(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 328 510**

(51) Int. Cl.:
C07K 14/605 (2006.01)
C07K 14/765 (2006.01)
C12N 15/62 (2006.01)
A61K 38/38 (2006.01)
A61K 38/26 (2006.01)
C07K 19/00 (2006.01)
A61P 3/10 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Número de solicitud europea: **06118975 .9**
(96) Fecha de presentación : **29.11.2001**
(97) Número de publicación de la solicitud: **1724284**
(97) Fecha de publicación de la solicitud: **22.11.2006**

(54) Título: **Proteínas de fusión GLP-1.**

(30) Prioridad: **07.12.2000 US 251954 P**

(45) Fecha de publicación de la mención BOPI:
13.11.2009

(45) Fecha de la publicación del folleto de la patente:
13.11.2009

(73) Titular/es: **Eli Lilly & Company**
Lilly Corporate Center
Indianapolis, Indiana 46285, US

(72) Inventor/es: **Glaesner, Wolfgang;**
Micanovic, Radmila y
Tschang, Sheng-Hung Rainbow

(74) Agente: **Carpintero López, Mario**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

Venta de fascículos: Oficina Española de Patentes y Marcas. Pº de la Castellana, 75 – 28071 Madrid

DESCRIPCIÓN

Proteínas de fusión GLP-1.

La presente invención se refiere a péptidos similares al glucagón, incluyendo análogos y derivados de los mismos, fusionados con proteínas que tienen el efecto de aumentar la vida media *in vivo* de los péptidos. Estas proteínas de fusión se pueden usar para tratar la diabetes mellitus no dependiente de la insulina, así como una variedad de otras afecciones.

El péptido similar al glucagón de tipo 1 (GLP-1) es un péptido de 37 aminoácidos que es secretado por las células L del intestino como respuesta a la ingestión de alimentos. Se ha descubierto que estimula la secreción de insulina (acción insulinotrópica), causando así la absorción de glucosa por las células y la disminución de los niveles de glucosa en suero [véase, p. ej., Mojsov, S., (1992) *Int. J. Peptide Protein Research*, 40: 333-343]. Sin embargo, el GLP-1 es poco activo. Una escisión endógena posterior entre la sexta y la séptima posición produce un péptido GLP-1(7-37) OH con una actividad biológica más potente. Se conocen numerosos análogos y derivados de GLP-1, y, en la presente memoria, se denominan "compuestos de GLP-1". Estos análogos de GLP-1 incluyen las exendinas, que son péptidos que se encuentran en el veneno de Heloderma. Las exendinas tienen una homología secuencial con el GLP-1 nativo y se pueden unir al receptor de GLP-1 e iniciar la cascada de transducción de señales responsable de las numerosas actividades que han sido atribuidas al GLP-1(7-37)OH.

Los compuestos de GLP-1 tienen una variedad de actividades fisiológicamente significativas. Por ejemplo, se ha observado que el GLP-1 estimula la liberación de insulina, disminuye la secreción del glucagón, inhibe la evacuación gástrica y aumenta la utilización de la glucosa. [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].

El GLP-1 supone la mayor promesa como tratamiento para la diabetes mellitus no dependiente de la insulina (DMNDI). Hay numerosos fármacos orales en el mercado para tratar la resistencia a la insulina asociada con la DMNDI. Sin embargo, a medida que va progresando la enfermedad, los pacientes deben pasar a tratamientos que estimulan la liberación de insulina y, finalmente, a tratamientos que implican inyecciones de insulina. Sin embargo, los fármacos actuales que estimulan la liberación de insulina también pueden causar hipoglicemia, al igual que la propia administración de insulina. No obstante, la actividad de GLP-1 está controlada por los niveles de glucosa en sangre. Cuando los niveles caen hasta un cierto nivel umbral, GLP-1 deja de ser activo. De este modo, no existe el riesgo de una hipoglicemia asociada con el tratamiento que implica el GLP-1.

Sin embargo, la utilidad de la terapia con péptidos GLP-1 se ha visto limitada por su rápido aclaramiento y sus cortas vidas medias. Por ejemplo, el GLP-1(7-37) tiene una vida media en suero de sólo 3 a 5 minutos. El GLP-1(7-36) amida tiene una acción temporal de aproximadamente 50 minutos cuando se administra subcutáneamente. Incluso los análogos y los derivados que son resistentes a la escisión por proteasas endógenas carecen de vidas medias lo suficientemente largas como para evitar las administraciones repetidas durante un período de 24 horas. El aclaramiento rápido de un agente terapéutico no es deseable en aquellos casos en los que se quiera mantener un nivel elevado en sangre del agente durante un período prolongado de tiempo, pues se crea la necesidad de administraciones repetidas. Además, los compuestos de actuación prolongada son particularmente importantes para los pacientes diabéticos cuyo régimen de tratamiento en el pasado ha implicado tomar únicamente una medicación oral. Estos pacientes tienen a menudo una transición temporal extremadamente difícil a un régimen que implica múltiples inyecciones de medicación.

La presente invención resuelve los problemas asociados con la administración de un compuesto que tiene una vida media en plasma corta. Los compuestos de la presente invención engloban compuestos de GLP-1 fusionados con otra proteína con una vida media en circulación larga, tal como la porción Fc de una inmunoglobulina o una albúmina.

En general, los péptidos terapéuticos pequeños son difíciles de manipular, porque incluso los pequeños cambios de su estructura pueden afectar a la estabilidad y/o a la actividad biológica. Esto ha sido especialmente cierto para los compuestos de GLP-1 que se encuentran actualmente en desarrollo. Por ejemplo, el GLP-1(7-37)OH tiene tendencia a sufrir un cambio de configuración de una estructura fundamentalmente de hélice alfa a una estructura fundamentalmente de lámina beta. Esta forma de lámina beta da como resultado un material agregado que se cree que es inactivo. Por lo tanto, resultó sorprendente poder desarrollar proteínas de fusión GLP-1 biológicamente activas con vidas medias más prolongadas. Esto era especialmente inesperado debido a la dificultad de trabajar con GLP-1 (7-37)OH solo y al gran tamaño de la pareja de fusión en comparación con el pequeño tamaño del péptido GLP-1 unido.

Los compuestos de la presente invención incluyen proteínas de fusión heterólogas que comprenden un primer polipéptido con un terminal N y un terminal C fusionado con un segundo polipéptido con un terminal N y un terminal C, siendo el primer polipéptido un compuesto de GLP-1 y seleccionándose el segundo polipéptido de entre:

a) albúmina humana;

b) análogos de albúmina humana; y

c) fragmentos de albúmina humana,

en los que el terminal C del primer polipéptido está fusionado con el terminal N del segundo polipéptido, y en los que la proteína de fusión es biológicamente activa y tiene una vida media en plasma más larga que el compuesto tomado individualmente.

Los compuestos de la presente invención también incluyen proteínas de fusión heterólogas que comprenden un primer polipéptido con un terminal N y un terminal C fusionados con un segundo polipéptido con un terminal N y un terminal C, siendo el primer polipéptido un compuesto de GLP-1 y seleccionándose entonces el segundo polipéptido del grupo constituido por:

a) albúmina humana;

b) análogos de albúmina humana; y

c) fragmentos de albúmina humana, y en los que el terminal C del primer polipéptido está fusionado con el terminal N del segundo polipéptido mediante un ligador peptídico. Es preferible que el ligador peptídico se seleccione del grupo constituido por:

a) un péptido rico en glicina;

b) un péptido que tenga una secuencia $[Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser]_n$ en la que n es 1, 2, 3, 4. 5 ó 6; y

c) un péptido que tenga una secuencia $[Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser]_3$.

En general, es preferible que el compuesto de GLP-1 que forma parte de la proteína de fusión heteróloga no tenga más de 6 aminoácidos que sean diferentes del correspondiente aminoácido de GLP-1(7-37)OH, GLP-1(7-36)OH o exendina-4. Es incluso más preferible que el compuesto de GLP-1 no tenga más de 5 aminoácidos que difieran del correspondiente aminoácido de GLP-1(7-37)OH, GLP-1(7-36)OH o exendina-4. Lo más preferible es que el compuesto de GLP-1 no tenga más de 4, 3 ó 2 aminoácidos que difieran del correspondiente aminoácido de GLP-1(7-37) OH, GLP-1(7-36)OH o exendina-4. Preferiblemente, el compuesto de GLP-1 que forma parte de la proteína de fusión heteróloga tiene glicina o valina en la posición 8.

La presente invención incluye también polinucleótidos codificantes de la proteína de fusión heteróloga descrita en la presente memoria, vectores que comprenden estos polinucleótidos y células huésped transfectadas o transformadas con los vectores descritos en la presente memoria. También se incluye un procedimiento para producir una proteína de fusión heteróloga que comprende las etapas de transcribir y traducir un polinucleótido descrito en la presente memoria en condiciones en las que la proteína de fusión heteróloga es expresada en cantidades detectables.

La presente invención engloba también un procedimiento para normalizar los niveles de glucosa en sangre en un mamífero en necesidad de ello que comprende la administración de una cantidad terapéuticamente eficaz de una proteína de fusión heteróloga descrita en la presente memoria.

La invención se ilustra en mayor profundidad con referencia a las siguientes figuras:

Figura 1: Secuencia de aminoácidos de Fc de la IgG1 que engloba la los dominios de la región bisagra, CH2 y CH3.

Figura 2: Secuencia de aminoácidos de la albúmina de suero humano.

Figura 3: A. gel para electroforesis en gel de poliacrilamida-SDS e inmunotransferencia del mismo gel que ilustra el peso molecular de las proteínas de fusión IgG1-Fc y GLP-1-Fc (Ruta 1, patrones de PM; Ruta 2, Fc purificado; Ruta 3, medios transfectados de prueba; Ruta 4, Val^8-GLP-1-Fc; Ruta 5, Exendina-4-Fc) B. gel para electroforesis en gel de poliacrilamida-SDS e inmunotransferencia del mismo gel que ilustra el peso molecular de ASH y las proteínas de fusión GLP-1-ASH (Ruta 1, patrones de PM; Ruta 2, ASH purificada; Ruta 3, medios transfectados de prueba; Ruta 4, Val^8-GLP-1-ASH; Ruta 5, Val^8-GLP-1-$[Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser]_3$-ASH; Ruta 6, Exendina-4-ASH; Ruta 7, Exendina-4-$[Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser]_3$-ASH).

Figura 4: gel de electroforesis en gel de poliacrilamida-SDS de Fc purificado, albúmina y proteínas de fusión GLP-1 (Ruta 1, patrones de PM; Ruta 2, Fc purificado; Ruta 3, Val^8-GLP-1-Fc; Ruta 4, Exendina-4-Fc; Ruta 5, patrón de PM; Ruta 6, Val^8-GLP-1-ASH; Ruta 7, Exendina-4-ASH; Ruta 8, Exendina-4-$[Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser]_3$-ASH).

Fig. 5: Vector de clonación de la expresión que contiene las regiones Fc ilustradas en la figura 1.

Fig. 6: Vector de clonación de la expresión que contiene la secuencia de albúmina ilustrada en la figura 2.

Fig. 7: Vector de clonación de la expresión que contiene ADN codificante de un ligador de 15 aminoácidos fusionado en el marco y en 5' de la secuencia de albúmina ilustrada en la figura 2.

Fig. 8: Actividad dosis-respuesta *in vitro* de proteínas de fusión GLP-1.

Fig. 9: Farmacocinéticas de proteínas de fusión de ASH y Fc con GLP-1.

Figura 13: Secuencia de ADN codificante de una proteína de albúmina humana.

Las proteínas de fusión heterólogas de la presente invención comprenden un compuesto de GLP-1 fusionado con una albúmina humana, un análogo de albúmina humana o un fragmento de albúmina humana.

El terminal C del compuesto de GLP-1 puede estar fusionado directamente, o mediante un ligador peptídico, con el terminal N de una albúmina. Estas proteínas de fusión heterólogas son biológicamente activas y tienen una mayor vida media en comparación con GLP-1 nativo.

Es preferible que los compuestos de GLP-1 que forman parte de la proteína de fusión heteróloga engloben polipéptidos que tengan de aproximadamente veinticinco a aproximadamente treinta y nueve aminoácidos naturales o no naturales que tengan una homología suficiente con el GLP-1(7-37)OH nativo tal que presenten una actividad insulinotrópica al unirse con el receptor de GLP-1 sobre células β del páncreas. Un compuesto de GLP-1 comprende comúnmente 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. GLP-1(7-37)OH tiene una secuencia de aminoácidos de SEC ID N.º 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
```

(SEC ID N.º 1)

Según lo acostumbrado en la técnica, al terminal amino de GLP-1(7-37)OH se le ha asignado el residuo número 7 y al terminal carboxilo, el número 37. El resto de aminoácidos del polipéptido están numerados consecutivamente, según lo mostrado en la SEC ID N.º 1. Por ejemplo, la posición 12 es fenilalanina y la posición 22 es glicina.

Los compuestos de GLP-1 también engloban "fragmentos de GLP-1". Un fragmento de GLP-1 es un polipéptido obtenido tras el truncamiento de uno o más aminoácidos desde el terminal N y/o el terminal C de GLP-1(7-37)OH, o un análogo o derivado del mismo. La nomenclatura usada para describir a GLP-1(7-37)OH también es aplicable a los fragmentos de GLP-1. Por ejemplo, GLP-1(9-36)OH denota un fragmento de GLP-1 obtenido mediante el truncamiento de dos aminoácidos del terminal N y de un aminoácido del terminal C. Los aminoácidos del fragmento se indican por el mismo número que el correspondiente aminoácido de GLP-1(7-37)OH. Por ejemplo, el ácido glutámico N-terminal de 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(7-36)OH, la glicina de la posición 37 de GLP-1(7-37)OH está eliminada.

Los compuestos de GLP-1 también incluyen polipéptidos en los que se ha añadido uno o más aminoácidos al terminal N y/o al terminal C de GLP-1(7-37)OH, o fragmentos o análogos del mismo. Es preferible que los compuestos de GLP-1 de este tipo tengan hasta aproximadamente treinta y nueve aminoácidos. Los aminoácidos del compuesto de GLP-1 "ampliado" se indican por el mismo número que el correspondiente aminoácido de GLP-1(7-37)OH. Por ejemplo, el aminoácido N-terminal de un compuesto de GLP-1 obtenido mediante la adición de dos aminoácidos en el terminal N de GLP-1(7-37)OH está en la posición 5; y el aminoácido C terminal de un compuesto de GLP-1 obtenido mediante la adición de un aminoácido en el terminal C 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 tanto en estos compuestos de GLP-1 "ampliados" como en GLP-1(7-37)OH. Los aminoácidos 1-6 de un compuesto de GLP-1 ampliado son preferiblemente los mismos que o una sustitución conservativa del aminoácido de la correspondiente posición de GLP-1(1-37)OH. Los aminoácidos 38-45 de un compuesto de GLP-1 ampliado son preferiblemente los mismos que o una sustitución conservativa del aminoácido de la correspondiente posición de glucagón o exendina-4.

Los compuestos de GLP-1 de la presente invención engloban "análogos de GLP-1". Un análogo de GLP-1 tiene una homología suficiente con GLP-1(7-37)OH, o con un fragmento de GLP-1(7-37)OH, tal que el análogo tiene actividad

insulinotrópica. Preferiblemente, un análogo de GLP-1 tiene una secuencia de aminoácidos de GLP-1(7-37)OH, o un fragmento de la misma, modificada tal que uno, dos, tres, cuatro o cinco aminoácidos difieren del aminoácido de la correspondiente posición de GLP-1(7-37)OH o de un fragmento de GLP-1(7-37)OH. En la nomenclatura usada en la presente memoria para designar los compuestos GLP-1, el aminoácido sustituyente y su posición están indicados antes de la estructura parental. Por ejemplo, Glu^{22}-GLP-1(7-37)OH designa un compuesto de GLP-1 en el que la glicina normalmente encontrada en la posición 22 de GLP-1(7-37)OH ha sido reemplazada por ácido glutámico; Val^{8}-Glu^{22}-GLP-1(7-37)OH designa un compuesto de GLP-1 en el que la alanina normalmente encontrada en la posición 8 y la glicina normalmente encontrada en la posición 22 de GLP-1(7-37)OH han sido reemplazadas por valina y ácido glutámico, respectivamente.

Los compuestos de 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 también tiene la modificación química de uno o más de sus grupos laterales de aminoácido, átomos α-carboxilo, grupo amino terminal o grupo de ácido carboxílico terminal. Una modificación química incluye, pero no se limita a, la adición de restos químicos, la creación de nuevos enlaces y la eliminación de restos químicos. Las modificaciones en los grupos laterales de aminoácido incluyen, sin limitación, la acilación de grupos ε-amino de lisina, la N-alquilación de arginina, histidina o lisina, la alquilación de grupos de ácido carboxílico glutámico o aspártico y la desamidación de glutamina o asparagina. Las modificaciones del grupo amino terminal incluyen, sin limitación, las modificaciones desamino, de *N*-alquilo inferior, de *N*-di-alquilo inferior y de *N*-acilo. Las modificaciones del grupo carboxilo terminal incluyen, sin limitación, las modificaciones de amidas, alquil-amidas inferiores, dialquil-amidas y alquil-ésteres inferiores. Alquilo inferior es alquilo(C_1-C_4). Además, se pueden proteger uno o más grupos laterales o grupos terminales mediante grupos protectores conocidos por el experto en Química de Proteínas. El α-carbono de un aminoácido puede estar mono o dimetilado.

Cualquier compuesto de GLP-1 puede formar parte de las proteínas de fusión heterólogas de la presente invención, siempre y cuando el propio compuesto de GLP-1 pueda unirse a e inducir la señalización a través del receptor de GLP-1. La unión al receptor de GLP-1 y la transducción de señales se puede evaluar usando análisis *in vitro* tales como los descritos en el documento EP 619.322 y la patente estadounidense n.º 5.120.712, respectivamente.

Se conocen numerosos fragmentos, análogos y derivados de GLP-1 activos en la técnica, y cualquiera de estos análogos y derivados también puede formar parte de las proteínas de fusión heterólogas de la presente invención. En la presente memoria, se proporcionan algunos ejemplos de nuevos análogos, 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), Gln^{9}-GLP-1(7-37), D-Gln^{9}-GLP-1(7-37), Thr^{16}-Lys^{18}-GLP-1(7-37) y Lys^{18}-GLP-1(7-37). En la patente estadounidense n.º 5.118.666, se revelan análogos de GLP-1 tales como GLP-1(7-34) y GLP-1(7-35). También se conocen formas procesadas biológicamente de GLP-1 que tienen propiedades insulinotrópicas, tales como GLP-1(7-36). En la patente estadounidense n.º 5.977.071 concedida a Hoffmann, *et al*., la patente estadounidense n.º 5.545.618 concedida a Buckley, *et al*., y Adelhorst, *et al., J. Biol. Chem*. 269: 6275 (1994), se revelan otros compuestos de GLP-1 biológicamente activos conocidos.

Un grupo preferido de análogos de GLP-1 está compuesto por análogos de GLP-1 de fórmula I (SEC ID N.º 2)

