305 310 315 320
 Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile
 325 330 335
 Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala
 340 345 350

Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala
 355 360 365
 Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His
 370 375 380
 Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu
 385 390 395 400
 Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr
 405 410 415
 Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln Asn
 420 425 430
 Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys
 435 440 445
 Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val
 450 455 460
 Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly
 465 470 475 480
 Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu
 485 490 495
 Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys
 500 505 510
 Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val
 515 520 525
 Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val
 530 535 540
 Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys
 545 550 555 560
 Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val
 565 570 575
 Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala
 580 585 590
 Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp
 595 600 605
 Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala
 610 615 620

Ser Gln Ala Ala Leu Gly Leu
625 630

<210> 15

<211> 640

<212> PRT

<213> secuencia artificial

<220>

<223> construcción sintética

<400> 15

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly
35 40 45

Gly Ser Gly Gly Gly Gly Ser Asp Ala His Lys Ser Glu Val Ala His
50 55 60

Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile
65 70 75 80

Ala Phe Ala Gln Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val Lys
85 90 95

Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu
100 105 110

Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys
115 120 125

Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp
130 135 140 1

Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His
145 150 155 160

Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp
165 170 175

Val Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys
180 185 190

Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu
 195 200 205
 Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys
 210 215 220
 Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu
 225 230 235 240
 Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala
 245 250 255
 Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala
 260 265 270
 Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys
 275 280 285
 Leu Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp
 290 295 300
 Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys
 305 310 315 320
 Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys
 325 330 335
 Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu
 340 345 350
 Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys
 355 360 365
 Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met
 370 375 380
 Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu
 385 390 395 400
 Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys
 405 410 415
 Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe
 420 425 430
 Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu
 435 440 445
 Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val
 450 455 460

Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu
465 470 475 480

Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro
485 490 495

Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu
500 505 510

Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val
515 520 525

Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser
530 535 540

Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu
545 550 555 560

Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg
565 570 575

Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro
580 585 590

Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala
595 600 605

Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala
610 615 620

Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
625 630 635 640

<210> 16

<211> 624

<212> PRT

<213> secuencia artificial

<220>

<223> construcción sintética

<400> 16

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

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

Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys
 305 310 315 320
 Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu
 325 330 335
 Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys
 340 345 350
 Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met
 355 360 365
 Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu
 370 375 380
 Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys
 385 390 395 400
 Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe
 405 410 415
 Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu
 420 425 430
 Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu Val
 435 440 445
 Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val Glu
 450 455 460
 Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His Pro
 465 470 475 480
 Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu
 485 490 495
 Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val
 500 505 510
 Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe Ser
 515 520 525
 Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala Glu
 530 535 540
 Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu Arg
 545 550 555 560
 Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro
 565 570 575

400
397

Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala
580 585 590

Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe Ala
595 600 605

Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
610 615 620

<210> 17

<211> 640

<212> PRT

<213> secuencia artificial

<220>

<223> construcción sintética

<400> 17

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly
35 40 45

Gly Ser Gly Gly Gly Gly Ser Asp Ala His Lys Ser Glu Val Ala His
50 55 60

Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu Ile
65 70 75 80

Ala Phe Ala Gln Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val Lys
85 90 95

Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu
100 105 110

Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys
115 120 125

Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala Asp
130 135 140

Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln His
145 150 155 160

Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val Asp
 165 170 175
 Val Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys
 180 185 190
 Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu
 195 200 205
 Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys Cys
 210 215 220
 Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu Leu
 225 230 235 240
 Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys Ala
 245 250 255
 Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala
 260 265 270
 Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys
 275 280 285
 Leu Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly Asp
 290 295 300
 Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile Cys
 305 310 315 320
 Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu Lys
 325 330 335
 Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu
 340 345 350
 Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys
 355 360 365
 Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly Met
 370 375 380
 Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val Leu
 385 390 395 400
 Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys Cys
 405 410 415
 Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe
 420 425 430

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

<210> 18

<211> 264

<212> PRT

<213> secuencia artificial

<220>

<223> construcción sintética

497
400

<400> 18

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ala
20 25 30

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
35 40 45

Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
50 55 60

Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
65 70 75 80

Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
85 90 95

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
100 105 110

Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln
115 120 125

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
130 135 140

Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
145 150 155 160

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
165 170 175

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
180 185 190

Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
195 200 205

Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
210 215 220

Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
225 230 235 240

Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
245 250 255

Ser Leu Ser Leu Ser Pro Gly Lys
260

<210> 19

<211> 272

<212> PRT

<213> secuencia artificial

<220>

<223> construcción sintética

<400> 19

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr
35 40 45

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
50 55 60

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
65 70 75 80

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
85 90 95

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
100 105 110

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
115 120 125

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
130 135 140

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
145 150 155 160

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
165 170 175

Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
180 185 190

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
195 200 205

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
210 215 220

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
225 230 235 240

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
245 250 255

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
260 265 270

<210> 20

<211> 264

<212> PRT

<213> secuencia artificial

<220>

<223> construcción sintética

<400> 20

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ala
20 25 30

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
35 40 45

Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
50 55 60

Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
65 70 75 80

Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
85 90 95

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
100 105 110

Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln
115 120 125

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
130 135 140

Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
145 150 155 160

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
165 170 175

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
180 185 190

Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
195 200 205

Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
210 215 220

Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
225 230 235 240

Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
245 250 255

Ser Leu Ser Leu Ser Pro Gly Lys
260

<210> 21

<211> 272

<212> PRT

<213> secuencia artificial

<220>

<223> construcción sintética

<400> 21

His Val Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr
35 40 45

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
50 55 60

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
65 70 75 80

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
85 90 95

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
100 105 110

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
115 120 125

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
130 135 140

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
145 150 155 160

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
165 170 175

Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
180 185 190

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
195 200 205

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
210 215 220

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
225 230 235 240

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
245 250 255

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
260 265 270

<210> 22

<211> 272

<212> PRT

<213> secuencia artificial

<220>

<223> construcción sintética

<400> 22

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

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

<211> 287

<212> PRT

<213> secuencia artificial

<220>

<223> construcción sintética

<400> 23

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
35 40 45

Ser Gly Gly Gly Gly Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr His
50 55 60

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
65 70 75 80

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
85 90 95

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu
100 105 110

Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
115 120 125

Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser
130 135 140

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
145 150 155 160

Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
165 170 175

Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
180 185 190

Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
195 200 205

Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
210 215 220

Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
 225 230 235 240

Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
 245 250 255

Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
 260 265 270

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
 275 280 285

<210> 24
 <211> 284
 <212> PRT
 <213> secuencia artificial

<220>
 <223> construcción sintética ..
 <400> 24

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
 1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
 20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
 35 40 45

Ser Gly Gly Gly Gly Ser Ala Glu Ser Lys Tyr Gly Pro Pro Cys Pro
 50 55 60

Ser Cys Pro Ala Pro Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe
 65 70 75 80

Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val
 85 90 95

Thr Cys Val Val Val Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe
 100 105 110

Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro
 115 120 125

Arg Glu Glu Gln Phe Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr
 130 135 140

Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val
 145 150 155 160
 Ser Asn Lys Gly Leu Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala
 165 170 175
 Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln
 180 185 190
 Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly
 195 200 205
 Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro
 210 215 220
 Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser
 225 230 235 240
 Phe Phe Leu Tyr Ser Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu
 245 250 255
 Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His
 260 265 270
 Tyr Thr Gln Lys Ser Leu Ser Leu Ser Leu Gly Lys
 275 280

<210> 25

<211> 302

<212> PRT

<213> secuencia artificial

<220>

<223> construcción sintética

<400> 25

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
 1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
 20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
 35 40 45

Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
 50 55 60

Gly Gly Gly Gly Ser Ala Glu Pro. Lys Ser Cys Asp Lys Thr His Thr
 65 70 75 80
 Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe
 85 90 95
 Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro
 100 105 110
 Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val
 115 120 125
 Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr
 130 135 140
 Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val
 145 150 155 160
 Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys
 165 170 175
 Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser
 180 185 190
 Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro
 195 200 205
 Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val
 210 215 220
 Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly
 225 230 235 240
 Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp
 245 250 255
 Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp
 260 265 270
 Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His
 275 280 285
 Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
 290 295 300