```
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
Ile-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-
40   41   42   43   44   45
Xaa-Xaa-Xaa-Xaa-Xaa-Xaa
```

Fórmula I (SEC ID N.º 2)

en la que:

El Xaa de la posición 8 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

El Xaa de la posición 9 es Glu, Asp o Lys;

El Xaa de la posición 11 es Thr, Ala, Gly, Ser, Leu, Ile, Val, Glu, Asp o Lys;

El Xaa de la posición 14 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp o Lys;

El Xaa de la posición 16 es Val, Ala, Gly, Ser, Thr, Leu, Ile, Tyr, Glu, Asp, Trp o Lys;

El Xaa de la posición 17 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp o Lys;

El Xaa de la posición 18 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, Trp, Tyr o Lys;

El Xaa de la posición 19 es Tyr, Phe, Trp, Glu, Asp, Gln o Lys;

El Xaa de la posición 20 es Leu, Ala, Gly, Ser, Thr, Ile, Val, Glu, Asp, Met, Trp, Tyr o Lys;

El Xaa de la posición 21 es Glu, Asp o Lys;

El Xaa de la posición 22 es Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

El Xaa de la posición 23 es Gln, Asn, Arg, Glu, Asp o Lys;

El Xaa de la posición 24 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Arg, Glu, Asp o Lys;

El Xaa de la posición 25 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

El Xaa de la posición 26 es Lys, Arg, Gln, Glu, Asp o His;

El Xaa de la posición 27 es Leu, Glu, Asp o Lys;

El Xaa de la posición 30 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

El Xaa de la posición 31 es Trp, Phe, Tyr, Glu, Asp o Lys;

El Xaa de la posición 32 es Leu, Gly, Ala, Ser, Thr, Ile, Val, Glu, Asp o Lys;

El Xaa de la posición 33 es Val, Gly, Ala, Ser, Thr, Leu, Ile, Glu, Asp o Lys;

El Xaa de la posición 34 es Asn, Lys, Arg, Glu, Asp o His;

El Xaa de la posición 35 es Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

El Xaa de la posición 36 es Gly, Arg, Lys, Glu, Asp o His;

El Xaa de la posición 37 es Pro, Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys, o está eliminado;

El Xaa de la posición 38 es Ser, Arg, Lys, Glu, Asp o His, o está eliminado;

El Xaa de la posición 39 es Ser, Arg, Lys, Glu, Asp o His, o está eliminado;

El Xaa de la posición 40 es Gly, Asp, Glu o Lys, o está eliminado;

El Xaa de la posición 41 es Ala, Phe, Trp, Tyr, Glu, Asp o Lys, o está eliminado;

El Xaa de la posición 42 es Ser, Pro, Lys, Glu o Asp, o está eliminado;

El Xaa de la posición 43 es Ser, Pro, Glu, Asp o Lys, o está eliminado;

El Xaa de la posición 44 es Gly, Pro, Glu, Asp o Lys, o está eliminado;

y

El Xaa de la posición 45 es Ala, Ser, Val, Glu, Asp o Lys, o está eliminado;

con la condición de que cuando el aminoácido de la posición 37, 38, 39, 40, 41, 42, 43 ó 44 esté eliminado, entonces cada aminoácido secuencia abajo del aminoácido también está eliminado.

Es preferible que el compuesto de GLP-1 de fórmula I contenga menos de seis aminoácidos que difieran del correspondiente aminoácido de GLP-1(7-37)OH o exendina-4. Es más preferible que menos de cinco aminoácidos difieran del correspondiente aminoácido de GLP-1(7-37)OH o exendina-4. Es incluso más preferible que menos de cuatro aminoácidos difieran del correspondiente aminoácido de GLP-1(7-37)OH o exendina-4.

Los compuestos de GLP-1 de la presente invención incluyen derivados de fórmula I tales como un éster(C_1-C_6), o una amida o una alquilamida(C_1-C_6) o una dialquilamida(C_1-C_6) de los mismos. El documento WO99/43706 describe derivados de compuestos de GLP-1 de fórmula I.

Los compuestos de fórmula I derivados según lo descrito en el documento WO99/43706 y no derivados están englobados por la presente invención.

Otro grupo preferido de compuestos de GLP-1 está compuesto por análogos de GLP-1 de fórmula II (SEC ID N.º 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
```

Fórmula II (SEC ID N.º 3)

en la que:

el Xaa de la posición 7 es: L-histidina, D-histidina, desamino-histidina, 2-amino-histidina, β-hidroxi-histidina, homohistidina, α-fluorometil-histidina o α-metil-histidina;

el Xaa de la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser o Thr;

el Xaa de la posición 9 es: Thr, Ser, Arg, Lys, Trp, Phe, Tyr, Glu o His;

el Xaa de la posición 11 es: Asp, Glu, Arg, Thr, Ala, Lys o His;

el Xaa de la posición 12 es: His, Trp, Phe o Tyr;

el Xaa de la posición 16 es: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Tyr, Glu o Ala;

el Xaa de la posición 18 es: His, Pro, Asp, Glu, Arg, Ser, Ala o Lys;

el Xaa de la posición 19 es: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;

el Xaa de la posición 23 es: His, Asp, Lys, Glu, Gln o Arg;

el Xaa de la posición 24 es: Glu, Arg, Ala o Lys;

el Xaa de la posición 26 es: Trp, Tyr, Phe, Asp, Lys, Glu o His;

el Xaa de la posición 27 es: Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;

el Xaa de la posición 30 es: Ala, Glu, Asp, Ser o His;

el Xaa de la posición 31 es: Asp, Glu, Ser, Thr, Arg, Trp o Lys;

el Xaa de la posición 33 es: Asp, Arg, Val, Lys, Ala, Gly o Glu;

el Xaa de la posición 34 es: Glu, Lys o Asp;

el Xaa de la posición 35 es: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o Glu;

el Xaa de la posición 36 es: Thr, Ser, Asp, Trp, Tyr, Phe, Arg, Glu o His;

el R de la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, o está eliminado.

Otro grupo preferido de compuestos de GLP-1 está compuesto por análogos de GLP-1 de fórmula III (SEC ID N.º 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-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.º 4)

en la que:

el Xaa de la posición 7 es: L-histidina, D-histidina, desamino-histidina, 2-amino-histidina, β-hidroxi-histidina, homohistidina, α-fluorometil-histidina α-metil-histidina;

el Xaa de la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser o Thr;

el Xaa de la posición 11 es: Asp, Glu, Arg, Thr, Ala, Lys o His;

el Xaa de la posición 12 es: His, Trp, Phe o Tyr;

el Xaa de la posición 16 es: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu o Ala;

el Xaa de la posición 22: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;

el Xaa de la posición 23 es: His, Asp, Lys, Glu o Gln;

el Xaa de la posición 24 es: Glu, His, Ala o Lys;

el Xaa de la posición 25 es: Asp, Lys, Glu o His;

el Xaa de la posición 27 es: Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;

el Xaa de la posición 30 es: Ala, Glu, Asp, Ser o His;

el Xaa de la posición 33 es: Asp, Arg, Val, Lys, Ala, Gly o Glu;

el Xaa de la posición 34 es: Glu, Lys o Asp;

el Xaa de la posición 35 es: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o Glu;

el Xaa de la posición 36 es: Arg, Glu o His;

R de la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, o está eliminado.

Otro grupo preferido de compuestos de GLP-1 está compuesto por análogos de GLP-1 de fórmula IV (SEC ID N.º 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-R
```

fórmula IV (SEC ID N.º 5)

en la que:

el Xaa de la posición 7 es: L-histidina, D-histidina, desamino-histidina, 2-amino-histidina, β-hidroxi-histidina, homohistidina, α-fluorometil-histidina α-metil-histidina;

el Xaa de la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser, Met o Thr;

el Xaa de la posición 12 es: His, Trp, Phe o Tyr;

el Xaa de la posición 16 es: Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu o Ala;

el Xaa de la posición 22 es: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;

el Xaa de la posición 23 es: His, Asp, Lys, Glu o Gln;

el Xaa de la posición 26 es: Asp, Lys, Glu o His;

el Xaa de la posición 30 es: Ala, Glu, Asp, Ser o His;

el Xaa de la posición 35 es: Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o Glu;

R de la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, o está eliminado.

Otro grupo preferido de compuestos de GLP-1 está compuesto por análogos de GLP-1 de fórmula V (SEC ID N.º 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-Xaa-Ala-Lys-Glu-Phe-
29  30  31  32  33  34  35  36  37
Ile-Xaa-Trp-Leu-Val-Lys-Gly-Arg-R
```

fórmula V (SEC ID N.º 6)

en la que:

el Xaa de la posición 7 es: L-histidina, D-histidina, desamino-histidina, 2-amino-histidina, β-hidroxi-histidina, homohistidina, α-fluorometil-histidina o α-metil-histidina;

el Xaa de la posición 8 es: Gly, Ala, Val, Leu, Ile, Ser o Thr;

el Xaa de la posición 22 es: Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;

el Xaa de la posición 23 es: His, Asp, Lys, Glu o Gln;

el Xaa de la posición 24 es: Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;

el Xaa de la posición 30 es: Ala, Glu, Asp, Ser o His;

el R de la posición 37 es: Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, o está eliminado.

Los compuestos de GLP-1 de fórmula I, II, III, IV y V preferidos comprenden análogos de GLP-1 o fragmentos de análogos de GLP-1, conteniendo los análogos o los fragmentos un aminoácido distinto de 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 en comparación con el correspondiente aminoácido del GLP-1(7-37)OH nativo. También es preferible que estos análogos tengan 6 o menos cambios en comparación con los correspondientes aminoácidos del GLP-1(7-37)OH o GLP-1(7-36)OH nativo. Los análogos más preferidos tienen 5 o menos cambios en comparación con los correspondientes aminoácidos del GLP-1(7-37)OH o GLP-1(7-36)OH nativo, o tienen 4 o menos cambios en comparación con los correspondientes aminoácidos del GLP-

1(7-37)OH o GLP-1(7-36)OH nativo. Es incluso más preferible que estos análogos tengan 3 o menos cambios en comparación con los correspondientes aminoácidos del GLP-1(7-37)OH o GLP-1(7-36)OH nativo. Lo más preferible es que estos análogos tengan 2 o menos cambios en comparación con los correspondientes aminoácidos del GLP-1(7-37)OH nativo.

Se ha descubierto que los compuestos de fórmula II, III, IV y V tienen una propensión reducida a agregarse y a generar formas insolubles. Esto también es importante en el contexto de las proteínas de fusión, en el que un péptido GLP-1 relativamente pequeño debe mantener una configuración activa a pesar de ser fusionado con una proteína mucho mayor. Los compuestos de GLP-1 de fórmula II, III, IV y V preferidos englobados 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 de la posición 22 y, preferiblemente, la alanina de la posición 8 han sido reemplazadas por otro aminoácido.

Cuando la posición 22 es ácido aspártico, ácido glutámico, arginina o lisina, la posición 8 es preferiblemente, glicina, valina, leucina, isolecina, serina, treonina o metionina, y más preferiblemente, valina o glicina. Cuando la posición 22 es ácido sulfónico, tal como ácido cisteico, la posición 8 es preferiblemente, glicina, valina, leucina, isolecina, serina, treonina o metionina, y más preferiblemente, valina o glicina.

Otros compuestos de GLP-1 preferidos incluyen análogos de GLP-1 de fórmula IV (SEC ID N.º 5), teniendo los análogos la secuencia de GLP-1(7-37)OH a excepción de que el aminoácido de la posición 8 es preferiblemente glicina, valina, leucina, isolecina, 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 glutamático.

Otros compuestos de GLP-1 preferidos incluyen análogos de GLP-1 de fórmula IV (SEC ID N.º 5), teniendo los análogos la secuencia de GLP-1(7-37)OH a excepción de que el aminoácido de la posición 8 es preferiblemente glicina, valina, leucina, isolecina, serina, treonina o metionina, y más preferiblemente, valina o glicina, y la posición 37 es histidina, lisina, arginina, treonina, serina, ácido glutámico, ácido aspártico, triptófano, tirosina, fenilalanina y más preferiblemente, histidina.

Otros compuestos de GLP-1 preferidos incluyen análogos de GLP-1 de fórmula IV (SEC ID N.º 5), teniendo los análogos la secuencia de GLP-1(7-37)OH a excepción de que el aminoácido de la posición 8 es preferiblemente glicina, valina, leucina, isoleucina, serina, treonina o metionina, y más preferiblemente, valina o glicina, y la posición 22 es ácido glutámico, lisina, ácido aspártico o arginina, y más preferiblemente, ácido glutámico o lisina, y la posición 23 es lisina, arginina, ácido glutámico, ácido aspártico e histidina, y más preferiblemente, lisina o ácido glutámico.

Otros compuestos de GLP-1 preferidos incluyen análogos de GLP-1 de fórmula V (SEC ID N.º 6), teniendo los análogos la secuencia de GLP-1(7-37)OH a excepción de que el aminoácido de la posición 8 es preferiblemente glicina, valina, leucina, isoleucina, serina, treonina o metionina, y más preferiblemente, valina o glicina, y la posición 22 es ácido glutámico, lisina, ácido aspártico o arginina, y más preferiblemente, ácido de glutamina o lisina, y la posición 27 es alanina, lisina, arginina, triptófano, tirosina, fenilalanina o histidina, y más preferiblemente, alanina.

Otros compuestos de GLP-1 preferidos incluyen análogos de GLP-1 de fórmula II, teniendo los análogos la secuencia GLP-1(7-37)OH a excepción de que el aminoácido de la posición 8 y 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 de la correspondiente posición del GLP-1(7-37)OH nativo.

Otros compuestos de GLP-1 de fórmula II preferidos incluyen: Val^{8}-GLP-1(7-37)OH, Gly^{8}-GLP-1(7-37)OH, Glu^{22}-GLP-1(7-37)OH, Asp^{22}-GLP-1(7-37)OH, Arg^{22}-GLP-1(7-37)OH, Lys^{22}-GLP-1(7-37)OH, Cys^{22}-GLP-1(7-37)OH, Val^{8}-Glu^{22}-GLP-1(7-37)OH, Val^{8}-Asp^{22}-GLP-1(7-37)OH, Val^{8}-Arg^{22}-GLP-1(7-37)OH, Val^{8}-Lys^{22}-GLP-1(7-37)OH, Val^{8}-Cys^{22}-GLP-1(7-37)OH, Gly^{8}-Glu^{22}-GLP-1(7-37)OH, Gly^{8}-Asp^{22}-GLP-1(7-37)OH, Gly^{8}-Arg^{22}-GLP-1(7-37)OH, Gly^{8}-Lys^{22}-GLP-1(7-37)OH, Gly^{8}-Cys^{22}-GLP-1(7-37)OH, Glu^{22}-GLP-1(7-36)OH, Asp^{22}-GLP-1(7-36)OH, Arg^{22}-GLP-1 (7-36)OH, Lys^{22}-GLP-1(7-36)OH, Cys^{22}-GLP-1(7-36)OH, Val^{8}-Glu^{22}-GLP-1(7-36)OH, Val^{8}-Asp^{22}-GLP-1(7-36)OH, Val^{8}-Arg^{22}-GLP-1(7-36)OH, Val^{8}-Lys^{22}-GLP-1(7-36)OH, Val^{8}-Cys^{22}-GLP-1(7-36)OH, Gly^{8}-Glu^{22}-GLP-1(7-36)OH, Gly^{8}-Asp^{22}-GLP-1(7-36)OH, Gly^{8}-Arg^{22}-GLP-1(7-36)OH, Gly^{8}-Lys^{22}-GLP-1(7-36)OH, Gly^{8}-Cys^{22}-GLP-1(7-36)OH, Lys^{23}-GLP-1(7-37)OH, Val^{8}-Lys^{23}-GLP-1(7-37)OH, Gly^{8}-Lys^{23}-GLP-1(7-37)OH, His^{24}-GLP-1(7-37)OH, Val^{8}-His^{24}-GLP-1(7-37)OH, Gly^{8}-His^{24}-GLP-1(7-37)OH, Lys^{24}-GLP-1(7-37)OH, Val^{8}-Lys^{24}-GLP-1(7-37)OH, Gly^{8}-Lys^{23}-GLP-1(7-37)OH, Glu^{30}-GLP-1(7-37)OH, Val^{8}-Glu^{30}-GLP-1(7-37)OH, Gly^{8}-Glu^{30}-GLP-1(7-37)OH, Asp^{30}-GLP-1 (7-37)OH, Val^{8}-Asp^{30}-GLP-1(7-37)OH, Gly^{8}-Asp^{30}-GLP-1(7-37)OH, Gln^{30}-GLP-1(7-37)OH, Val^{8}-Gln^{30}-GLP-1(7-37)OH, Gly^{8}-Gln^{30}-GLP-1(7-37)OH, Tyr^{30}-GLP-1(7-37)OH, Val^{8}-Tyr^{30}-GLP-1(7-37)OH, Gly^{8}-Tyr^{30}-GLP-1(7-37)OH, Ser^{30}-GLP-1(7-37)OH, Val^{8}-Ser^{30}-GLP-1(7-37)OH, Gly^{8}-Ser^{30}-GLP-1(7-37)OH, His^{30}-GLP-1(7-37)OH, Val^{8}-His^{30}-GLP-1(7-37)OH, Gly^{8}-His^{30}-GLP-1(7-37)OH, Glu^{34}-GLP-1(7-37)OH, Val^{8}-Glu^{34}-GLP-1(7-37)OH, Gly^{8}-Glu^{34}-GLP-1(7-37)OH, Ala^{34}-GLP-1(7-37)OH, Val^{8}-Ala^{34}-GLP-1(7-37)OH, Gly^{8}-Ala^{34}-GLP-1(7-37)OH, Gly^{34}-GLP-1(7-37)OH, Val^{8}-Gly^{34}-GLP-1(7-37)OH, Gly^{8}-Gly^{34}-GLP-1(7-37)OH, Ala^{35}-GLP-1(7-37)OH, Val^{8}-Ala^{35}-GLP-1(7-37)OH, Gly^{8}-Ala^{35}-GLP-1(7-37)OH, Lys^{35}-GLP-1(7-37)OH, Val^{8}-Lys^{35}-GLP-1(7-37)OH, Gly^{8}-Lys^{35}-GLP-1(7-37)OH, His^{35}-GLP-1(7-37)OH Val^{8}-His^{35}-GLP-1(7-37)OH, Gly^{8}-His^{35}-GLP-1(7-37)OH, Pro^{35}-GLP-1(7-37)OH, Val^{8}-Pro^{35}-GLP-1(7-37)OH, Gly^{8}-Pro^{35}-GLP-1(7-37)OH, Glu^{35}-GLP-1(7-37)OH Val^{8}-Glu^{35}-GLP-1(7-37)OH, Gly^{8}-

Glu^{35}-GLP-1(7-37)OH, Val^{8}-Ala27-GLP-1(7-37)OH, Val^{8}-His37-GLP-1(7-37)OH, Val^{8}-Glu^{22}-Lys^{23}-GLP-1(7-37) OH, Val^{8}-Glu^{22}-Glu^{23}-GLP-1(7-37)OH, Val^{8}-Glu^{22}-Ala^{27}-GLP-1(7-37)OH, Val^{8}-Gly^{34}-Lys^{35}-GLP-1(7-37)OH, Val^{8}-His^{37}-GLP-1(7-37) OH, Gly^{8}-His^{37}-GLP-1(7-37)OH, Val^{8}-Glu^{22}-Ala^{27}-GLP-1(7-37)OH, Gly^{8}-Glu^{22}-Ala^{27}-GLP-1(7-37)OH, Val^{8}-Lys^{22}-Glu^{23}-GLP-1(7-37)OH, y Gly^{8}-Lys^{22}-Glu^{23}-GLP-1(7-37)OH.

Otro grupo preferido de análogos y derivados de GLP-1 para su uso en la presente invención está compuesto por moléculas de fórmula VI (SEC ID N.º 7):

$$R_1\text{-}X\text{-Glu-Gly}^{10}\text{-Thr-Phe-Thr-Ser-Asp}^{15}\text{-Val-Ser-Ser-Tyr-Leu}^{20}\text{-}Y\text{-Gly-Gln-Ala-Ala}^{25}\text{-Lys-}Z\text{-Phe-Ile-Ala}^{30}\text{-Trp-Leu-Val-Lys-Gly}^{35}\text{-Arg-}R_2$$

fórmula VI (SEC ID N.º 7)

en la que: R_1 se selecciona del grupo constituido por L-histidina, D-histidina, desamino-histidina, 2-amino-histidina, β-hidroxi-histidina, homohistidina, alfa-fluorometil-histidina y alfa-metil-histidina; 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 R_2 es Gly-OH.

Otro grupo preferido de compuestos de GLP-1 para su uso en la presente invención se revela en el documento WO91/11457, y está esencialmente constituido por GLP-1(7-34), GLP-1(7-35), GLP-1(7-36) o GLP-1(7-37), o la forma amida de los mismos, y sus sales farmacéuticamente aceptables, que tienen al menos una modificación seleccionada del grupo constituido por:

(a) sustitución de glicina, serina, cisteína, treonina, asparagina, glutamina, tirosina, alanina, valina, isoleucina, leucina, metionina, fenilalanina, arginina o D-lisina por lisina en la posición 26 y/o en la posición 34; o la sustitución de glicina, serina, cisteína, treonina, asparagina, glutamina, tirosina, alanina, valina, isoleucina, leucina, metionina, fenilalanina, lisina o D-arginina por arginina en la posición 36;

(b) sustitución de un aminoácido resistente a la oxidación por triptófano en la posición 31:

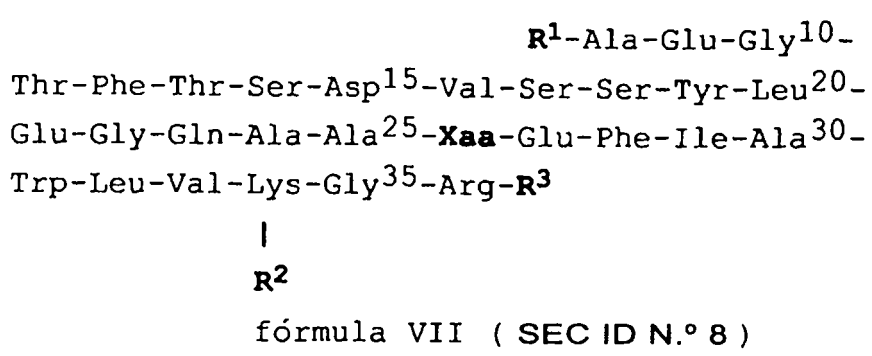
(c) sustitución de al menos uno entre: tirosina por valina en la posición 16; lisina por serina en la posición 18; ácido aspártico por ácido glutámico en la posición 21; serina por glicina en la posición 22; arginina por glutamina en la posición 23; arginina por alanina en la posición 24; y glutamina por lisina en la posición 26; y

(d) sustitución de al menos una entre: glicina, serina o cisteína por alanina en la posición 8; ácido aspártico, glicina, serina, cisteína, treonina, asparagina, glutamina, tirosina, alanina, valina, isoleucina, leucina, metionina o fenilalanina por ácido glutámico en la posición 9; serina, cisteína, treonina, asparagina, glutamina, tirosina, alanina, valina, isoleucina, leucina, metionina o fenilalanina por glicina en la posición 10; y ácido glutámico por ácido aspártico en la posición 15; y

(e) sustitución de 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 por histidina en la posición 7; en la que, en las sustituciones (a), (b), (d) y (e), los aminoácidos sustituidos pueden estar opcionalmente en la forma D y los aminoácidos sustituidos en la posición 7 pueden estar opcionalmente en la forma N-acilada o N-alquilada.

Como la enzima dipeptidil-peptidasa IV (DPP IV) puede ser responsable de la rápida desactivación *in vivo* observada del GLP-1 administrado, [véase, p.ej., Mentlein, R., *et al., Eur. J. Biochem.*, 214: 829-835 (1993)], se prefieren los análogos y derivados de GLP-1 que están protegidos de la actividad de la DPP IV en el contexto de una proteína de fusión, siendo más preferidas las proteínas de fusión en las que el compuesto de GLP-1 es Gly^{8}-GLP-1(7-37)OH, Val^{8}-GLP-1(7-37)OH, α-metil-Ala^{8}-GLP-1(7-37)OH o Gly^{8}-Gln^{21}-GLP-1(7-37)OH.

Otro grupo preferido de compuestos de GLP-1 para su uso en la presente invención está constituido por compuestos de fórmula VII (SEC ID N.º 8) reivindicados en la patente estadounidense n.º 5.512.549:



fórmula VII (SEC ID N.º 8)

en la que R^1 se selecciona del grupo constituido por 4-imidazopropionilo, 4-imidazoacetilo o 4-imidazo-α,α-dimetilacetilo; R^2 se selecciona del grupo constituido por acilo(C_6-C_{10}) sin ramificar, o está ausente; R^3 se selecciona del grupo constituido por Gly-OH o NH_2; y Xaa es Lys o Arg.

Los compuestos de fórmula IV más preferidos para su uso en la presente invención son aquellos en los que Xaa es Arg y R^2 es acilo(C_6-C_{10}) no ramificado. Los compuestos de fórmula IV todavía más preferidos para su uso en la presente invención son aquellos en los que Xaa es Arg, R^2 es acilo(C_6-C_{10}) no ramificado y R^3 es Gly-OH. Otros compuestos de fórmula IV muy preferidos para su uso en la presente invención son aquellos en los que Xaa es Arg, R^2 es acilo(C_6-C_{10}) no ramificado, R^3 es Gly-OH y R^1 es 4-imidazopropionilo. Un compuesto de fórmula IV especialmente preferido para su uso en la presente invención es aquél en el que Xaa es Arg, R^2 es acilo(C_8) no ramificado, R^3 es Gly-OH y R^1 es 4-imidazopropionilo.

Preferiblemente, los compuestos de GLP-1 comprenden análogos de GLP-1 en los que la estructura para tales 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 desamino-histidina, 2-amino-histidina, β-hidroxi-histidina, homohistidina, α-fluorometil-histidina o α-metil-histidina en la posición 7. Es preferible que estos análogos de la posición 8 contengan uno o más cambios adicionales en las posiciones 12, 16, 18, 19, 20, 22, 25, 27, 30, 33 y 37 en comparación con el correspondiente aminoácido del GLP-1(7-37)OH nativo. Es más preferible que estos análogos de posición 8 contengan uno o más cambios adicionales en las posiciones 16, 18, 22, 25 y 33 en comparación con el correspondiente aminoácido del 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 de la posición 12 se selecciona del grupo constituido por triptófano o tirosina. Es más preferible que, además de la sustitución de la posición 12, el aminoácido de la posición 8 esté sustituido por glicina, valina, leucina, isoleucina, serina, treonina o metionina, y más preferiblemente, por valina o glicina. Es incluso más preferible que además de las sustituciones de la posición 12 y 8, el aminoácido de la posición 22 esté sustituido por ácido glutámico.

En otra realización preferida, el análogo de GLP-1 es GLP-1(7-37)OH, en el que el aminoácido de la posición 16 se selecciona del grupo constituido por triptófano, isoleucina, leucina, fenilalanina o tirosina. Es más preferible que además de la sustitución de la posición 16, el aminoácido de la posición 8 esté sustituido por glicina, valina, leucina, isolecina, serina, treonina o metionina, y más preferiblemente, por valina o glicina. Es incluso más preferible que además de las sustituciones de las posiciones 16 y 8, el aminoácido de la posición 22 esté sustituido por ácido glutámico. También es preferible que además de las sustituciones de las posiciones 16 y 8, el aminoácido de la posición 30 esté sustituido por ácido glutámico. También es preferible que además de las sustituciones de las posiciones 16 y 8, el aminoácido de la posición 37 esté sustituido por histidina.

En otra realización preferida, el análogo de GLP-1 es GLP-1(7-37)OH, en el que el aminoácido de la posición 18 se selecciona del grupo constituido por triptófano, tirosina, fenilalanina, lisina, leucina o isoleucina, preferiblemente, triptófano, tirosina e isoleucina. Es más preferible que además de la sustitución de la posición 18, el aminoácido de la posición 8 esté sustituido por glicina, valina, leucina, isolecina, serina, treonina o metionina, y más preferiblemente, por valina o glicina. Es incluso más preferible que además de las sustituciones de las posiciones 18 y 8, el aminoácido de la posición 22 esté sustituido por ácido glutámico. También es preferible que además de las sustituciones de las posiciones 18 y 8, el aminoácido de la posición 30 esté sustituido por ácido glutámico. También es preferible que además de las sustituciones de las posiciones 18 y 8, el aminoácido de la posición 37 esté sustituido por histidina.

En otra realización preferida, el análogo de GLP-1 es GLP-1(7-37)OH, en el que el aminoácido de la posición 19 se selecciona del grupo constituido por triptófano, o fenilalanina, preferiblemente, triptófano. Es más preferible que además de la sustitución de la posición 19, el aminoácido de la posición 8 esté sustituido por glicina, valina, leucina, isolecina, serina, treonina o metionina, y más preferiblemente, por valina o glicina. Es incluso más preferible que además de las sustituciones de las posiciones 19 y 8, el aminoácido de la posición 22 esté sustituido por ácido glutámico. También es preferible que además de las sustituciones de las posiciones 19 y 8, el aminoácido de la posición 30 esté sustituido por ácido glutámico. También es preferible que además de las sustituciones de las posiciones 19 y 8, el aminoácido de la posición 37 esté sustituido por histidina.

En otra realización preferida, el análogo de GLP-1 es GLP-1(7-37)OH, en el que el aminoácido de la posición 20 es fenilalanina, tirosina o triptófano. Es más preferible que, además de la sustitución de la posición 20, el aminoácido de la posición 8 esté sustituido por glicina, valina, leucina, isolecina, serina, treonina o metionina, y más preferiblemente, por valina o glicina. Es incluso más preferible que, además de las sustituciones de las posiciones 20 y 8, el aminoácido de la posición 22 esté sustituido por ácido glutámico. También es preferible que, además de las sustituciones de las posiciones 20 y 8, el aminoácido de la posición 30 esté sustituido por ácido glutámico. También es preferible que, además de las sustituciones de las posiciones 20 y 8, el aminoácido de la posición 37 esté sustituido por histidina.

En otra realización preferida, el análogo de GLP-1 es GLP-1(7-37)OH, en el que el aminoácido de la posición 25 se selecciona del grupo constituido por valina, isoleucina y leucina, preferiblemente, valina. Es más preferible que, además de la sustitución de la posición 25, el aminoácido de la posición 8 esté sustituido por glicina, valina, leucina, isoleucina, serina, treonina o metionina, y más preferiblemente, por valina o glicina. Es incluso más preferible que además de las sustituciones de las posiciones 25 y 8, el aminoácido de la posición 22 esté sustituido por ácido glutámico. También es preferible que, además de las sustituciones de las posiciones 25 y 8, el aminoácido de la posición 30 esté sustituido por ácido glutámico. También es preferible que además de las sustituciones de las posiciones 25 y 8, el aminoácido de la posición 37 esté sustituido por histidina.

En otra realización preferida, el análogo de GLP-1 es GLP-1(7-37)OH, en el que el aminoácido de la posición 27 se selecciona del grupo constituido por isoleucina o alanina. Es más preferible que, además de la sustitución de la posición 27, el aminoácido de la posición 8 esté sustituido por glicina, valina, leucina, isolecina, serina, treonina o metionina, y más preferiblemente, por valina o glicina. Es incluso más preferible que, además de las sustituciones de las posiciones 27 y 8, el aminoácido de la posición 22 esté sustituido por ácido glutámico. También es preferible que, además de las sustituciones de las posiciones 27 y 8, el aminoácido de la posición 30 esté sustituido por ácido glutámico. También es preferible que además de las sustituciones de las posiciones 27 y 8, el aminoácido de la posición 37 esté sustituido por histidina.

En otra realización preferida, el análogo de GLP-1 es GLP-1(7-37)OH, en el que el aminoácido de la posición 33 es isoleucina. Es más preferible que, además de la sustitución de la posición 33, el aminoácido de la posición 8 esté sustituido por glicina, valina, leucina, isoleucina, serina, treonina o metionina, y más preferiblemente, por valina o glicina. Es incluso más preferible que, además de las sustituciones de las posiciones 33 y 8, el aminoácido de la posición 22 esté sustituido por ácido glutámico. También es preferible que además de las sustituciones de las posiciones 33 y 8, el aminoácido de la posición 30 esté sustituido por ácido glutámico. También es preferible que además de las sustituciones de las posiciones 33 y 8, el aminoácido de la posición 37 esté sustituido por 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. Estos compuestos de GLP-1 muestran una mayor potencia en comparación con GLP-1(7-37) OH y comprenden la secuencia de aminoácidos de fórmula IX (SEC ID N.º 12)

Xaa_7-Xaa_8-Glu-Gly-Thr-Xaa_{12}-Thr-Ser-Asp-Xaa_{16}-Ser-
Xaa_{18}-Xaa_{19}-Xaa_{20}-Glu-Xaa_{22}-Gln-Ala-Xaa_{25}-Lys-Xaa_{27}-
Phe-Ile-Xaa_{30}-Trp-Leu-Xaa_{33}-Lys-Gly-Arg-Xaa_{37}

Fórmula IX (SEC ID N.º 12)

en la que:

Xaa_7 es: L-histidina, D-histidina, desamino-histidina, 2-amino-histidina, β-hidroxi-histidina, homohistidina, α-fluorometil-histidina o α-metil- histidina;

Xaa_8 es: Ala, Gly, Val, Leu, Ile, Ser o Thr;

Xaa_{12} es: Phe, Trp o Tyr;

Xaa_{16} es: Val, Trp, Ile, Leu, Phe o Tyr;

Xaa_{18} es: Ser, Trp, Tyr, Phe, Lys, Ile, Leu, Val;

Xaa_{19} es: Tyr, Trp o Phe;

Xaa_{20} es: Leu, Phe, Tyr o Trp;

Xaa_{22} es: Gly, Glu, Asp o Lys;

Xaa_{25} es: Ala, Val, Ile o Leu;

Xaa_{27} es: Glu, Ile o Ala;

Xaa_{30} es: Ala o Glu;

Xaa_{33} es: Val o Ile; y

Xaa_{37} es: Gly, His, NH_2 o está ausente.

Algunos compuestos de GLP-1 de fórmula IX preferidos incluyen GLP-1(7-37)OH, GLP-1(7-36)-NH_2, Gly^8-GLP-1(7-37)OH, Gly^8-GLP-1(7-36)NH_2, Val^8-GLP-1(7-37)OH, Val^8-GLP-1(7-36)NH_2, Leu^8-GLP-1(7-37)OH, Leu^8-GLP-1(7-36)NH_2, Ile^8-GLP-1(7-37)OH, Ile^8-GLP-1(7-36)NH_2, Ser^8-GLP-1(7-37)OH, Ser^8-GLP-1(7-36)NH_2, Thr^8-GLP-1(7-37)OH, Thr^8-GLP-1(7-36)NH_2, Val^8-Tyr^{12}-GLP-1(7-37)OH, Val^8-Tyr^{12}-GLP-1(7-36)NH_2, Val^8-Tyr^{16}-GLP-1(7-37)OH, Val^8-Ty^{16}-GLP-1(7-36)NH_2, Val^8-Glu^{22}-GLP-1(7-37)OH, Val^8-Glu^{22}-GLP-1(7-36)NH_2, Gly^8-Glu^{22}-GLP-1(7-37)OH, Gly^8-Glu^{22}-GLP-1(7-36)NH_2, Val^8-Asp^{22}-GLP-1(7-37)OH, Val^8-Asp^{22}-GLP-1(7-36)NH_2, Gly^8-Asp^{22}-GLP-1(7-37)OH, Gly^8-Asp^{22}-GLP-1(7-36)NH_2, Val^8-Lys^{22}-GLP-1(7-37)OH, Val^8-Lys^{22}-GLP-1(7-36)NH_2, Gly^8-Lys^{22}-GLP-1(7-37)OH, Gly^8-Lys^{22}-GLP-1(7-36)NH_2, Leu^8-Glu^{22}-GLP-1(7-37)OH, Leu^8-Glu^{22}-GLP-1(7-36)NH_2, Ile^8-Glu^{22}-GLP-1(7-37)OH, Ile^8-Glu^{22}-GLP-1(7-36)NH_2, Leu^8-Asp^{22}-GLP-1(7-37)OH, Leu^8-Asp^{22}-GLP-1(7-36)NH_2, Ile^8-Asp^{22}-GLP-1(7-37)OH, Ile^8-Asp^{22}-GLP-1(7-36)NH_2, Leu^8-Lys^{22}-GLP-1(7-37)OH, Leu^8-Lys^{22}-GLP-1(7-36)NH_2, Ile^6-Lys^{22}-GLP-1(7-37)OH, Ile^8-Lys^{22}-GLP-1(7-36)NH_2, Ser^8-Glu^{22}-GLP-1(7-37)OH, Ser^8-Glu^{22}-GLP-1(7-36)NH_2, Thr^8-Glu^{22}-GLP-1(7-37)OH, Thr^8-Glu^{22}-GLP-1(7-36)NH_2, Ser^8-Asp^{22}-GLP-1(7-37)OH, Ser^8-Asp^{22}-GLP-1(7-36)NH_2, Thr^8-Asp^{22}-GLP-1(7-37)OH, Thr^8-Asp^{22}-GLP-1(7-36)NH_2, Ser^8-Lys^{22}-GLP-1(7-37)OH, Ser^8-Lys^{22}-GLP-1(7-36)NH_2, Thr^8-Lys^{22}-GLP-1(7-37)OH, Thr^8-Lys^{22}-GLP-1(7-36)NH_2, Glu^{22}-GLP-1(7-37)OH, Glu^{22}-GLP-1(7-36)NH_2, Asp^{22}-GLP-1(7-37)OH, Asp^{22}-GLP-14(7-36)NH_2, Lys^{22}-GLP-1(7-37)OH, Lys^{22}-GLP-1(7-36)NH_2, Val^8-Ala^{27}-GLP-1(7-37)OH, Val^8-Glu^{22}-Ala^{27}-GLP-1(7-37)OH, Val^8-Glu^{30}-GLP-1(7-37)OH, Val^8-Glu^{30}-GLP-1(7-36)NH_2, Gly^8-Glu^{30}-GLP-1(7-37)OH, Gly^8-Glu^{30}-GLP-1(7-36)NH_2, Leu^6-Glu^{30}-GLP-1(7-37)OH, Leu^8-Glu^{30}-GLP-1(7-36)NH_2, Ile^8-Glu^{30}-GLP-1(7-37)OH, Ile^8-Glu^{30}-GLP-1(7-36) NH_2, Ser^8-Glu^{30}-GLP-1(7-37)OH, Ser^8-Glu^{30}-GLP-1(7-36)NH_2, Thr^8-Glu^{30}-GLP-1(7-37)OH, Thr^8-Glu^{30}-GLP-1(7-36)NH_2, Val^8-His^{37}-GLP-1(7-37)OH, Val^8-His^{37}-GLP-1(7-36)NH_2, Gly^8-His^{37}-GLP-1(7-37)OH, Gly^8-His^{37}-GLP-1(7-36)NH_2, Leu^8-His^{37}-GLP-1(7-37)OH, Leu^8-His^{37}-GLP-1(7-36)NH_2, Ile^8-His^{37}-GLP-1(7-37)OH, Ile^8-His^{37}-GLP-1(7-36)NH_2, Ser^8-His^{37}-GLP-1(7-37)OH, Ser^8-His^{37}-GLP-1(7-36)NH_2, Thr^8-His^{37}-GLP-1(7-37)OH, Thr^8-His^{37}-GLP-1(7-36)NH_2.