<210> 26

<211> 294

<212> PRT

<213> secuencia artificial

<220>

<223> construcción sintética

<400> 26

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Gly
20 25 30

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
35 40 45

Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Glu Pro
50 55 60

Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu
65 70 75 80

Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp
85 90 95

Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp
100 105 110

Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly
115 120 125

Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn
130 135 140

Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp
145 150 155 160

Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro
165 170 175

Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu
180 185 190

Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn
195 200 205

Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile
210 215 220

Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr
225 230 235 240

Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys
245 250 255

Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys
260 265 270

Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu
275 280 285

Ser Leu Ser Pro Gly Lys
290

<210> 27

<211> 280

<212> PRT

<213> secuencia artificial

<220>

<223> construcción sintética

<400> 27

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Ser Ser Gly Ala Pro Pro Pro Ser Ala
35 40 45

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
50 55 60

Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
65 70 75 80

Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
85 90 95

Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
100 105 110

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
115 120 125

Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln
130 135 140

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
145 150 155 160

Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
165 170 175

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
180 185 190

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
195 200 205

Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
210 215 220

Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
225 230 235 240

Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
245 250 255

Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
260 265 270

Ser Leu Ser Leu Ser Pro Gly Lys
275 280

<210> 28

<211> 287

<212> PRT

<213> secuencia artificial

<220>

<223> construcción sintética

<400> 28

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Glu
1 5 10 15

Gln Ala Val Lys Glu Phe Ile Ala Trp Leu Ile Lys Gly Arg Gly Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Ser Gly Gly Gly Gly
35 40 45

Ser Gly Gly Gly Gly Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr His
50 55 60

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
 65 70 75 80
 Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
 85 90 95
 Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu
 100 105 110
 Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
 115 120 125
 Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser
 130 135 140
 Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
 145 150 155 160
 Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
 165 170 175
 Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
 180 185 190
 Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
 195 200 205
 Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
 210 215 220
 Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
 225 230 235 240
 Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
 245 250 255
 Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
 260 265 270
 His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
 275 280 285
 <210> 29
 <211> 272
 <212> PKT
 <213> secuencia artificial

X-

<220>

<223> construcción sintética

<400> 29

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr
35 40 45

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
50 55 60

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
65 70 75 80

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
85 90 95

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
100 105 110

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
115 120 125

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
130 135 140

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
145 150 155 160

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
165 170 175

Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
180 185 190

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
195 200 205

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
210 215 220

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
225 230 235 240

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
245 250 255

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys .
 260 265 270

- <210> 30
- <211> 272
- <212> PRT
- <213> secuencia artificial

- <220>
- <223> construcción sintética

<400> 30

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
 1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
 20 25 30

Ser Gly Ala Ser Ser Gly Ala Ala Glu Pro Lys Ser Cys Asp Lys Thr
 35 40 45

His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser
 50 55 60

Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg
 65 70 75 80

Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro
 85 90 95

Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala
 100 105 110

Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val
 115 120 125

Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr
 130 135 140

Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr
 145 150 155 160

Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu
 165 170 175

Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys
 180 185 190

Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser
195 200 205

Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp
210 215 220

Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser
225 230 235 240

Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala
245 250 255

Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
260 265 270

<210> 31
<211> 287
<212> PKT
<213> secuencia artificial

<220>
<223> construcción sintética
<400> 31

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

Ser Gly Ala Pro Pro Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
35 40 45

Ser Gly Gly Gly Gly Ser Ala Glu Pro Lys Ser Cys Asp Lys Thr His
50 55 60

Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val
65 70 75 80

Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
85 90 95

Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu
100 105 110

Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
115 120 125

Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser
130 135 140

Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
145 150 155 160

Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile
165 170 175

Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
180 185 190

Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
195 200 205

Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
210 215 220

Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser
225 230 235 240

Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
245 250 255

Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
260 265 270

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
275 280 285

<210> 32
<211> 232
<212> PRT
<213> Homo sapiens

<400> 32.