Algunos compuestos de GLP-1 de fórmula IX preferidos que tienen múltiples sustituciones 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 posición 18 es tirosina, triptófano o isoleucina, la posición 25 es valina y la posición 33 es isoleucina. Otros compuestos de GLP-1 preferidos incluyen los siguientes: Val^8-Tyr^{16}-GLP-1(7-37)OH, Val^8-Tyr^{12}-Glu^{22}-GLP-1(7-37)OH, Val^8-Tyr^{16}-Phe^{19}-GLP-1(7-37)OH, Val^8-Tyr^{16}-Glu^{22}-GLP-1(7-37)OH, Val^8-Trp^{16}-Glu^{22}-GLP-1(7-37)OH, Val^{18}-Leu^{16}-Glu^{22}-GLP-1(7-37)OH, Val^8-Ile^6-Glu^{22}-GLP-1(7-37)OH, Val^8-Phe^{16}-Glu^{22}-GLP-1(7-37)OH; Val^8-Trp^{18}-Glu^{22}-GLP-1(7-37)OH, Val^8-Tyr^{18}-Glu^{22}-GLP-1(7-37)OH, Val^8-Phe^{18}-Glu^{22}-GLP-1(7-37)OH y Val^8-Ile^{18}-Glu^{22}-GLP-1(7-37)OH.

Los compuestos de GLP-1 de la presente invención también engloban compuestos de exendina. La exendina-3 y la exendina-4 son péptidos biológicamente activos que fueron aislados por primera vez del veneno de lagartos Heloderma, y han demostrado unirse al receptor de GLP-1 y estimular la producción de H^+ dependiente del AMP cíclico en células parietales de mamífero. La exendina 3 y la exendina-4 son ambos péptidos de 39 aminoácidos que tienen una homología del aproximadamente 53% con el GLP-1. Actúan como potentes agonistas de la actividad de GLP-1. En particular, un derivado truncado N-terminalmente de la exendina, conocido como exendina(aminoácidos 9-39), es un inhibidor de la exendina-3, la exendina-4 y el GLP-1.

Un compuesto de exendina comprende comúnmente un polipéptido que tiene la secuencia de aminoácidos de la exendina-3, la exendina-4, o un análogo o un fragmento de las mismas. La exendina 3 y la exendina 4 se revelan en la patente estadounidense n.º: 5.424.286.

La exendina-3 tiene la secuencia de aminoácidos de SEC ID N.º 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
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
```

(SEC ID N.º 9)

La exendina-4 tiene la secuencia de aminoácidos de SEC ID N.º 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
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
```

(SEC ID N.º 10)

Los compuestos de GLP-1 también incluyen fragmentos de exendina que son polipéptidos obtenidos tras el truncamiento de uno o más aminoácidos del terminal N y/o del terminal C de la exendina o de un análogo de exendina. Además, los compuestos de GLP-1 incluyen polipéptidos de exendina a los que se ha añadido uno o más aminoácidos al terminal N y/o al terminal C de la exendina o fragmentos de la misma. Los compuestos de exendina de este tipo tienen hasta aproximadamente cuarenta y cinco aminoácidos.

Los compuestos de GLP-1 también incluyen "análogos de exendina". Un análogo de exendina tiene una homología suficiente con la exendina-4, la exendina-3 o un fragmento de las mismas tal que el análogo tenga actividad insulinotrópica. Se puede analizar la actividad de los fragmentos y/o los análogos de exendina usando análisis *in vitro* tales como los descritos en el documento EP 619.322 y en la patente estadounidense n.º: 5.120.712.

Preferiblemente, un análogo de exendina tiene una secuencia de aminoácidos de exendina-4 o un fragmento de la misma, modificada tal que uno, dos, tres, cuatro o cinco aminoácidos difieran del aminoácido de la correspondiente posición de la exendina-4 o del fragmento de exendina-4. En la nomenclatura usada en la presente memoria para designar los compuestos de exendina, el aminoácido sustituyente y su posición están indicados antes de la estructura precursora. Por ejemplo, Val^{8}-exendina-4 designa un compuesto de exendina en el que la glicina normalmente encontrada en la posición 8 de la exendina-4 ha sido sustituida por valina.

Otro grupo preferido de compuestos de GLP-1 está compuesto por análogos de GLP-1/exendina-4 de fórmula VIII (SEC ID N.º11):

```
 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
```

fórmula VIII (SEC ID N.º 11)

en la que:

el Xaa de la posición 7 es: L-histidina, D-histidina, desamino-histidina, 2-amino-histidina, β-hidroxi-histidina, homohistidina, α-fluorometil-histidina o α-metil-histidina;

el Xaa de la posición 8 es: Gly, Ala o Val;

el Xaa de la posición 16 es: Leu o Val;

el Xaa de la posición 18 es Lys o Ser;

el Xaa de la posición 19 es: Gln o Tyr;

el Xaa de la posición 20 es: Met o Leu;

el Xaa de la posición 22 es: Glu o Gln;

el Xaa de la posición 23 es: Glu o Gln;

el Xaa de la posición 25 es: Val o Ala;

el Xaa de la posición 26 es: Arg o Lys;

el Xaa de la posición 27 es Leu o Glu;

el Xaa de la posición 30 es: Glu o Ala;

el Xaa de la posición 33 es: Val o Lys;

el Xaa de la posición 34 es: Asn o Lys;

el Xaa de la posición 36 es: Gly o Arg; y

el R de la posición 37 es: Gly, Pro, Pro-Ser-Ser-Gly-Ala-Pro-Pro-Pro-Ser o está ausente. En la tabla 6, se proporciona la actividad de 18 especies diferentes que pertenecen a este género.

En las publicaciones de patente conforme al PCT 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*) y WO 99/43708 (Knudsen *et al*), se describen más análogos de exendina que son útiles para la presente invención.

Las proteínas de fusión GLP-1 de la presente invención pueden comprender sitios de glicosilación. La glicosilación es una modificación química en la que se añaden restos de azúcar a la proteína en sitios específicos. La glicosilación de proteínas desempeña un papel en garantizar la carga, la confirmación y la estabilidad correctas de las proteínas en maduración, y puede dirigir a la proteína a la superficie celular y a la secreción final de la misma. Lo más importante es que la glicosilación efectúa la tasa de aclaramiento *in vivo* para muchas proteínas. Los azúcares pueden estar enlazados a O o a N. En general, los azúcares enlazados a O se añaden al oxígeno del grupo hidroxilo de la serina y la treonina, mientras que los azúcares enlazados a N se añaden al nitrógeno de la amida de la asparagina. El sitio consenso para la N-glicosilación es Asn X1 X2, en el que X1 es cualquier aminoácido a excepción de Pro, y X2 es Ser o Thr.

Proteínas de fusión de albúmina heterólogas

Los compuestos de GLP-1 descritos anteriormente pueden estar fusionados directamente o mediante un ligador peptídico a albúmina o a un análogo, fragmento o derivado de la misma.

Generalmente, las proteínas albúminas que forman parte de las proteínas de fusión de la presente invención pueden estar derivadas de albúmina clonada de cualquier especie. Sin embargo, se prefieren la albúmina humana, y los fragmentos y análogos de la misma, para reducir el riesgo de que la proteína de fusión sea inmunogénica en seres humanos. La albúmina de suero humano (ASH) está constituida por una sola cadena de polipéptido no glicosilado de 585 aminoácidos con un peso molecular en fórmula de 66.500. En la figura 2, se muestra la secuencia de aminoácidos de la ASH humana. (Véase 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]. Se ha descrito una variedad de variantes polimórficas, así como análogos y fragmentos de albúmina. [Véase Weitkamp, *et al.,* (1973) *Ann. Hum. Genet*. 37:219]. Por ejemplo, en el documento EP 322.094, los inventores revelan diversas formas más cortas de ASH. Algunos de estos fragmentos incluyen ASH(1-373), ASH(1-388), ASH(1-389), ASH(1-369) y ASH (1-419), y fragmentos entre 1-369 y 1-419. El documento EP 399.666 revela fragmentos de albúmina que incluyen ASH(1-177) y ASH (1-200), y fragmentos entre ASH(1-177) y ASH(1-200).

Se entiende que las proteínas de fusión heterólogas de la presente invención incluyen compuestos de GLP-1 que están acoplados a cualquier proteína albúmina, incluyendo fragmentos, análogos y derivados, siendo tal proteína de fusión biológicamente activa y teniendo una vida media en plasma más larga que el compuesto de GLP-1 solo. De este modo, la parte de albúmina de la proteína de fusión no necesita tener una vida media en plasma igual a la de la albúmina humana nativa. Los fragmentos, análogos y derivados son conocidos o pueden ser generados tal que tengan vidas medias más largas o vidas medias intermedias a las de la albúmina humana nativa y el compuesto de GLP-1 de interés.

Las proteínas de fusión heterólogas de la presente invención engloban proteínas que tienen sustituciones conservadoras de aminoácidos en el compuesto de GLP-1 y/o en la parte de albúmina de la proteína de fusión. Una "sustitución conservadora" es el reemplazo de un aminoácido por otro aminoácido que tiene la misma carga electrónica neta, y aproximadamente el mismo tamaño y la misma forma. Los aminoácidos con cadenas laterales alifáticas o alifáticas sustituidas de aminoácidos tienen aproximadamente el mismo tamaño cuando el número total de carbonos y de heteroátomos de sus cadenas laterales difiere en no más de aproximadamente cuatro. Tienen aproximadamente la misma forma cuando el número de ramificaciones de sus cadenas laterales difiere en no más de uno. Se considera que los aminoácidos con grupos fenilo o fenilo sustituidos en sus cadenas laterales tienen aproximadamente el mismo tamaño y la misma forma. Excepto que se proporcionen específicamente de otra manera en la presente memoria, las sustituciones conservadoras están hechas preferiblemente con aminoácidos naturales.

Sin embargo, el término "aminoácido" se usa en la presente memoria en su sentido más amplio e incluye aminoácidos naturales, así como aminoácidos no naturales, incluyendo análogos y derivados de aminoácido. Estos últimos incluyen moléculas que contienen un resto de aminoácido. Cualquier experto en la técnica entenderá, en vista de esta amplia definición, que la referencia que se hace en la presente memoria a un aminoácido incluye, por ejemplo, L-aminoácidos proteogénicos naturales; D-aminoácidos; aminoácidos químicamente modificados, tales como análogos y derivados de aminoácido; aminoácidos no proteogénicos naturales, tales como norleucina, β-alanina, ornitina, GABA, etc.; y compuestos sintetizados químicamente que tienen propiedades conocidas en la técnica por ser características de los aminoácidos. Como se usa en la presente memoria, el término "proteogénico" indica que el aminoácido puede ser incorporado en un péptido, polipéptido o proteína de una célula a través de una ruta metabólica.

La incorporación de aminoácidos no naturales, incluyendo aminoácidos no nativos sintéticos, aminoácidos sustituidos, o uno o más D-aminoácidos en las proteínas de fusión heterólogas de la presente invención puede ser ventajosa en un número de diferentes modos. Los péptidos que contienen D-aminoácidos, etc., presentan una mayor estabilidad *in vitro* o *in vivo* en comparación con sus homólogos que contienen L-aminoácidos. De este modo, la construcción de péptidos, etc., incorporando D-aminoácidos puede ser particularmente útil cuando se desee o se necesite una mayor estabilidad intracelular. Más específicamente, los D-péptidos etc., son resistentes a las peptidasas y a las proteasas endógenas, proporcionando así una mejor biodisponibilidad de la molécula y tiempos de vida prolongados *in vivo* cuando se desean tales propiedades. Además, los D-péptidos, etc., no pueden ser procesados eficazmente para una presentación restringida por el complejo de histocompatibilidad de clase II a células auxiliares T, y son por tanto menos propensos a inducir respuestas inmunes humorales en todo el organismo.

Además de los análisis de estructura/función de los diversos polipéptidos englobados por la presente invención, hay numerosos factores que pueden ser considerados al seleccionar los aminoácidos para una sustitución. Un factor que puede ser considerado en la elaboración de tales cambios es el índice hidropático de los aminoácidos. La importancia del índice hidropático de los aminoácidos en conferir función biológica interactiva a una proteína ha sido tratada por Kyte y Doolittle (1982, *J. Mol. Biol.,* 157: 105-132). Está aceptado 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, ADN, anticuerpos, antígenos, etc. En base a su hidrofobicidad y a sus características de carga, se ha asignado a cada aminoácido un índice hidropático según se presenta a continuación: 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, es posible sustituir ciertos aminoácidos de un péptido, polipéptido o proteína por otros aminoácidos que tengan un índice hidropático similar y producir un péptido resultante, etc., que tenga una actividad biológica similar o incluso mejor. Al realizar tales cambios, es preferible que los aminoácidos que tengan índices hidropáticos de ±2 sean sustituidos por otro. Es más preferible que las sustituciones sean aquéllas en las que los aminoácidos tengan índices hidropáticos de ±1. Las sustituciones más preferidas son aquéllas en las que los aminoácidos tienen índices hidropáticos de ±0,5.

Los aminoácidos similares también pueden ser sustituidos en base a la hidrofilidad. La patente estadounidense n.º: 4.554.101 revela que la mayor hidrofilidad media local de una proteína, gobernada por la hidrofilidad de sus aminoácidos adyacentes, se correlaciona con una propiedad biológica de la proteína. Se han asignado los siguientes valores de hidrofilidad a los aminoácidos: arginina/lisina (+3,0); aspartato/glutamato (+3,0±1); serina (+0,3); asparagina/glutamina (+0,2); glicina (0); treonina (-0,4); prolina (-0,51±1); alanina/histidina (-0,5); cisteína (-1,0); metionina (-1,3); valina (-1,5); leucina/isoleucina (-1,8); tirosina (-2,3); fenilalanina (-2,5) y triptófano (-3,4). De este modo, se puede sustituir un aminoácido de un péptido, polipéptido o proteína por otro aminoácido que tenga un índice de hidrofilidad similar y seguir produciendo un péptido resultante, etc., que tenga una actividad biológica similar, i.e., que sigue conservando una función biológica correcta. Al realizar tales cambios, es preferible que los aminoácidos que tengan índices hidropáticos de ±2 sean sustituidos por otro, siendo más preferidos los de ±1 y los más preferidos, los de ±0,5.

Como se señaló anteriormente, las sustituciones de aminoácidos de las proteínas de fusión de la presente invención pueden estar basadas en la similitud relativa de los sustituyentes de cadena lateral de aminoácidos, por ejemplo, su hidrofobicidad, hidrofilidad, carga, tamaño, etc. Además, las sustituciones se pueden hacer en base a la tendencia a una estructura secundaria. Por ejemplo, se puede reemplazar un aminoácido helicoidal por un aminoácido que conserve la estructura helicoidal. Es posible seleccionar ejemplos de sustituciones que tengan en consideración varias de las características anteriores con el fin de producir cambios de aminoácidos conservadores que den como resultado cambios silenciosos dentro de los presentes péptidos, etc., de otros miembros de la clase a la que pertenece el aminoácido natural. Los aminoácidos se pueden dividir en los cuatro siguientes grupos: (1) aminoácidos ácidos; (2) aminoácidos básicos; (3) aminoácidos polares neutros; y (4) aminoácidos no polares neutros.

Procedimientos generales para elaborar las proteínas de fusión heterólogas de la presente invención

Aunque las proteínas de fusión heterólogas de la presente invención se pueden elaborar mediante una variedad de diferentes procedimientos, se prefieren los procedimientos recombinantes. A efectos de la presente invención, según lo revelado y reivindicado en la presente memoria, se definen a continuación los siguientes términos y abreviaturas generales de Biología Molecular. Los términos y las abreviaturas usados en este documento tienen sus significados normales a no ser que se designe lo contrario. Por ejemplo, "ºC" 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 se exponen en 37 C. F. R. § 1.822 (b) (2) (1994).

"Par de bases" o "pb", como se usa en la presente memoria, se refiere a ADN o ARN. Las abreviaturas A, C, G y T corresponden a las formas 5'-monofosfato de los desoxirribonucleósidos (desoxi)adenosina, (desoxi)citidina, (desoxi)guanosina y timidina, respectivamente, cuando tienen lugar en moléculas de ADN. Las abreviaturas U, C, G y A corresponden a las formas 5'-monofosfato de los ribonucleósidos uridina, citidina, guanosina y adenosina, respectivamente, cuando tienen lugar en moléculas de ARN. En el ADN bicatenario, el par de bases puede referirse a una pareja de A con T o de C con G. En un ADN/ARN, el par de bases del heterodúplex puede referirse a una pareja de A con U o de C con G. (Véase la definición de "complementario", *infra*).

"Digestión" o "restricción" de ADN se refiere a la escisión catalítica del ADN con una enzima de restricción que sólo actúa en ciertas secuencias del ADN ("endonucleasas específicas de secuencia"). Las diversas enzimas de restricción usadas en la presente memoria se encuentran comercialmente disponibles y sus condiciones de reacción, cofactores y otros requisitos se usaron tal y como es conocido por cualquier experto habitual en la técnica. Los tampones y las cantidades de sustrato apropiados para las enzimas de restricción concretas son especificados por el fabricante y pueden encontrarse fácilmente en la bibliografía.

"Ligadura" se refiere al procedimiento de formación de enlaces de tipo fosfodiéster entre dos fragmentos de ácido nucleico bicatenarios. A no ser que se proporcione de otro modo, la ligadura se puede realizar usando tampones y condiciones conocidos con una ADN ligasa, tal como la ADN ligasa de T4.

"Plásmido" se refiere a un elemento genético (habitualmente) auto-replicativo extracromosómico. Los plásmidos se designan generalmente por una "p" minúscula seguida por letras y/o números. Los plásmidos iniciales de la presente memoria bien se encuentran comercialmente disponibles, accesibles al público en una base no restringida o se pueden construir a partir de plásmidos disponibles conforme a los procedimientos publicados. Además, los plásmidos equivalentes a los descritos son conocidos en la técnica y le resultarán evidentes al experto habitual en la técnica.

"Vector de clonación de ADN recombinante", como se usa en la presente memoria, se refiere a cualquier agente replicante de manera autónoma, incluyendo, pero no limitándose a, plásmidos y fagos que comprenden una molécula de ADN a la que se puede añadir o se han añadido uno o más segmentos de ADN adicionales.

"Vector de expresión de ADN recombinante", como se usa en la presente memoria, se refiere a cualquier vector de clonación de ADN recombinante al que se haya incorporado un promotor para controlar la transcripción del ADN insertado.

"Transcripción" se refiere al procedimiento mediante el cual la información contenida en una secuencia de nucleótidos de ADN se transfiere a una secuencia de ARN complementaria.

"Transfección" se refiere a la absorción de un vector de expresión por una célula huésped independientemente de si se expresa alguna secuencia codificante o no. El experto habitual en la técnica conoce numerosos procedimientos de transfección, por ejemplo, la coprecipitación con fosfato de calcio, la transfección de liposomas y la electroporación. En general, se reconoce que una transfección se ha realizado satisfactoriamente cuando tiene lugar cualquier indicio del funcionamiento de este vector dentro de la célula huésped.

"Transformación" se refiere a la introducción de ADN en un organismo tal que el ADN es replicable, bien como un elemento extracromosómico o mediante la integración cromosómica. Los procedimientos de transformación de huéspedes bacterianos y eucariotas son conocidos en la técnica, resumiéndose muchos de ellos, tales como la inyección nuclear, la fusión protoplástica o el tratamiento con calcio usando cloruro de calcio, en J. Sambrook, *et al.,* "Molecular Cloning: A Laboratory Manual", (1989). En general, cuando se introduce ADN en levadura, el término transformación se usa en contraposición con el término transfección.

"Traducción", como se usa en la presente memoria, se refiere al procedimiento mediante el cual la información genética del ARN mensajero (ARNm) se usa para especificar y dirigir la síntesis de una cadena polipeptídica.

"Vector" se refiere a un compuesto de ácido nucleico usado para la transfección y/o la transformación de células en secuencias de polinucleótidos que portan manipulaciones genéticas correspondientes a moléculas proteínicas apropiadas que, cuando se combina con secuencias control apropiadas, confiere propiedades específicas a la célula huésped por ser transfectada y/o transformada. Los plásmidos, los virus y los bacteriófagos son vectores adecuados. Los vectores artificiales se construyen cortando y uniendo moléculas de ADN de diferentes fuentes usando enzimas de restricción y ligasas. El término "vector", como se usa en la presente memoria, incluye vectores de clonación de ADN recombinante y vectores de expresión de ADN recombinante.

"Complementario" o "complementariedad", como se usa en la presente memoria, se refiere a pares de bases (purinas y pirimidinas) que se asocian a través de enlaces de hidrógeno en un ácido nucleico bicatenario. Son complementarios los siguientes pares de bases: guanina y citosina; adenina y timina; y adenina y uracilo.

"Hibridación", como se usa en la presente memoria, se refiere a un procedimiento en el que una cadena de ácido nucleico se une con una cadena complementaria a través del apareamiento de bases. Las condiciones empleadas en la hibridación de dos ácidos nucleicos complementarios no idénticos, pero muy similares, varían con el grado de complementariedad de las dos cadenas y la longitud de las mismas. Tales técnicas y condiciones son conocidas por los profesionales de este campo.

"Secuencia de aminoácidos aislada" se refiere a cualquier secuencia de aminoácidos, bien construida o sintetizada, que tenga una ubicación diferente a su ubicación natural.

"Compuesto de ADN aislado" se refiere a cualquier secuencia de ADN, bien construida o sintetizada, que tenga una ubicación diferente a su ubicación natural en el ADN genómico.

"Compuesto de ácido nucleico aislado" se refiere a cualquier secuencia de ADN o ARN, bien construida o sintetizada, que tenga una ubicación diferente a su ubicación natural.

"Cebador" se refiere a un fragmento de ácido nucleico que funciona como un sustrato iniciador del alargamiento enzimático o sintético.

"Promotor" se refiere a una secuencia de ADN que dirige la transcripción del ADN a ARN.

"Sonda" se refiere a un compuesto de ácido nucleico o un fragmento del mismo que se hibrida con otro compuesto de ácido nucleico.

Las "condiciones restrictivas" de las reacciones de hibridación son fácilmente determinables por cualquier experto habitual en la técnica, y generalmente se trata de un cálculo empírico dependiente de la longitud de la sonda, de la temperatura de lavado y de la concentración salina. En general, las sondas más largas requieren temperaturas más elevadas para aparearse apropiadamente, mientras que las sondas cortas necesitan temperaturas más bajas. La hibridación depende generalmente de la capacidad del ADN desnaturalizado para volverse a aparear cuando hay presentes cadenas complementarias en un medio por debajo de su temperatura de fusión. Cuanto más elevado es el grado de homología deseado entre la sonda y la secuencia hibridable, más elevada será la temperatura relativa que se pueda usar. Como resultado, las temperaturas relativas más elevadas tenderían a hacer posibles reacciones con condiciones más restrictivas, mientras que las temperaturas más bajas, con condiciones menos restrictivas. Para más información sobre las condiciones restrictivas de las reacciones de hibridación, véase Ausubel *et al*., "Current Protocols in Molecular Biology", Wiley Interscience Publishers, 1995.

"Las condiciones restrictivas" o "las condiciones muy restrictivas", según lo definido en la presente memoria, se pueden identificar por aquéllas que (1) emplean una fuerza iónica baja y una temperatura elevada para lavar, por ejemplo, cloruro de sodio 15 mM/citrato de sodio 1,5 mM/dodecil sulfato de sodio al 0,1% a 50ºC; (2) emplean durante la hibridación un agente desnaturalizante, tal como formamida, por ejemplo, formamida al 50% (v/v) con albúmina de suero bovino al 0,1%/ficoll al 0,1%/polivinilpirrolidona al 0,10%/tampón de fosfato de sodio 50 mM a pH 6,5 con cloruro de sodio 750 mM/citrato de sodio 75 mM a 42ºC; o (3) emplean formamida al 50%, 5 x SSC (cloruro de sodio 750 mM, citrato de sodio 75 mM), fosfato de sodio 50 mM (pH 6,8), pirofosfato de sodio al 0,1%, 5 x solución de Denhardt, ADN de esperma de salmón sometido a tratamiento con ultrasonidos (50 μg/ml), SDS al 0,1% y sulfato de dextrano al 10% a 42ºC con lavados a 42ºC en 0,2 x SSC (cloruro de sodio 30 mM/citrato de sodio 3 mM) y formamida al 50% a 55ºC, siguiendo con un lavado en condiciones muy restrictivas que consiste en 0,1 x SSC que contiene EDTA a 55ºC.

Las "condiciones moderadamente restrictivas" se pueden identificar según lo descrito por Sambrook *et al*. ["Molecular Cloning: A Laboratory Manual", Nueva York: Cold Spring Harbor Press, (1989)], e incluyen el uso de solución de lavado y condiciones de hibridación (p.ej., temperatura, fuerza iónica y % de SDS) menos restrictivas que las descritas anteriormente. Un ejemplo de condiciones moderadamente restrictivas es la incubación durante una noche a 37ºC en una solución que comprende: formamida al 20%, 5 x SSC (cloruro de sodio 750 mM, citrato de sodio 75 mM), fosfato de sodio 50 mM a pH 7,6; 5 x solución de Denhardt, sulfato de dextrano al 10% y 20 mg/ml de ADN de esperma de salmón triturado desnaturalizado, seguido por un lavado de los filtros en 1 x SSC a aproximadamente 37-50ºC. El experto en la técnica sabrá cómo ajustar la temperatura, la fuerza iónica, etc., según sus necesidades en virtud de factores tales como la longitud de sonda y similares.

"PCR" se refiere a la tan conocida reacción en cadena de la polimerasa que emplea una ADN polimerasa térmicamente estable.

"Secuencia líder" se refiere a una secuencia de aminoácidos que puede ser eliminada enzimática o químicamente para producir el polipéptido de interés deseado.

"La secuencia de señales de secreción" se refiere a una secuencia de aminoácidos generalmente presente en la región N-terminal de un polipéptido más largo que funciona para iniciar la asociación de ese polipéptido con los compartimentos de la membrana celular como el retículo endoplasmático y la secreción de ese polipéptido a través de la membrana plasmática.

Construcción de ADN codificante de las proteínas de fusión heterólogas de la presente invención

Las proteínas inmunoglobulinas y albúminas de tipo natural se pueden obtener de una variedad de fuentes. Por ejemplo, estas proteínas se pueden obtener de una genoteca de ADNc preparada a partir de tejido o células que expresen el ARNm de interés a un nivel detectable. Las genotecas se pueden rastrear con sondas diseñadas usando el ADN o la secuencia de proteínas publicados para la proteína concreta de interés. En Adams, *et al.* (1980) *Biochemistry* 19: 2711-2719; Goughet, *et al.* (1980) *Biochemistry* 19: 2702-2710; Dolby, *et al*. (1980) *Proc. Natl. Acad. Sci.* EE.UU. 77: 6027-6031; Rice *et al.* (1982) *Proc. Natl. Acad. Sci.* EE.UU. 79: 7862-7862; Falkner, *et al.*(1982) *Nature* 298: 286-288; y Morrison, *et al.* (1984) *Ann. Rev. Immunol*. 2: 239-256, se describen, por ejemplo, las regiones constantes de cadena ligera o pesada de las inmunoglobulinas. Algunas referencias que revelan secuencias de proteínas albúminas y de ADN incluyen 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; y Minghetti, *et al.* (1986) *J. Biol. Chem*. 261: 6747.

El rastreo de una genoteca de ADNc o genómica con la sonda seleccionada se puede realizar usando procedimientos estándar, tales como los descritos en Sambrook *et al.,* "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press, NY (1989). Un procedimiento alternativo para aislar un gen codificante de una proteína inmunoglobulina o albúmina consiste en usar una metodología de PCR [Sambrook *et al., supra*; Dieffenbach *et al.,* "PCR Primer: A Laboratory Manual", Cold Spring Harbor Laboratory Press, NY (1995)]. Los cebadores PCR se pueden diseñar en base a las secuencias publicadas.

En general, las secuencias de tipo natural de longitud completa clonadas a partir de una determinada especie pueden servir como molde para crear análogos, fragmentos y derivados que conserven la capacidad de conferir una vida media en plasma más larga al compuesto de GLP-1 que forma parte de la proteína de fusión. Es preferible que las partes de albúmina de las proteínas de fusión heterólogas de la presente invención sean derivadas de la secuencia humana nativa con el fin de reducir el riesgo de una posible inmunogenicidad de la proteína de fusión en seres humanos.

Dependiendo de si se desea un determinado efecto o determinadas funciones, o de las características deseadas de la proteína de fusión, un fragmento Fc puede contener la región bisagra junto con los dominios CH_2 y CH_3 o alguna otra combinación de los mismos. Estos fragmentos Fc se pueden generar usando técnicas PCR con cebadores diseñados para hibridarse a secuencias correspondientes a los extremos deseados del fragmento. De manera similar, si se desean fragmentos de albúmina, se pueden diseñar cebadores PCR que sean complementarios a secuencias de albúminas internas. También se pueden diseñar cebadores PCR para crear sitios de enzimas de restricción para facilitar la clonación en vectores de expresión.

Se puede producir ADN codificante de los compuestos de GLP-1 de la presente invención mediante una variedad de diferentes procedimientos incluyendo procedimientos de clonación como aquéllos descritos anteriormente, así como ADN sintetizado químicamente. La síntesis química puede resultar atractiva dada la corta longitud del péptido codificado. La secuencia de aminoácidos para GLP-1 ha sido publicada, así como la secuencia del gen preproglucagón. [Lopez, *et al*. (1983) *Proc. Natl. Acad. Sci.*, EE.UU. 80: 5485-5489; Bell, *et al.* (1983) *Nature*, 302: 716-718; Heinrich, G., *et al*. (1984) *Endocrinol*, 115: 2176-2181; Ghiglione, M., *et al*. 91984) *Diabetologia* 27: 599-600]. De este modo, se pueden diseñar cebadores para aplicar la PCR en los compuestos de GLP-1 nativos y en los fragmentos de los mismos.

Entonces se puede construir el gen codificante de una proteína de fusión ligando ADN codificante de un compuesto de GLP-1 en marco con ADN codificante de una proteína albúmina. El gen codificante del compuesto de GLP-1 y el gen codificante de la proteína albúmina también pueden unirse en marco a través de ADN codificante de un péptido ligador.

Se puede optimizar la función y la estabilidad *in vivo* de las proteínas de fusión heterólogas de la presente invención añadiendo pequeños ligadores peptídicos para evitar las posibles interacciones no deseadas entre dominios. Aunque estos ligadores pueden ser de cualquier longitud y estar constituidos por cualquier combinación de aminoácidos, es preferible que la longitud no sea mayor de lo necesario para evitar las interacciones no deseadas entre dominios y/o optimizar la actividad biológica y/o la estabilidad. Generalmente, los ligadores no deberían contener aminoácidos con cadenas laterales extremadamente voluminosas ni aminoácidos propensos a introducir una estructura secundaria significativa. Es preferible que el ligador sea rico en serina-glicina y que tenga una longitud de menos de 30 aminoácidos. Es más preferible que el ligador tenga una longitud de no más de 20 aminoácidos. Es incluso más preferible que el ligador tenga una longitud de no más de 15 aminoácidos. Un ligador preferido contiene repeticiones de la secuencia Gly-Gly-Gly-Gly-Ser. Es preferible que haya entre 2 y 6 repeticiones de esta secuencia. Es incluso más preferible que haya entre 3 y 4 repeticiones de esta secuencia.

El ADN codificante del GLP-1 de tipo natural, los polipéptidos de albúmina y los fragmentos de los mismos puede ser mutado bien antes de la ligadura o en el contexto de un ADNc codificante de una proteína de fusión entera. Hay una variedad de técnicas de mutagénesis conocidas en la materia. Por ejemplo, el procedimiento de PCR mutagénica utiliza la extensión por solapamiento de cadenas para crear mutaciones de bases específicas a efectos de cambiar la secuencia de aminoácidos específica de la proteína correspondiente. Esta mutagénesis de PCR requiere el uso de cuatro cebadores, dos en sentido directo (cebadores A y C) y dos en sentido inverso (cebadores B y D). Los genes mutados se amplifican desde el 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 son dirigidos a la zona del gen por ser mutado. Cuando se alinean estos cebadores con la zona diana, contienen apareamientos erróneos para las bases dirigidas a ser cambiadas. Una vez completadas las reacciones de A a B y de C a D, se aíslan los productos de reacción y se mezclan para su uso como el molde para la reacción de A a D. Entonces, la reacción genera el producto mutado completo.

Una vez que se produce un gen codificante de una proteína de fusión entera, se puede clonar en un vector de expresión apropiado. En el ejemplo 1, se describen las estrategias específicas que se pueden emplear para formar las proteínas de fusión GLP-1 de la presente invención.

Procedimientos generales para expresar recombinantemente las proteínas de fusión heterólogas de la presente invención

Se transfectan o transforman células huésped con los vectores de expresión o clonación descritos en la presente memoria para la producción de proteínas de fusión heterólogas, y se cultivan en medios de nutrientes convencionales modificados según proceda para inducir promotores, seleccionar transformantes o amplificar los genes codificantes de las secuencias deseadas. Las condiciones de cultivo, tales como los medios, la temperatura, el pH y similares, pueden ser seleccionadas por el experto en la técnica sin la necesidad de una experimentación. En general, los principios, los protocolos y las técnicas prácticas para maximizar la productividad de los cultivos celulares se pueden encontrar en "Mammalian Cell Biotechnology: A Practical Approach", M. Butler, ed. (IRL Press, 1991) y en Sambrook, *et al., supra*. El experto habitual en la técnica conoce procedimientos de transfección, por ejemplo, $CaPO_4$ y la electroporación. En la patente estadounidense n.º: 4.399.216, se han descrito los aspectos generales de las transformaciones de los sistemas huésped de células de mamífero. Las transformaciones en levadura se llevan a cabo comúnmente según el procedimiento de van Solingen *et al., J Bact.* 130 (2): 946-7 (1977) y Hsiao *et al., Proc. Natl. Acad. Sci*. EE.UU. 76 (8): 3829-33 (1979). Sin embargo, también se pueden usar otros procedimientos para introducir ADN en células, tales como mediante microinyección nuclear, electroporación, fusión de protoplastos bacterianos con células intactas o policationes, p.ej., polibreno o poliomitina. Para diversas técnicas de transformación de células de mamífero, véase Keown, *et al.*, "Methods in Enzymology" 185: 527-37 (1990) y Mansour, *et al., Nature* 336 (6197): 348-52 (1988).

Las células huésped adecuadas para clonar o expresar el ácido nucleico (p.ej., ADN) en los vectores de la presente memoria incluyen células procariotas, de levadura o eucariotas superiores. Los procariotas adecuados incluyen, pero no se limitan a, eubacterias, tales como organismos Gram-negativos o Gram-positivos, por ejemplo, enterobacterias tales como *E. coli*. Hay diversas cepas de *E. coli* accesibles al público, tales como *E. coli* K12 cepa MM294 (ATCC 3 1.446); *E. coli* XI 776 (ATCC 3 1.537); *E. coli.* cepa W3 110 (ATCC 27.325) y K5 772 (ATCC 53.635). Otras células

huésped procariotas adecuadas incluyen géneros de la familia *Enterobactericeae* tales como *Escherichia*, p.ej., *E. coli, Enterobacter*, *Erwinia*, *Klebisella*, *Proteus*, *Salmonella*, p.ej., *Salmonella typhimurium*, *Serratia*, p.ej., *Serratia marcescans* y *Shigeila*, así como *Bacilli* tales como *B. subtilis* y *B. lichentformis* (p.ej., *B. licheniformis* 4 1 P revelado en el documento DD266.710, *publicado el 12 de abril de 1989), Pseudomonas* tales como *P. aeruginosa* y *Streptomyces*. Estos ejemplos son ilustrativos y no restrictivos. La cepa W3110 es un huésped o huésped parental particularmente preferido, porque es una cepa huésped común para las fermentaciones de productos de ADN recombinante. Preferiblemente, la célula huésped secreta cantidades mínimas de enzimas proteolíticas. Por ejemplo, la cepa W3110 se puede modificar para que realice una mutación genética en los genes codificantes de proteínas endógenas para el huésped, incluyendo como ejemplos de tales huéspedes *E.coli* W3110 cepa 1A2, que tiene el genotipo completo ronA; *E. coli* W3110 cepa 9E4, que tiene el genotipo completo ton4 ptr3; *E.coli* W3110 cepa 27C7 (ATCC 55.244), que tiene el genotipo completo tonA, ptr3 phoA E15 (argF-lac) 169 degP ompT/can'; *E. coli* W3110 cepa 40B4, que es la cepa 37D6 con una mutación por eliminación de degP no resistente a la canamicina; y una cepa de *E. coli* que tiene la proteasa periplásmica mutante revelada en la patente estadounidense n.º: 4.946.783 concedida el 7 de agosto de 1990. Alternativamente, son adecuados procedimientos de clonación *in vivo*, p.ej., PCR u otras reacciones de la polimerasa de ácidos nucleicos.

Además de los procariotas, los microbios eucariotas tales como los hongos filamentosos o las levaduras son huéspedes de clonación o expresión adecuados para los vectores de proteínas de fusión. *Saccharomyces cerevisiae* es un microorganismo huésped eucariota inferior comúnmente usado. Otros 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 estadounidense n.º: 4.943.529; Fleer, *et al., BiolTechnology* 9 (10): 968-75 (1991)] tales como, p.ej., *K lactis* (MW98-8C, CBS683, CBS4574) [de Louvencourt *et al.*, *J. Bacteriol.* 154 (2): 737-42 (1983)]; *K. fiagilis* (ATCC 12.424), *K. bulgaricus* (ATCC 16.045), *K wickeramii* (ATCC 24.178), *K waltii* (ATCC 56.500), *K. drosophilarum* (ATCC 36.906) [Van den Berg *et al., BiolTechnology* 8 (2): 135-9 (1990)]; *K. thermotoierans* y *K. marxianus*; *yarrowia* (EP 402,226); *Pichia pastoris* (EP 183.070) [Sreekrishna *et al., J. Basic Microbiol.* 28 (4): 265-78 (1988)]; *Candid*; *Trichoderma reesia* (EP 244.234); *Neurospora crassa* [Case, *et al., Proc. Natl. Acad Sci.* EE.UU. 76 (10): 5259-63 (1979)]; *Schwanniomyces* tales como *Schwanniomyces occidentulis* (documento EP 394.538 publicado el 31 de octubre de 1990); y hongos filamentosos tal como, p.ej., *Neurospora*, *Penicillium*, *Tolypocladium* (documento WO 91/00357 publicado el 10 de enero de 1991) y huéspedes *Aspergillus* tales como *A. nidulans* [Ballance *et al., Biochem. Biophys. Res. Comm.* 112 (1): 284-9 (1983)]; Tilburn, *et al., Gene* 26 (2-3): 205-21 (1983); Yelton, *et al., Proc. Natl. Acad. Sci*. EE.UU. 81 (5): 1470-4 (1984)] y *A. niger* [Kelly e Hynes, *EMBO J.* 4 (2): 475-9 (1985)]. Las levaduras metilotrópicas se seleccionan del género constituido por *Hansenula, Candida, Kloeckera, Pichia, Saccharomyces, Torulopsia* y *Rhodotoruia*. En C. Antony, "The Biochemistry of Methylotrophs" 269 (1982), se puede encontrar una lista de especies específicas que son ejemplos de esta clase de levadura.