Ala Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro
1 5 10 15

Ala Pro Glu Lys Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro
20 25 30

Lys Asp Thr Lys Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
35 40 45

Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
50 55 60

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
65 70 75 80

Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln
85 90 95

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
100 105 110

Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
115 120 125

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
130 135 140

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
145 150 155 160

Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
165 170 175

Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
180 185 190

Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
195 200 205

Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
210 215 220

Ser Leu Ser Leu Ser Pro Gly Lys
225 230

<210> 33

<211> 703

<212> DNA

<213> Homo sapiens

<400> 33
gagcccaaat cttgtgacaa aactcacaca tgcccacagt gccagcacc tgaactcctg 60
gggggaccgt cagtcttctt cttccccca aaaccaagg acacctcat gatctcccg 120
acccctgagg tcacatgctt ggtggtggac gtgagccacg aagacctga ggtcaagttc 180
aactggtacg tggacggcgt ggaggtgcat aatgccaaga caagccgcg ggaggagcag 240
tacaacagca cgtaccgtgt ggtcagcgtc ctcaccgtcc tgcaccagga ctggctgaat 300
ggcaaggagt acaagtgcaa ggtctccaac aaagccctcc cagcccccat cpagaaaacc 360

```

-----
atctccaaag ccaaggggca gccccgagaa ccacaggtgt acaccctgcc cccatcccgg 420
gaggagatga ccaagaacca ggtcagcctg acctgcctgg tcaaaggctt ctatcccagc 480
gacatcgccg tggagtggga gagcaatggg cagccggaga acaactacaa gaccacgcct 540
cccgtgctgg actccgacgg ctctctcttc ctctatagca agctcaccgt ggacaagagc 600
aggtggcagc aggggaacct cttctcatgc tccgtgatgc atgaggctct gcacaaccac 660
tacacgcaga agagcctctc cctgtctccg ggtaaatgat agt 703

```

```

<210> 34
<211> 585
<212> PRT
<213> Homo sapiens

```

```

<400> 34
Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
1      5      10     15
Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
20     25     30
Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
35     40     45
Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
50     55     60
Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
65     70     75     80
Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
85     90     95
Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
100    105   110
Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
115    120   125
Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
130    135   140
Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
145    150   155   160
Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
165    170   175

```

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser
 180 185 190
 Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
 195 200 205
 Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro
 210 215 220
 Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys
 225 230 235 240
 Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp
 245 250 255
 Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser
 260 265 270
 Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His
 275 280 285
 Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser
 290 295 300
 Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala
 305 310 315 320
 Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg
 325 330 335
 Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr
 340 345 350
 Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu
 355 360 365
 Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro
 370 375 380
 Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Asn Leu Gly Glu
 385 390 395 400
 Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro
 405 410 415
 Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys
 420 425 430
 Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys
 435 440 445

Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His
 450 455 460

Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser
 465 470 475 480

Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr
 485 490 495

Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp
 500 505 510

Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
 515 520 525

Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
 530 535 540

Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
 545 550 555 560

Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
 565 570 575

Ala Ala Ser Gln Ala Ala Leu Gly Leu
 580 585

<210> 35
 <211> 1762
 <212> DNA
 <213> Homo sapiens

<400> 35
 gatgcgcaca agagtgaggt tgctcatcgg tttaaagatt tgggagaaga aaatttcaaa 60
 gccttggtgt tgattgcctt tgctcagtat cttcagcagt gtccatttga agatcatgta 120
 aaattagtga atgaagtaac tgaatttgca aaaacatgtg ttgctgatga gtcagctgaa 180
 aattgtgaca aatcacttca tacccttttt ggagacaaat tatgcacagt tgcaactctt 240
 cgtgaaacct atggtgaaat ggctgactgc tgtgcaaaac aagaacctga gagaaatgaa 300
 tgcttcttgc aacacaaaga tgacaacca aacctcccc gattggtgag accagaggtt 360
 gatgtgatgt gcactgcttt tcatgacant gaagagacat ttttgaaaa atacttatat 420
 gaaattgcca gaagacatcc ttacttttat gccccggaac tccttttctt tgctaaaagg 480
 tataaagctg cttttacaga atgttgccaa gctgctgata aagctgcctg cctgttgcca 540
 aagctcgatg aacttcggga tgaagggaag gcttcgtctg ccaaacagag actcaagtgt 600