Las células huésped adecuadas para la expresión de las proteínas de fusión de la presente invención están derivadas de organismos multicelulares. Los ejemplos de células de invertebrado incluyen células de insecto tales como *Drosophila* S2 y *Spodoptera* Sp., *Spodoptera* high5, así como células vegetales. Los ejemplos de líneas de células huésped de mamífero útiles incluyen células de ovario de hámster chino (CHO) y células COS. Los ejemplos más específicos incluyen la línea CV1 de riñón de mono transformada por SV40 (COS-7, ATCC CRL 1651); línea de riñón embrionario humano [células 293 ó 293 subclonadas para el crecimiento en cultivo de suspensiones, Graham, *et al., J. Gen Tirol.,* 36 (1): 59-74 (1977)]; células de ovario de hámster chino/-DHFR [CHO, Urlaub y Chasin, *Proc. Natl. Acad. Sci*. EE.UU., 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 mamario de ratón (MMT 060562, ATCC CCL51). Se considera que la selección de la célula huésped apropiada es competencia del experto en la técnica.

Las proteínas de fusión de la presente invención se pueden producir recombinantemente directamente o como una proteína que tenga una secuencia señal u otras secuencias adicionales que creen un sitio de escisión específico en el terminal N de la proteína de fusión madura. En general, la secuencia señal puede ser un componente del vector o puede ser una parte del ADN codificante de la proteína de fusión que esté insertado en el vector. La secuencia señal puede ser una secuencia señal procariota seleccionada, por ejemplo, del grupo de líderes de la alcalino-fosfatasa, penicilina, lpp o enterotoxina II estable al calor. Para la secreción de levaduras la secuencia señal puede ser, p.ej., la invertasa de levadura líder, el factor alfa líder (incluyendo los factores cc líderes de *Saccharomyces* y *Kluyveromyces*, estando el último descrito en la patente estadounidense n.º: 5.010.182) o la fosfatasa ácida líder, la glucoamilasa de *C. albicans* líder (EP 362.179) o la señal descrita en el documento WO 90/13646. En la expresión de células de mamífero, se pueden usar secuencias señal de mamífero para dirigir la secreción de la proteína, tales como secuencias señal de polipéptidos secretados de la misma especie o de una especie relacionada, así como secretores víricos líderes.

Tanto los vectores de expresión como los de clonación contienen una secuencia de ácidos nucleicos que permite la replicación del vector en una o más células huésped seleccionadas. Tales secuencias son conocidas para una variedad de bacterias, levaduras y virus. El origen de replicación del plásmido pBR322 es adecuado para la mayoría de las bacterias Gram-negativas, el origen del plásmido 2u es adecuado para la levadura y diversos orígenes víricos (SV40, polioma, adenovirus, VSV o BPV) son útiles para los vectores de clonación en células de mamífero.

Los vectores de expresión y de clonación contendrán comúnmente un gen de selección, también denominado marcador seleccionable. Los genes de selección más comunes codifican proteínas que (a) confieren resistencia a antibióticos o a otras toxinas, p.ej., ampicilina, neomicina, metotrexato o tetraciclina, (b) deficiencias autotróficas de

complemento o (c) suministran nutrientes fundamentales no disponibles en los medios complejos, p. ej., el gen codificante de D-alanina racemasa para *Bacilli*.

Un ejemplo de marcadores seleccionables adecuados para las células de mamífero son aquéllos que permiten la identificación de células competentes para la absorción de ácido nucleico codificante de proteínas de fusión, tal como DHFR o timidina quinasa. Una célula huésped apropiada cuando se emplea DHFR de tipo natural es la línea celular CHO deficiente en actividad de DHFR, preparada y propagada según lo descrito [Urlaub y Chasin, *Proc. Natl. Acad. Sci.* EE.UU., 77 (7): 4216-20 (1980)]. Un gen de selección adecuado para su uso en levadura es el gen trpl presente en el plásmido de levadura YRp7 [Stinchcomb, *et al., Nature* 282 (5734): 39-43 (1979); Kingsman, *et al., Gene* 7(2): 141-52 (1979); Tschumper, *et al., Gene* 10(2): 157-66 (1980)]. El gen trpl proporciona un marcador de selección para una cepa mutante de levadura carente de la capacidad de crecer en triptófano, por ejemplo, ATCC N.º 44076 o PEPC1 [Jones, Genetics 85: 23-33 (1977)].

Los vectores de expresión y de clonación habitualmente contienen un promotor ligado operativamente a la secuencia de ácido nucleico codificante de proteína de fusión para dirigir la síntesis de ARNm. Los promotores reconocidos por una variedad de posibles células huésped son conocidos. Los promotores adecuados para su uso con huéspedes procariotas incluyen los sistemas promotores de β-lactamasa y lactosa [Chang, *et al., Nature* 275 (5681): 617-24 (1978); Goeddel, *et al., Nature* 281 (5732): 544-8 (1979)], alcalino fosfatasa, un sistema promotor de triptófano (up) [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, *et al., Proc. Natl. Acad. Sci*. EE.UU. 80 (1): 21-5 (1983)]. Los promotores para su uso en sistemas bacterianos también contendrán una secuencia de Shine-Dalgarno (S.D.) ligada operativamente al ADN codificante de la proteína de fusión.

Los ejemplos de secuencias promotoras adecuadas para su uso con huéspedes de levadura incluyen los promotores para la 3-fosfoglicerato quinasa [Hitzeman, *et al., J. Biol. Chem*. 255 (24): 12073-80 (1980)] u otras enzimas glicolíticas [Hess *et al., J. Adv. Enzyme Reg.* 7: 149 (1968); Holland, *Biochemistry* 17 (23): 4900-7 (1978)], tales como enolasa, gliceraldehído-3-fosfato deshidrogenasa, hexoquinasa, piruvato descarboxilasa, fosfofructoquinasa, glucosa-6-fosfato isomerasa, 3-fosfoglicerato mutasa, piruvato quinasa, triosafosfato isomerasa, fosfoglucosa isomerasa y glucoquinasa.

Otros promotores de levadura que son promotores inducibles que tienen la ventaja adicional de la transcripción controlada por condiciones de crecimiento son las regiones promotoras para alcohol deshidrogenasa 2, isocitocromo C y fosfatasa ácida, enzimas degradantes asociadas con el metabolismo del nitrógeno, metalotioneína, gliceraldehído-3-fosfato deshidrogenasa y enzimas responsables de la utilización de la maltosa y la galactosa. En el documento EP 73.657, se describen más detalladamente los vectores y los promotores adecuados para su uso en la expresión de levaduras. La transcripción de ARNm codificante de la proteína de fusión procedente de vectores de células huésped de mamífero se puede controlar, por ejemplo, mediante promotores obtenidos de los genomas de virus tales como virus del polioma, virus de la viruela aviar, adenovirus (tal como el adenovirus 2), virus del papiloma bovino, virus del sarcoma aviar, citomegalovirus, un retrovirus, virus de la hepatitis B y virus 40 del simio (SV40), de promotores de mamífero heterólogos, p.ej., el promotor de la actina o un promotor de inmunoglobulinas, y de promotores de choque térmico, siempre y cuando tales promotores sean compatibles con los sistemas de células huésped.

Se puede aumentar la transcripción de un polinucleótido codificante de una proteína de fusión mediante eucariotas superiores insertando una secuencia potenciadora en el vector. Los potenciadores son elementos de ADN que actúan en *cis*, habitualmente de aproximadamente 10 a 300 pb, que actúan sobre un promotor para aumentar su transcripción. Ahora se conocen muchas secuencias potenciadoras de genes de mamífero (globina, elastasa, albúmina, α-ketoproteína e insulina). Sin embargo, lo común es usar un potenciador de un virus de célula eucariota. Los ejemplos incluyen el potenciador SV40 en el último lado del origen de replicación (100-270 pb), el potenciador del promotor temprano de citomegalovirus, el potenciador del polioma sobre el último lado del origen de replicación y potenciadores de adenovirus. Se puede cortar el potenciador y empalmarlo en el vector en una posición 5' o 3' con respecto a la secuencia codificante de la proteína de fusión, pero se localiza preferiblemente en un sitio 5' del promotor.

Los vectores de expresión usados en células huésped eucariotas (células de levadura, hongos, insecto, planta, animal, ser humano o células nucleadas de otros organismos multicelulares) también contendrán las secuencias necesarias para la terminación de la transcripción y para la estabilización del ARNm. Tales secuencias se encuentran comúnmente disponibles en las regiones no traducidas 5' y ocasionalmente 3' de ADN o ADNc eucariotas o víricos. Estas regiones contienen segmentos de nucleótidos transcritos como fragmentos poliadenilados en la parte no traducida del ARNm codificante de la proteína de fusión.

Se pueden recuperar diversas formas de una proteína de fusión de un medio de cultivo o de lisados de célula huésped. Si está unido a la membrana, puede ser liberado de la membrana usando una solución detergente adecuada (p.ej., Triton-X 100) o mediante escisión enzimática. Las células empleadas en la expresión de una proteína de fusión pueden ser alteradas mediante diversos procedimientos físicos o químicos, tales como ciclos de congelación-descongelación, tratamiento de ultrasonidos, alteración mecánica o con agentes de lisado celular.

Purificación de las proteínas de fusión heterólogas de la presente invención

Una vez expresadas las proteínas de fusión heterólogas de la presente invención en la célula huésped apropiada, se pueden aislar y purificar los análogos. Los siguientes procedimientos son ejemplos de procedimientos de purificación

adecuados: fraccionamiento sobre carboximetilcelulosa; filtración sobre gel tal como Sephadex G-75; resina de intercambio aniónico tal como DEAE o Mono-Q; intercambio catiónico tal como CM- o Mono-S; proteína A sefarosa para eliminar contaminantes tales como IgG; columnas quelantes de metales para unir formas del polipéptido marcadas con epítopos; CLAR en fase inversa; cromatoenfocamiento; gel de sílice; precipitación en etanol; y precipitación en sulfato de amonio.

Se pueden emplear diversos procedimientos de purificación de proteínas, y tales procedimientos son conocidos en la técnica y se encuentran descritos, 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 usado y de la proteína de fusión producida en particular. Por ejemplo, se pueden purificar eficazmente proteínas de fusión que comprendan un fragmento Fc usando una matriz de afinidad con proteína A o proteína G. Se pueden usar tampones de pH bajo o alto pare eluir la proteína de fusión de la matriz de afinidad. Las condiciones de elución suaves ayudarán en la prevención de la desnaturalización irreversible de la proteína de fusión. También se pueden usar tampones que contengan imidazol. El ejemplo 3 describe algunos protocolos de purificación satisfactorios 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 presente invención. Algunos de estos procedimientos incluyen: la electroforesis en gel de poliacrilamida-SDS junto con procedimientos de tinción de proteínas o inmunotransferencia usando anticuerpos anti-IgG o anti-ASH. Otros procedimientos incluyen la espectrometría de masas de desorción/ionización en matriz inducida por láser (MALDI-SM), cromatografía en fase líquida/espectrometría de masas, isoelectroenfoque, intercambio aniónico analítico, cromatoenfoque y dicroísmo circular por nombrar unos cuantos. Se caracterizó un número representativo de proteínas de fusión heterólogas usando electroforesis en gel de poliacrilamida-SDS junto con una inmunotransferencia, así como espectrometría de masas (véanse los ejemplos 4 y 5 y las figuras 3 y 4).

Por ejemplo, la tabla 3 (véase el ejemplo 5) ilustra la masa molecular calculada para un número representativo de proteínas de fusión, así como la masa determinada mediante espectrometría de masas. Además, las figuras 3 y 4 ilustran los pesos moleculares de un número representativo de proteínas de fusión determinado mediante electroforesis en gel de poliacrilamida-SDS. Todas las proteínas de fusión heterólogas analizadas fueron expresadas y secretadas transitoriamente. Además, se escindió la secuencia señal de Igκ para producir proteínas con el terminal N correcto.

Además, la tabla 3 ilustra que, en algunos casos, la masa determinada por espectrometría de masas es mayor de lo esperado. Esto es el resultado de la glicosilación de la parte Fc y de la extensión del terminal C. La digestión enzimática de las proteínas de fusión seguida por una CLAR en fase inversa y una espectrometría de masas pueden identificar las fracciones peptídicas que contienen restos de azúcar. Estas fracciones luego pueden ser secuenciadas con aminoácidos N-terminales para identificar el posible sitio de glicosilación. Por ejemplo, la caracterización de exendina-4-Fc (SEC ID N.º 29) muestra que la serina de la posición 39 y la treonina de la posición 50 están glicosiladas enlazadas a O, y la asparagina de la posición 122 está glicosilada enlazada a N.

También se analizó un número representativo de proteínas de fusión GLP-1 en cuanto a su actividad. Existen numerosos procedimientos para detectar la actividad de GLP-1 *in vitro* e *in vivo* (véanse los ejemplos 6, 7, 8 y 9). La tabla 4 (ejemplo 6) ilustra la actividad del receptor de GLP-1 asociado con varias fusiones de GLP-1. Los números se refieren a la actividad asociada con Val^8-GLP-1(7-37)OH. Todas las proteínas de fusión analizadas 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 sustancial aumento de la vida media de estas proteínas de fusión, una actividad *in vitro* débil generalmente no sirve para predecir una actividad *in vivo* débil. La figura 7 y el ejemplo 7 ilustran la vida media prolongada asociada con las proteínas de fusión de la presente invención. Por ejemplo, Val^8-GLP-1-Fc tenía una vida media de aproximadamente 45 horas en monos, Val^8-GLP-1-ASH tenía una vida media de aproximadamente 87 horas en monos, Gly^8-Glu^{22}-GLP-1-CEx-ligador-IgG1 tenía una vida media tras la administración i.v. de aproximadamente 55 horas en perros y Gly^8-Glu^{22}-GLP-1-CEx-ligador-IgG1 tenía una vida media tras la administración s.c. de aproximadamente 38 horas en perros.

Composiciones de la invención

La estabilidad física también es una característica esencial de las formulaciones de proteínas terapéuticas. Los compuestos GLP-1 han sido especialmente difíciles de fabricar y formular debido a los cambios estructurales que tienen lugar durante su procesamiento. Por ejemplo, algunos compuestos de GLP-1 tienen tienden en general a agregarse. Además, se ha observado que algunos compuestos de GLP-1 pasan de una forma soluble y de hélice α activa a una forma insoluble y de lámina β potencialmente inactiva. La fusión de compuestos de GLP-1 con proteínas de gran tamaño tales como la albúmina no sólo actúa para aumentar la vida media del compuesto de GLP-1, sino que además contribuye a la estabilidad física y configuracional del compuesto de GLP-1. Por ejemplo, Val^8-GLP-1-Ligador-ASH en PBS es estable a 37ºC fuera hasta aproximadamente 30 días.

Las proteínas de fusión heterólogas de la presente invención se pueden formular con uno o más excipientes. Es posible combinar las proteínas de fusión activas de la presente invención con un tampón farmacéuticamente aceptable, y ajustar el pH para proporcionar una estabilidad aceptable y un pH aceptable para la administración tal como una administración parenteral.

Opcionalmente, se pueden añadir uno o más agente antimicrobianos farmacéuticamente aceptables. El meta-cresol y el fenol son agentes microbianos farmacéuticamente aceptables preferidos. Se pueden añadir una o más sales farmacéuticamente aceptables para ajustar la fuerza iónica o la tonicidad. Se pueden añadir uno o más excipientes para ajustar más la isotonicidad de la formulación. La glicerina es un ejemplo de excipiente ajustador de la isotonicidad. "farmacéuticamente aceptable" significa adecuado para la administración a un ser humano o a otro animal y, por tanto, que no contiene elementos tóxicos o contaminantes no deseados ni interfiere en la actividad de los compuestos activos presentes.

Se puede usar una forma de una sal farmacéuticamente aceptable de las proteínas de fusión heterólogas de la presente invención en la presente invención. Los ácidos comúnmente empleados para formar las sales de adición ácida son ácidos inorgánicos tales como ácido clorhídrico, ácido bromhídrico, ácido yodhídrico, ácido sulfúrico, ácido fosfórico, y similares, y los ácidos orgánicos tales como ácido *p*-toluenosulfónico, ácido metanosulfónico, ácido oxálico, ácido *p*-bromofenil-sulfónico, ácido carbónico, ácido succínico, ácido cítrico, ácido benzoico, ácido acético y similares. Las sales de adición ácida preferidas son aquéllas formadas con ácidos minerales tales como ácido clorhídrico y ácido bromhídrico.

Las sales de adición básica incluyen aquéllas derivadas de bases inorgánicas, tales como hidróxidos de metales alcalinos o alcalinotérreos o de amonio, carbonatos, bicarbonatos y similares. Tales bases útiles en la preparación de sales de esta invención incluyen por tanto hidróxido de sodio, hidróxido de potasio, hidróxido de amonio, carbonato de potasio y similares.

Administración de las composiciones

La administración se puede realizar por cualquier vía conocida que le resulte eficaz al profesional habitual en la técnica. La administración periférica parenteral es uno de tales procedimientos. La administración parenteral se encuentra ampliamente descrita en la bibliografía médica 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 también se pueden adaptar a la administración mediante 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 variedad de enfermedades y afecciones. Las proteínas de fusión de la presente invención ejercen principalmente sus efectos biológicos actuando en un receptor denominado "receptor de GLP-1". Los sujetos con enfermedades y/o condiciones que responden favorablemente a la estimulación con receptor de GLP-1 o a la administración de compuestos de GLP-1 pueden ser, por tanto, tratados con las proteínas de fusión GLP-1 de la presente invención. Se dice que estos sujetos "están en necesidad del tratamiento con compuestos de GLP-1" o "en necesidad de la estimulación con receptores de GLP-1". Se incluyen sujetos con diabetes no dependiente de la insulina, diabetes dependiente de la insulina, derrame cerebral (véase el documento WO 00/16797), infarto de miocardio (véase el documento WO 98/08531), obesidad (véase el documento WO 98/19698), cambios catabólicos tras una cirugía (véase la patente estadounidense n.º: 6.006.753), dispepsia funcional y síndrome del intestino irritable (véase el documento WO 99/64060). También se incluyen los sujetos que requieren un tratamiento profiláctico con un compuesto de GLP-1, p.ej., sujetos con riesgo de desarrollar diabetes no dependiente de la insulina (véase el documento WO 00/07617). Los sujetos con problemas de tolerancia a la glucosa o de disminución de la glucosa en ayunas, los sujetos cuyo peso corporal es de aproximadamente un 25% superior al peso corporal normal correspondiente a la altura y la constitución corporal del sujeto, los sujetos con una pancreatectomía parcial, los sujetos que tienen uno o más padres con diabetes no dependiente de la insulina, los sujetos que han tenido diabetes gestacional y los sujetos que han padecido pancreatitis aguda o crónica están en riesgo de desarrollar diabetes no dependiente de la insulina.

Una "cantidad eficaz" de un compuesto de 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 una estimulación con receptores de GLP-1. Un "efecto terapéutico deseado" incluye uno o más de los siguientes efectos: 1) una mejoría del/de los síntoma(s) asociados con la enfermedad o la condición; 2) un retraso de la aparición de los síntomas asociados con la enfermedad o la condición; 3) un aumento de la longevidad en comparación con la ausencia del tratamiento; y 4) una mayor calidad de vida en comparación con la ausencia del tratamiento. Por ejemplo, una "cantidad eficaz" de un compuesto de GLP-1 para el tratamiento de la diabetes es la cantidad que daría como resultado un mayor control de la concentración de glucosa en sangre del que habría en ausencia del tratamiento, dando así como resultado un retraso de la aparición de complicaciones diabéticas tales como retinopatía, neuropatía o enfermedad de riñón. Una "cantidad eficaz" de un compuesto de GLP-1 para la prevención de la diabetes es la cantidad que retardaría, en comparación con la ausencia del tratamiento, la aparición de niveles elevados de glucosa en sangre que requieren el tratamiento con fármacos anti-hipoglicémicos tales como sulfonilureas, tiazolidinodionas, insulina y/o bisguanidinas.

La dosis de proteína de fusión eficaz para normalizar la glucosa en sangre de un paciente dependerá de un número de factores, entre los que se incluyen, sin limitación, el sexo, el peso, la edad del sujeto, la gravedad de la imposibilidad de regular la glucosa en sangre, la vía de administración y la biodisponibilidad, el perfil farmacocinético de la proteína de fusión, la potencia y la formulación.

La presente invención comprende compuestos de GLP-1 que tienen mejores propiedades bioquímicas y biofísicas en virtud de ser fusionados a una proteína albúmina, un fragmento de albúmina, un análogo de albúmina, una proteína Fc, un fragmento Fc o un análogo de Fc. Estas proteínas heterólogas se pueden expresar satisfactoriamente en células huésped, conservar las actividades de señalización asociadas con la activación del receptor de GLP-1 y tener vidas medias prolongadas.

Los siguientes ejemplos se presentan para describir en mayor profundidad la presente invención. No se pretende que el ámbito de la presente invención esté únicamente constituido por los siguientes ejemplos. Los expertos en la técnica reconocerán que los reactivos concretos, el equipo y los procedimientos descritos son meramente ilustrativos y no pretenden limitar la presente invención de ningún modo.

Ejemplo 1

Construcción de ADN codificante de proteínas de fusión heterólogas

Ejemplo comparativo 1a

Construcción de ADN codificante de de Val^8-GLP-1(7-37)-Fc

Se aisló una porción Fc de IgG1 humana de una genoteca de ADNc y que contenía la región bisagra completa y los dominios CH_2 y CH_3. Se subclonó un fragmento que contenía 696 pares de bases de esta porción Fc de la IgG1 humana en los sitios *Nhe*I y *Eco47III* del vector de expresión mamífero pJB02 para crear pJB02/Fc (véase la figura 5). El ADN codificante de la secuencia de señales de secreción de lgκ fusionada a Val^8-GLP-1(7-37) fue generado mediante hibridación *in vitro* de cuatro oligonucleótidos solapantes y complementarios:

5'- CTAGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
CCAGGTTCCACTGGTGACCAGTG - 3' [SEC ID N.º 12]

5'- GAGGGCACCTTCACCTCCGACGTGTCCTCCTATCTGGAGGGCCAGGCCGCCAAGGA
GTTCATCGCCTGGCTGGTGAAGGGAAGAGGC - 3' [SEC ID N.º 13]

5'- TGAAGGTGCCCTCCACGTGGTCACCAGTGGAACCTGGAACCCAGAGCAGCAGTA
CCCATAGCAGGAGTGTGTCTGTCTCCATGGTGG - 3' (SEC ID N.º 14)

5'- GCCTCTTCCCTTCACCAGCCAGGCGATGAACTCCTTGGCGGCCTGGCCCTCCAGA
TAGGAGGACACGTCGGAGG - 3' [SEC ID N.º 15]

La reacción de hibridación se llevó a cabo usando cantidades equivalentes de cada oligonucleótido (1 pm/μl de concentración final para cada oligo). Se calentó la mezcla de oligonucleótidos durante 5 min a 100ºC en tampón de unión (Tris-HCl 50 mM; pH 7,5; $MgCl_2$ 10 mM; DTT 10 mM; ATP 1 mM; 25 μg/ml de albúmina de suero bovino) y luego se enfrió durante al menos 2 horas a 30ºC.

Se unió el producto de hibridación resultante durante 2 horas a temperatura ambiente o durante una noche a 16ºC a la estructura del vector pJB02/Fc que había sido digerido con *NheI* y *Eco47III*. Los productos de unión se usaron para transformar células azules XL-1 competentes (Stratagene). Se rastrearon los plásmidos recombinantes para detectar la presencia de insertos codificantes de péptidos mediante la digestión de clones con *NcoI* (codificando la secuencia Kozak y la primera Met del péptido señal) y se secuenciaron. El plásmido de expresión resultante usado para los ensayos de transfección fue denominado pJB02-V8-GLP-1-Fc (Figura 5).

Ejemplo 1b

Construcción de ADN codificante de Val^8-GLP-1(7-37)-ASH

El plásmido ASH/pcDNA3.1GS fue adquirido en Invitrogen (n.º de catálogo H-M12523M-pcDNA3.1/GS) y se usó como un molde para aislar el ADNc codificante de la albúmina de suero humano (ASH). El ADNc de ASH fue preparado usando una PCR en la que se retiró el ADN codificante de la secuencia líder, así como el pro-péptido de seis aminoácidos del extremo 5'. Además, se añadieron directamente codones de terminación en el extremo 3' de la secuencia codificante de la ASH. Finalmente, se diseñaron genéticamente sitios de enzimas de restricción en el

extremo 5' y el extremo 3' para facilitar la clonación. La secuencia de ADN de la ASH presente en el vector original adquirido en Invitrogen contenía un cambio en una sola base en la región 3' del gen (posición 667) en comparación con la secuencia humana nativa. Este cambio resultaría en un codón para Asn en lugar de para Asp. De este modo, usando el procedimiento de mutagénesis dirigida por PCR de solapamiento de cadenas tratado anteriormente, se cambió el codón para que codificara Asp en esta posición. Se clonó el ADN codificante de la ASH resultante en los sitios *NheI* y *HindIII* de pJB02 para crear pJB02-ASH (Figura 6).

La secuencia líder de Igκ fusionada con la secuencia Val^8-GLP-1(7-37) fue generada según lo tratado en el ejemplo 1a. Se unió el ADN en los sitios *NheI* y *FspI* de pJB02-ASH para crear pJB02-Val^8-GLP-1-ASH.

Ejemplo 1c

Construcción de ADN codificante de Val^8-GLP-1-(7-37)-ligador-ASH

Se preparó el vector pJB02-ASH según lo descrito en el ejemplo 1b. Se unió el ADN codificante de la secuencia ligadora $[GGGGS]_3$ en marco con el extremo 5' del ADN codificante de la ASH para crear pJB02-ligador-ASH (Figura 7). El ADN codificante de la secuencia líder de Igκ fusionada con la secuencia de Val^8-GLP-1(7-37) y la parte 5' de la secuencia ligadora fue generado según lo tratado en el ejemplo 1a. Este ADN fue unido en los sitos *NheI* y *BspEI* de pJB02 para crear pJB02-Val^8-GLP-1-ligador-ASH.

Ejemplo comparativo 1d

Construcción de ADN codificante de exendina-4-Fc

Se preparó el plásmido pJB02/Fc según lo descrito en el ejemplo 1a. El ADN codificante de la secuencia señal de Igκ fusionada con la exendina-4 fue generado mediante una hibridación *in vitro* de los siguientes oligonucleótidos solapantes y complementarios:

5' - CTAGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTG
GGTTCCAGGTTCCACCGGTCAC - 3' [SEC ID N.º 16]

5' - GGAGAGGGAACCTTCACCAGCGACCTGAGCAAGCAGATGGAGGAGGAGGCCGT
GAGACTG - 3' [SEC ID N.º 17]

5' - TTCATCGAGTGGCTGAAGAACGGAGGACCAAGCAGCGGAGCCCCTCCTCCT
AGC - 3' [SEC ID N.º 18]

5' - GAACCTGGAACCCAGAGCAGCAGTACCCATAGCAGGAGTGTGTCTGTCTCCA
TGGTGG - 3' [SEC ID N.º 19]

5' - CTCCTCCTCCATCTGCTTGCTCAGGTCGCTGGTGAAGGTTCCCTCTCCGTGA
CCGGTG - 3' [SEC ID N.º 20]

5' - GCTAGGAGGAGGGGCTCCGCTGCTTGGTCCTCCGTTCTTCAGCCACTCGAT
GAACAGTCTCACGGC - 3' [SEC ID N.º 21]

La reacción de hibridación se llevó a cabo según lo descrito en el ejemplo 1a. El producto hibridado fue enlazado al vector pJB02 que había sido digerido con *NheI* y *Eco47III* según lo descrito en el ejemplo 1a para crear pJB02-exendina-4-Fc.

Ejemplo 1e

Construcción de ADN codificante de exendina-4-ASH

Se preparó el plásmido pJB02-ASH según lo descrito en el ejemplo 1b. El ADN codificante de la secuencia señal de Ig*κ* fusionada con la exendina-4 fue generado mediante la hibridación *in vitro* de los mismos oligonucleótidos solapantes y complementarios descritos en el ejemplo 1d. Las reacciones de hibridación se llevaron a cabo según lo descrito anteriormente. Se clonó el ADN en los sitios únicos *NheI* y *FspI* de pJB02-ASH para crear pJB02-exendina-4-ASH.

Ejemplo 1f

Construcción de ADN codificante de exendina-4-ligador-ASH

Se construyó el plásmido pJB02-ligador-ASH según lo descrito en el ejemplo 1c. El ADN codificante de la secuencia líder de Ig*κ* fusionada con la exendina-4 y con la parte 5' de la secuencia ligadora fue generado según lo tratado en el ejemplo 1d. Se clonó este ADN en los sitios únicos *NheI* y *BspEI* de pJB02-ligador-ASH para crear pJB02-exendina-4-ligador-ASH.

Ejemplo 1g

Construcción de ADN codificante de Val^{8}-GLP-1/C-Ex-Fc

Se preparó el plásmido pJB02-Exendina-4-Fc según lo descrito en el ejemplo 1d. Se escindió el ADN codificante de exendina-4 del vector con *AgeI* y *Eco47III*. El ADN codificante de Val^{8}-GLP-1/C-Ex fue generado mediante hibridación *in vitro* de los siguientes oligonucleótidos solapantes y complementarios:

5' -CCGGTCACGTGGAGGGCACCTTCACCTCCGACGTGTCCTCCTATCTGGA
GGGCCAGGCCGCCA - 3' [SEC ID N.º 22]

5' - AGGAATTCATCGCCTGGCTGGTGAAGGGCCGGGGCAGCAGCGG
AGCCCCTCCTCCTAGC - 3' [SEC ID N.º 23]

5' - CTCCAGATAGGAGGACACGTCGGAGGTGAAGGTGCCCTCCAC
GTGA - 3' [SEC ID N.º 24]

5' - GCTAGGAGGAGGGGCTCCGCTGCTGCCCCGGCCCTTCACCAGCCAGGCGA
TGAATTCCTTGGCGGCCTGGCC - 3' [SEC ID N.º 25]

La reacción de hibridación se llevó a cabo según lo descrito en el ejemplo 1a. Se unió el producto hibridado en lugar de Exendina-4 en el vector de expresión pJB02-Exendina-4-Fc para crear pJB02-Val^{8}-GLP-1/C-Ex-Fc.

Ejemplo comparativo 1h

Construcción de ADN codificante de Val^{8}-Glu^{22}-GLP-1-Fc

Se preparó el plásmido pJB02-Exendina-4-Fc según lo descrito en el ejemplo 1d. Se escindió el ADN codificante de exendina-4 del vector con *AgeI* y *Eco47III*. El ADN codificante de Val^{8}-Glu^{22}-GLP-1 fue generado mediante hibridación *in vitro* de los siguientes oligonucleótidos solapantes y complementarios:

5' -CCGGTCACGTGGAGGGCACCTTCACCTCCGACGTGTCCTCCTATCTCGA
GGAGCAGGCCGCCA - 3' [SEC ID N.º 26]

5' - AGGAGTTCATCGCCTGGCTGGTGAAGGGCCGGGGC - 3' [SEC ID N.º 27]

5' - GCCCCGGCCCTTCACCAGCCAGGCGATGAACTCCTTGGCGGCC
TGCTC - 3' [SEC ID N.º 28]

5' - CTCGAGATAGGAGGACACGTCGGAGGTGAAGGTGCCCT
CCACGTGA - 3' [SEC ID N.º 29]

La reacción de hibridación se llevó a cabo según lo descrito en el ejemplo 1a. Se unió el producto hibridado en lugar de Exendina-4 en el vector de expresión pJB02-Exendina-4-Fc para crear pJB02-Val^{8}-Glu^{22}-GLP-1-Fc.

Ejemplo comparativo 1i

Construcción de ADN codificante de Val^{8}-Glu^{22}GLP-1/C-Ex-Fc

Se preparó el plásmido pJB02-Exendina-4-Fc según lo descrito en el ejemplo 1d. Se escindió el ADN codificante de exendina-4 del vector con *AgeI* y *Eco47III*. El ADN codificante de Val^{8}-Glu^{22}GLP-1/C-Ex fue generado mediante hibridación *in vitro* de los siguientes oligonucleótidos solapantes y complementarios:

5' - CCGGTCACGTGGAGGGCACCTTCACCTCCGACGTGTCCTCCTATCTCGA
GGAGCAGGCCGCCA - 3' [SEC ID N.º 30]

5' - ACGAATTCATCGCCTGGCTGGTGAAGGGCCGGGGCAGCAGCGGA
GCCCCTCCTCCTAGC - 3' [SEC ID N.º 31]

5' - CTCGAGATAGGAGGACACGTCGGAGGTGAAGGTGCCC
TCCACGTGA - 3' [SEC ID N.º 32]

5' - GCTAGGAGGAGGGGCTCCGCTGCTGCCCCGGCCCTTCACCAGCCAGGCGA
TGAATTCCTTGGCGGCCTGCTC - 3' [SEC ID N.º 33]

La reacción de hibridación se llevó a cabo según lo descrito en el ejemplo 1a. Se unió el producto hibridado en lugar de Exendina-4 en el vector de expresión pJB02-Exendina-4-Fc para crear pJB02-Val^{8}-Glu^{22}-GLP-1/C-Ex-Fc.

Ejemplo 1j

Construcción de ADN codificante de Gly^{8}-GLP-1-Fc

Se preparó el plásmido pJB02-Exendina-4-Fc según lo descrito en el ejemplo 1d. Se escindió el ADN codificante de exendina-4 del vector con *AgeI* y *Eco47III*. El ADN codificante de Gly^{8}-GLP-1 fue generado mediante hibridación *in vitro* de los siguientes oligonucleótidos solapantes y complementarios:

5' - CCGGTCACGGCGAGGGCACCTTCACTAGTGACGTGTCCTCCTATCTGGA
GGGCCAGGCCGCCA - 3' [SEC ID N.º 34]

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

5' - CTCCAGATAGGAGGACACGTCACTAGTGAAGGTGCCCTC
GCCGTGA - 3' [SEC ID N.º 36]

5′ - GCCCCGGCCCTTCACCAGCCAGGCGATGAACTCCTTGGCGGC
CTGGCC - 3′ [SEC ID N.º 37]

La reacción de hibridación se llevó a cabo según lo descrito en el ejemplo 1a. Se unió el producto hibridado en lugar de Exendina-4 en el vector de expresión pJB02-Exendina-4-Fc para crear pJB02-Gly^{8}-GLP-1-Fc.

Ejemplo 2

Expresión de proteínas de fusión heterólogas

La expresión de las proteínas de fusión codificadas por los constructos de ADN del ejemplo 1 se llevó a cabo transfectando transitoriamente células HEK 293ANEB (tanto adherentes como en suspensión). Se contaron y se sembraron las células 24 horas antes de la transfección. La mezcla de transfección se preparó mezclando reactivo de transfección FuGene®6 (Roche Molecular Biochemicals, n.º de catálogo 1814443) con OptiMEM (Gibco/BRL) y se incubaron a temperatura ambiente durante 5 min, tras lo que se añadió ADN y se incubó la mezcla durante 15 min más. Inmediatamente antes de la transfección, se añadió medio de crecimiento recién preparado a la placa. Las tablas 1 y 2 proporcionan más datos sobre la transfección.

TABLA 1

Reactivos usados en las transfecciones transitorias de células 293ANEB

Recipiente de cultivo tisular	Número de células sembradas	ADN (µg)	FuGene (µl)	Medio optiMEM (ml)	Vol. del medio de crecimiento (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

TABLA 2

Composición de los medios

Medio de crecimiento y de transfección	Medio de cosecha
DMEM F12 3:1	Base de Hybritech
SBF al 5%	Ca^{2+} 1mM
HEPES 20mM	HEPES 20mM
L–glutamina 2mM	1 µg/ml de Nuselin (insulina humana)
50 µg/ml de geneticina (G418 NEO)	1 µg/ml de transferrina humana
50 µg/ml de tobromicina	50 µg/ml de tobromicina

Para las transfecciones a pequeña escala (recipientes de 35 mm-10 mm), se aclararon las células con PBS y se cambiaron a medios de cosecha 24 horas después de la transfección, recogiéndose y reemplazándose los medios cada 24 horas durante varios días. En el caso de las transfecciones a gran escala (frascos rotativos de 700 cm^2), se aclararon los frascos rotativos con PBS 48 horas después de la transfección y se cambiaron a medios de cosecha. Los medios

fueron recogidos y cambiados cada 24 horas durante al menos 10 días consecutivos. Rutinariamente, sólo se usaron 10 cosechas para la purificación posterior de las proteínas.

Ejemplo comparativo 3

Purificación de proteínas de fusión heterólogas

Ejemplo 3a

Purificación de Val^8-GLP-1-Fc

Se filtraron aproximadamente 4,5 litros de medio acondicionado (nivel de expresión de proteínas de fusión de aproximadamente 20 μg/ml) de las transfecciones a gran escala usando un sistema de filtros CUNO, y se concentraron hasta 250 ml usando un sistema de filtración de flujo tangencial Pro Flux con una membrana de filtración de 10 K. Se capturó la proteína Val^8-GLP-1-Fc con una columna de proteínas A HiTrap de 5 ml en 1 x PBS, pH de 7,4 a un caudal de 2 ml/min y se eluyó con ácido cítrico 50 mM, pH 3,3. Las fracciones (1 ml) se recogieron en tubos que contenían 4 ml de 1 x PBS y 100 μl de Tris 1M, pH 8.

Se mezclaron las fracciones que contenían la proteína de fusión determinadas mediante electroforesis en gel de poliacrilamida-SDS y CLAR en fase inversa sobre Zorbax C8, y se aplicaron a una columna Superdex 75 60/60 en 1 x PBS, pH 7,4 a un caudal de 10 ml/min. Se recogieron y se mezclaron las fracciones positivas (20 ml/tubo). Entonces se sometieron las fracciones mezcladas a una cromatografía en fase inversa C4 en agua y TFA al 0,1% a un caudal de 3 ml/min. Se eluyó Val^8-GLP-1-Fc usando un gradiente de B al 5% (TFA al 0,1% en acetonitrilo) hasta B al 100% en 70 min. Se recogieron las fracciones de eluyente (3 ml/tubo). Se retiró el acetonitrilo mediante secado al vacío y se añadió 1 ml de H_2O. Se sometió a diálisis la mezcla purificada (aproximadamente 32 ml) dos veces frente a 4 litros de 1 x PBS, pH 7,4.

Entonces se filtró la muestra sometida a diálisis usando una unidad de filtración de 0,22 um de MILLEX-GV, y se determinó la concentración usando una absorción a 280 nm.

Ejemplo 3b

Purificación de Val^8-GLP-1-ASH o Val^8-GLP-1-ligador-ASH

Se filtraron aproximadamente 6,5 litros de medio acondicionado (nivel de expresión de proteínas de fusión de aproximadamente 10 μg/ml) usando un sistema de filtros CUNO, y se concentraron hasta 380 ml usando un sistema de filtración de flujo tangencial ProFlux con una membrana de filtración de 10 K.

Se capturó la proteína de fusión usando una columna Q Fast Flow de 50 ml (Pharmacia) en Tris 20 mM, pH 7,4 a un caudal de 5ml/min. Se eluyó la proteína usando un gradiente del 0% al 50% de Tris 20 mM, pH 7,4; NaCl 1M en 10 CV, luego hasta B al 100% en 2 CV.

Entonces se mezclaron y se sometieron las fracciones que contenían la proteína de fusión a una cromatografía en fase inversa C4 en agua y TFA al 0,1% a un caudal de 5 ml/min. Se eluyó la proteína de fusión usando un gradiente de B al 20% (TFA al 0,1% en acetonitrilo) hasta B al 90% en 120 min. Se recogieron las fracciones (3,5 ml/tubo). Se retiró el acetonitrilo mediante secado al vacío.