gccagtctcc aaaaatttgg agaaagagct ttcaaagcat gggcagtagc tcgcctgagc 660
 cagagatttc ccaaagctga gtttgcagaa gtttccaagt tagtgacaga tcttaccaaa 720
 gtccacacgg aatgctgcca tggagatctg cttgaatgtg ctgatgacag ggcggacctt 780
 gccaaagtata tctgtgaaa tcaagattcg atctccagta aactgaagga atgctgtgaa 840
 aaacctctgt tggaaaaatc cactgcatg gccgaagtgg aaaatgatga gatgcctgct 900
 gacttgccctt cattagctgc tgattttgtt gaaagtaagg atgtttgcaa aaactatgct 960
 gaggcaagg atgtcttcct gggcatgttt ttgtatgaat atgcaagaag gcctcctgat 1020
 tactctgtcg tgctgctgct gagacttgcc aagacatatg aaaccactct agagaagtgc 1080
 tgtgccgctg cagatcctca tgaatgctat gccaaagtgt tcgatgaatt taaacctctt 1140
 gtggaagagc ctcaagaattt aatcaacaa aattgtgagc tttttgagca gcttggagag 1200
 tacaaattcc agaatgcgct attagttcgt tacaccaaga aagtacccca agtgtcaact 1260
 ccaactcttg tagaggtctc aagaaccta ggaaaagtgg gcagcaaatg ttgtaaacat 1320
 cctgaagcaa aaagaatgcc ctgtgcagaa pactatctat ccgtggctct gaaccagtta 1380
 tgtgtgttgc atgagaaaac gccagtaagt pacagagtca ccaaatgctg cacagaatcc 1440
 ttggtgaaca ggcgaccatg cttttcagct ctggaagtgc atgaaacata cgttcccaaa 1500
 gagttaaatg ctgaaacatt caccttccat gcagatatat gcacactttc tgagaaggag 1560
 agacaaatca agaaacaaac tgcacttggt gagctcgtga aacacaagcc caaggcaaca 1620
 aaagagcaac tgaagctgt tatggatgat ttgcagctt ttgtagagaa gtgctgcaag 1680
 gctgacgata aggagacctg ctttgccgag gagggtaaaa aacttgttgc tgcaagtcaa 1740
 gctgccttag gcttataatg ac 1762

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference X-13991	FOR FURTHER ACTION see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
International application No. PCT/US 01/ 43165	International filing date (day/month/year) 29/11/2001	(Earliest) Priority Date (day/month/year) 07/12/2000
Applicant ELI LILLY AND COMPANY		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 3 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. Basis of the report

- a. With regard to the language, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.

☐ the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

- b. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of the sequence listing:

☒ contained in the international application in written form.

☒ filed together with the international application in computer readable form.

☐ furnished subsequently to this Authority in written form.

☐ furnished subsequently to this Authority in computer readable form.

☐ the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

☐ the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

2. ☐ Certain claims were found unsearchable (See Box I).

3. ☐ Unity of invention is lacking (see Box II).

4. With regard to the title,

☒ the text is approved as submitted by the applicant.

☐ the text has been established by this Authority to read as follows:

5. With regard to the abstract,

☒ the text is approved as submitted by the applicant.

☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the drawings to be published with the abstract is Figure No.

☐ as suggested by the applicant.

☐ because the applicant failed to suggest a figure.

☐ because this figure better characterizes the invention.

☒ None of the figures.

INTERNATI NAL SEARCH REPORT

International Application No.

PCT/US 01/43165

424

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C07K14/605 C07K14/765 C12N15/62 A61K38/38 A61K38/26
C07K19/00 A61P3/10

426

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)

IPC 7 C07K C12N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, MEDLINE, CHEM ABS Data, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
✓ A	WO 00 69911 A (ARCHEVEQUE BENOIT L ;CONJUCHEM INC (CA); LEBLANC ANOUK (CA); HOLME) 23 November 2000 (2000-11-23) claims the whole document	1-50
✓ A	KNUDSEN LOTTE B ET AL: "Potent derivatives of glucagon-like peptide-1 with pharmacokinetic properties suitable for once daily administration." JOURNAL OF MEDICINAL CHEMISTRY, vol. 43, no. 9, 4 May 2000 (2000-05-04), pages 1664-1669, XP002222050 ISSN: 0022-2623 the whole document	1-50

☐ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex

* Special categories of cited documents

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"Z" document member of the same patent family

Date of the actual completion of the international search

22 November 2002

Date of mailing of the international search report

21/01/2003

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentkan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-3040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Döpfer, K-P

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/43165

425 429

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0069911	A	23-11-2000	AT 226593 T 15-11-2002
			AU 4774800 A 05-12-2000
			AU 4855500 A 05-12-2000
			AU 5027100 A 05-12-2000
			AU 5139300 A 05-12-2000
			BR 0010750 A 26-02-2002
			BR 0010757 A 19-02-2002
			CN 1350548 T 22-05-2002
			CN 1351611 T 29-05-2002
			EP 1171582 A2 16-01-2002
			EP 1180121 A1 20-02-2002
			EP 1179012 A1 13-02-2002
			EP 1105409 A2 13-06-2001
			WO 0070665 A2 23-11-2000
			NO 20015584 A 03-01-2002
			WO 0069911 A1 23-11-2000
			WO 0069900 A2 23-11-2000
			WO 0069902 A1 23-11-2000
			US 6329336 B1 11-12-2001
			US 2002049153 A1 25-04-2002

REQUISITOS NECESARIOS PARA PRESENTACION Y AGILIZACION DEL TRAMITE

	USUARIO	RADICADOR
PATENTE DE INVENCION	()	()
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	()	()

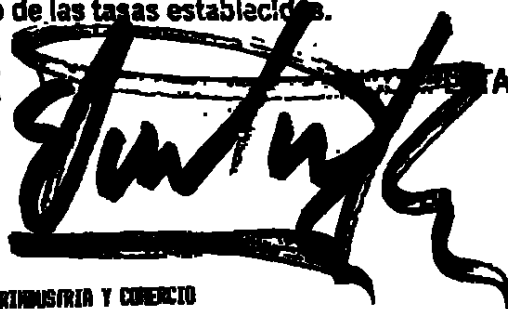
COMPLETA () INCOMPLETA ()

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:



SUPERINDUSTRIA Y COMERCIO

Fecha Radicación: 03042706 00000000

Folios: 429

Fecha (UMB): 2003-06-09 16:53:52

Trámite: 011 PCT-PATENT 379 PAGE NUCI 411 PRESENTAC

Dependencia: 2020 DIVISION DE NUEVAS CREACIONES



2020
Bogotá, D.C.,

Folio 431

Doctor:
Carlos R. Olarte
Apoderado
El Lilly and Company

Oficio N° 07509

ASUNTO:	Radicación N°	03-048708
	Solicitud internacional	PCT/US2001/43165
	Fecha inicio fase nacional	07/07/2003
	Trámite	011
	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).
- ☐ Los dibujos adjuntos de las figuras 3 y 8 no son legibles o copiables para trámites posteriores de publicación y título.
- ☐ La solicitud no contiene la página número 50 de la traducción de la descripción de la invención.

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.

12 SET. 2003

El anterior requerimiento fue notificado por fijación en lista de fecha. _____