Se diluyeron aproximadamente 9 ml de muestra mezclada con 1 x PBS, pH 7,4 hasta un volumen de 40 ml, y se sometieron a diálisis frente a 4 litros de 1 x PBS, pH 7,4 durante una noche. Se filtró la muestra y se determinó la concentración mediante absorbancia a 280 nm.

Ejemplo 3c comparativo

Purificación de Exendina-4-Fc

Se filtraron aproximadamente 4 litros de medio acondicionado (nivel de expresión de proteínas de fusión de aproximadamente 8 μg/ml) usando un sistema de filtros CUNO, y se concentraron hasta 250 ml usando un sistema de filtración de flujo tangencial Pro Flux con una membrana de filtración de 30 K.

La proteína exendina-4-Fc fue capturada con una columna proteínas A HiTrap de 5 ml en 1 x PBS, 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 fueron mezcladas, filtradas y sometidas a diálisis frente a 4 litros de 1 x PBS durante una noche. Se aplicó la mezcla sometida a diálisis a una columna Superdex 75 60/60 en 1 x PBS, pH 7,4; NaCl 0,5M, a un caudal de 10 ml/min. Las fracciones (20 ml/tubo) que contenían la proteína de fusión fueron recogidas, mezcladas y concentradas hasta aproximadamente 1 mg/ml. Entonces se filtraron las muestras concentradas usando una unidad de filtración de 0,22 um de MILLEX-GV.

Ejemplo 3d

Purificación de Exendina-4-ASH y de Exendina-4-ligador-ASH

Se filtraron aproximadamente 1,1 litros de medio acondicionado (nivel de expresión de proteínas de fusión de aproximadamente 6 μg/ml) usando un sistema de filtros CUNO, y se concentraron hasta 175 ml usando un sistema de filtración de flujo tangencial ProFlux con una membrana de filtración de 30 K.

Se capturó la proteína de fusión usando una columna de sefarosa-Q HiTrap de 5 ml (Pharmacia) en Tris 20 mM, pH 7,4 a un caudal de 2 ml/min. Se eluyó la proteína usando un gradiente del 0% al 50% de Tris 20 mM, pH 7,4; NaCl 1M en 12 CV, luego hasta B al 100% en 4 CV.

Entonces se mezclaron y se sometieron las fracciones que contenían la proteína de fusión a una cromatografía en fase inversa C4 en agua y TFA al 0,1% a un caudal de 5 ml/min. Se eluyó la proteína de fusión usando un gradiente de B al 10% (TFA al 0,1% en acetonitrilo) hasta B al 100% en 70 min. Se recogieron las fracciones (10 ml/tubo) que contenían la proteína de fusión. Se retiró el acetonitrilo usando una secadora de vacío.

Se sometieron a diálisis aproximadamente 8 ml de muestra mezclada frente a 4 litros de 1 x PBS, pH 7,4 durante toda la noche. Se filtró la muestra y se determinó la concentración mediante absorbancia a 280 nm. Se aplicó la muestra sometida a diálisis a una columna 26/60 Superdex 200 en 1 x PBS, pH 7,4; NaCl 0,5M, a un caudal de 2 ml/min. Las fracciones (3 ml/tubo) que contenían la proteína de fusión fueron recogidas, mezcladas, concentradas y filtradas.

Ejemplo 4

Caracterización de proteínas de fusión mediante electroforesis en gel de poliacrilamida-SDS

Se usó una electroforesis en gel de poliacrilamida-SDS seguida por una inmunotransferencia para analizar tanto la proteína de fusión purificada como el medio acondicionado de células transfectadas con diversos vectores de expresión de proteínas de fusión. La electroforesis en gel de poliacrilamida-SDS fue realizada en un sistema Novex Powerease 500 usando geles Precast de tris-glicina al 16% de Novex (EC6498), tampón de ejecución (10 x LC2675) y tampón de muestra (L2676). Se redujeron las muestras con DTT 50 mM y se calentaron durante 3-5 min a 95ºC antes de cargarlas.

Tras aplicar el gel de la electroforesis en gel de poliacrilamida-SDS, se usaron agua y tampón de transferencia (1 x Tris-Glicina Seprabuff (Owl Scientific n.º de cat. ER26-S) con metanol al 20%) para aclarar el SDS de los geles. Se usó un aparato de transferencia de Novex con PVDF (BioRad, n.º de cat. 162-0174) y membranas de nitrocelulosa (BioRad, n.º de cat. 1703965 ó 1703932). La transferencia se llevó a cabo a temperatura ambiente durante 90 min a 30-35 V. Se bloquearon las membranas en 1 x PBS con Tween-20 al 0,1% (Sigma, n.º de cat. P-7949) y leche al 5% (BioRad, n.º de cat. 170-6404) durante 1-12 horas a 4ºC. Se diluyen anticuerpos en 1 x PBS + leche al 5% y se incuban las transferencias en estas soluciones durante 1-2 h a 4ºC. Entre las incubaciones, se lavaron las transferencias cuatro veces durante 5 min cada una con 1 x PBS y Tween-20 al 0,2% a temperatura ambiente. El PBS se elaboró bien a partir de GIBCO 10 X PBS (n.º de cat. 70011) para proporcionar una composición final de fosfato de potasio monobásico 1 mM, fosfato de sodio dibásico 3 mM, cloruro de sodio 153 mM, pH 7,4, o a partir de bolsas de PBS de Sigma (n.º de cat. 1000-3) para proporcionar NaCl 120 mM, KCl 2,7 mM y fosfato 10 mM, pH 7,4 a 25ºC.

Los anticuerpos primarios eran bien anticuerpos policlonales anti-IgG1 de cabra o anti-ASH de conejo. El anticuerpo secundario era bien HRP de IgG anti-cabra o HRP de IgG anti-conejo. El anticuerpo secundario fue diluido 1:5000. Se usó un sistema ECL (Amersham Pharmacia Biotech, n.º de cat. RN2108 y n.º de cat. RPN1674H) para desarrollar las transferencias.

La figura comparativa 3A compara la proteína Fc purificada con el medio acondicionado de células transfectadas con pJB02-Val^{8}-GLP-1-Fc y pJB02-exendina-4-Fc. La disminución en la movilidad coincide con el aumento de tamaño debido a la parte de GLP-1 de la proteína de fusión. La figura 3B compara de manera similar la ASH purificada con el medio acondicionado de células transfectadas con pJB02-Val^{8}-GLP-2-ASH, pJB02-Val^{8}-GLP-1-ligador-ASH, pJB02-exendina-4-ASH o pJB02-exendina-4-ligador-ASH. La figura 4 identifica preparaciones de proteínas de fusión purificadas.

Ejemplo 5

Caracterización de proteínas de fusión usando espectrometría de masas

Todos los experimentos se realizaron en un espectrómetro de masas TofSpec-2E de Micromass equipado con electrónica de enfoque retardado, un reflectrón (usado en el análisis del intervalo peptídico de 0-8000 Da), un detector lineal (usado durante el análisis de señales de masa elevada/buenas) y un detector de post-aceleración (o P.A.D., usado para el análisis de señales de masa elevada/extremadamente bajas). La longitud de la trayectoria eficaz del instrumento en el modo lineal es de 1,2 metros, en el modo Reflectrón es de 2,3 metros. Se conectan dos detectores duales de placa

de microcanal para una detección del modo lineal y del modo reflectrón. El láser usado es el láser de nitrógeno de Laser Science Inc. VSL-337i, que funciona a 337 mm a 5 pulsaciones de láser por segundo. Todos los datos fueron obtenidos usando un digitalizador interno de 8 bits, 2 GHz, haciéndose una media de hasta 50 pulsaciones de láser por espectro.

El instrumento fue utilizado en el modo lineal para el análisis de las proteínas de fusión GLP-1 en cuestión. El detector lineal es un dispositivo que detecta iones que se transportan por el tubo de vuelo del instrumento MALDI-ToF-EM. Mide la abundancia iónica a lo largo del tiempo y envía una señal al digitalizador para su conversión. El digitalizador es un convertidor de los datos analógicos en digitales que permite la transferencia de la señal desde el espectrómetro de masas al ordenador, donde es reconstruida en el espectro m/z usable.

Se utilizó una solución de ácido sinapínico saturada recristalizada (diluida en Acn/H_2O 50/50 y TFA al 0,1%) como la matriz de ionización. El ácido sinapínico es una matriz apropiada para las proteínas de más de 10 kDa. Se usaron proteínas de referencia de masa apropiada para los archivos de calibración internos y externos con el fin de obtener determinaciones de masas exactas para las muestras analizadas. Todas las muestras fueron analizadas usando una proporción de 1:2 de muestra con respecto a la dilución matriz. El instrumento se configuró inicialmente según las siguientes condiciones del detector lineal:

Tensión de la fuente: 20,0 keV Tensión de los impulsos: 3,0 keV

Tensión de extracción: 20,0 keV Grueso láser: 50

Tensión de focalización: 16,0 keV Fino láser: 50

Detector lineal: 3,7 keV

P.A.D.: (fuera de línea)

Estas configuraciones fueron modificadas (según lo necesario) para proporcionar la mejor proporción señal/ruido y la resolución más elevada. La tabla 3 proporciona una caracterización de las diferentes proteínas de fusión GLP-1.

TABLA 3

Proteína de fusión	Masa esperada (kDa)	Masa determinada por EM (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
Exendina–4–IgG1	60,69	65,86
Val^8–GLP–1–ligador–ASH	70,70	69,89; 70,74
Exendina–4–ASH	70,56	70,62
Exendina–4–ligdor–ASH	71,56	71,62

CEx se refiere a una extensión C-terminal y comprende la secuencia de Ser-Ser-Gly-Ala-Pro-Pro-Pro-Ser.

El ligador es Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Gly-Ser.

Ejemplo 6

Actividad de las proteínas de fusión heterólogas

Se analizó la capacidad de las proteínas de fusión de la presente invención para activar al receptor de GLP-1 usando análisis *in vitro* tales como aquéllos descritos en el documento EP 619.322 concedido a Gelfand, *et al.,* y la patente estadounidense n.º: 5.120.712, respectivamente. La actividad de estos compuestos relativa a la actividad de Val^8-GLP-1(7-37)OH se presenta en la tabla 4. La figura 8 representa las curvas de dosis-respuesta *in vitro* para las proteínas de fusión Val^8-GLP-1 y exendina-4. Además, la tabla 5a y 5b proporcionan la actividad *in vitro* de una amplio grupo de análogos de GLP-1 que se pueden fusionar a una proteína albúmina para formar proteínas de fusión biológicamente activas. Estas actividades se comparan con GLP-1(7-37)OH.

TABLA 4

Actividad in vitro de proteínas de fusión GLP-1

Proteína de fusión	Actividad *in vitro* (% de Val^8–GLP–1)
Val^8–GLP–1–IgG1	1
Exendina–4–IgG1	240
Val^8–GLP–1–ligador–ASH	0,2
Exendina–4–ASH	20
Exendina–4–ligador–ASH	90
Exendina–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–ligador–IgG1	150
Exendina–4–C2–IgG1	250
Exendina–4–ligador–IgG1	330
Gly^8–Glu^{22}–GLP–1–CEx–ligador–ASH	4
Gly^8–Glu^{22}–GLP–1–CEx–ligador–IgG4	80

CEx se refiere a una extensión C-terminal y comprende la secuencia de Ser-Ser-Gly-Ala-Pro-Pro-Pro-Ser. El ligador es Gly-Gly-Gly-Gly-Ser-Gly-Gly-Gly-Gly-Ser-Gly-Gly-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 SEC ID N.º 13 a SEC ID N.º 31. La secuencia de aminoácidos de Val^8-GLP-1-albúmina de suero humano está representada por SEC ID N.º 13.

```
  1 HVEGTFTSDV SSYLEGQAAK EFIAWLVKGR GDAHKSEVAH RFKDLGEENF KALVLIAFAQ
 61 YLQQCPFEDH VKLVNEVTEF AKTCVADESA ENCDKSLHTL FGDKLCTVAT LRETYGEMAD
121 CCAKQEPERN ECFLQHKDDN PNLPRLVRPE VDVMCTAFHD NEETFLKKYL YEIARRHPYF
181 YAPELLFFAK RYKAAFTECC QAADKAACLL PKLDELRDEG KASSAKQRLK CASLQKFGER
241 AFKAWAVARL SQRFPKAEFA EVSKLVTDLT KVHTECCHGD LLECADDRAD LAKYICENQD
301 SISSKLKECC EKPLLEKSHC IAEVENDEMP ADLPSLAADF VESKDVCKNY AEAKDVFLGM
361 FLYEYARRHP DYSVVLLLRL AKTYETTLEK CCAAADPHEC YAKVFDEFKP LVEEPQNLIK
421 QNCELFEQLG EYKFQNALLV RYTKKVPQVS TPTLVEVSRN LGKVGSKCCK HPEAKRMPCA
481 EDYLSVVLNQ LCVLHEKTPV SDRVTKCCTE SLVNRRPCFS ALEVDETYVP KEFNAETFTF
541 HADICTLSEK ERQIKKQTAL VELVKHKPKA TKEQLKAVMD DFAAFVEKCC KADDKETCFA
601 EEGKKLVAAS QAALGL [SEC ID N.º 13]
```

La secuencia de aminoácidos de Val^8-GLP-1-ligador-albúmina de suero humano está representada por SEC ID N.º 14.

```
  1 HVEGTFTSDV SSYLEGQAAK EFIAWLVKGR GGGGGSGGGG SGGGGSDAHK SEVAHRFKDL
 61 GEENFKALVL IAFAQYLQQC PFEDHVKLVN EVTEFAKTCV ADESAENCDK SLHTLFGDKL
121 CTVATLRETY GEMADCCAKQ EPERNECFLQ HKDDNPNLPR LVRPEVDVMC TAFHDNEETF
181 LKKYLYEIAR RHPYFYAPEL LFFAKRYKAA FTECCQAADK AACLLPKLDE LRDEGKASSA
241 KQRLKCASLQ KFGERAFKAW AVARLSQRFP KAEFAEVSKL VTDLTKVHTE CCHGDLLECA
301 DDRADLAKYI CENQDSISSK LKECCEKPLL EKSHCIAEVE NDEMPADLPS LAADFVESKD
361 VCKNYAEAKD VFLGMFLYEY ARRHPDYSVV LLLRLAKTYE TTLEKCCAAA DPHECYAKVF
421 DEFKPLVEEP QNLIKQNCEL FEQLGEYKFQ NALLVRYTKK VPQVSTPTLV EVSRNLGKVG
481 SKCCKHPEAK RMPCAEDYLS VVLNQLCVLH EKTPVSDRVT KCCTESLVNR RPCFSALEVD
541 ETYVPKEFNA ETFTFHADIC TLSEKERQIK KQTALVELVK HKPKATKEQL KAVMDDFAAF
601 VEKCCKADDK ETCFAEEGKK LVAASQAALG L [SEC ID N.º 14]
```

La secuencia de aminoácidos de Gly^8-Glu^{22}-GLP-1-CEx-ligador-albúmina de suero humano está representada por SEC ID N.º 15.

```
  1 HGEGTFTSDV SSYLEEQAAK EFIAWLVKGR GSSGAPPPSG GGGGSGGGGS GGGGSDAHKS
 61 EVAHRFKDLG EENFKALVLI AFAQYLQQCP FEDHVKLVNE VTEFAKTCVA DESAENCDKS
121 LHTLFGDKLC TVATLRETYG EMADCCAKQE PERNECFLQH KDDNPNLPRL VRPEVDVMCT
181 AFHDNEETFL KKYLYEIARR HPYFYAPELL FFAKRYKAAF TECCQAADKA ACLLPKLDEL
241 RDEGKASSAK QRLKCASLQK FGERAFKAWA VARLSQRFPK AEFAEVSKLV TDLTKVHTEC
301 CHGDLLECAD DRADLAKYIC ENQDSISSKL KECCEKPLLE KSHCIAEVEN DEMPADLPSL
361 AADFVESKDV CKNYAEAKDV FLGMFLYEYA RRHPDYSVVL LLRLAKTYET TLEKCCAAAD
421 PHECYAKVFD EFKPLVEEPQ NLIKQNCELF EQLGEYKFQN ALLVRYTKKV PQVSTPTLVE
481 VSRNLGKVGS KCCKHPEAKR MPCAEDYLSV VLNQLCVLHE KTPVSDRVTK CCTESLVNRR
541 PCFSALEVDE TYVPKEFNAE TFTFHADICT LSEKERQIKK QTALVELVKH KPKATKEQLK
601 AVMDDFAAFV EKCCKADDKE TCFAEEGKKL VAASQAALGL [SEC ID N.º 15]
```

La secuencia de aminoácidos de exendina-4-albúmina de suero humano está representada por SEC ID N.º 16.

```
  1 HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSD AHKSEVAHRF KDLGEENFKA
 61 LVLIAFAQYL QQCPFEDHVK LVNEVTEFAK TCVADESAEN CDKSLHTLFG DKLCTVATLR
121 ETYGEMADCC AKQEPERNEC FLQHKDDNPN LPRLVRPEVD VMCTAFHDNE ETFLKKYLYE
181 IARRHPYFYA PELLFFAKRY KAAFTECCQA ADKAACLLPK LDELRDEGKA SSAKQRLKCA
241 SLQKFGERAF KAWAVARLSQ RFPKAEFAEV SKLVTDLTKV HTECCHGDLL ECADDRADLA
301 KYICENQDSI SSKLKECCEK PLLEKSHCIA EVENDEMPAD LPSLAADFVE SKDVCKNYAE
361 AKDVFLGMFL YEYARRHPDY SVVLLLRLAK TYETTLEKCC AAADPHECYA KVFDEFKPLV
421 EEPQNLIKQN CELFEQLGEY KFQNALLVRY TKKVPQVSTP TLVEVSRNLG KVGSKCCKHP
481 EAKRMPCAED YLSVVLNQLC VLHEKTPVSD RVTKCCTESL VNRRPCFSAL EVDETYVPKE
541 FNAETFTFHA DICTLSEKER QIKKQTALVE LVKHKPKATK EQLKAVMDDF AAFVEKCCKA
601 DDKETCFAEE GKKLVAASQA ALGL [SEC ID N.º 16]
```

La secuencia de aminoácidos de exendina-4-ligador-albúmina de suero humano está representada por SEC ID N.º 17.

```
  1 HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGGSGGGGS GGGGSDAHKS
 61 EVAHRFKDLG EENFKALVLI AFAQYLQQCP FEDHVKLVNE VTEFAKTCVA DESAENCDKS
121 LHTLFGDKLC TVATLRETYG EMADCCAKQE PERNECFLQH KDDNPNLPRL VRPEVDVMCT
181 AFHDNEETFL KKYLYEIARR HPYFYAPELL FFAKRYKAAF TECCQAADKA ACLLPKLDEL
241 RDEGKASSAK QRLKCASLQK FGERAFKAWA VARLSQRFPK AEFAEVSKLV TDLTKVHTEC
301 CHGDLLECAD DRADLAKYIC ENQDSISSKL KECCEKPLLE KSHCIAEVEN DEMPADLPSL
361 AADFVESKDV CKNYAEAKDV FLGMFLYEYA RRHPDYSVVL LLRLAKTYET TLEKCCAAAD
421 PHECYAKVFD EFKPLVEEPQ NLIKQNCELF EQLGEYKFQN ALLVRYTKKV PQVSTPTLVE
481 VSRNLGKVGS KCCKHPEAKR MPCAEDYLSV VLNQLCVLHE KTPVSDRVTK CCTESLVNRR
541 PCFSALEVDE TYVPKEFNAE TFTFHADICT LSEKERQIKK QTALVELVKH KPKATKEQLK
601 AVMDDFAAFV EKCCKADDKE TCFAEEGKKL VAASQAALGL [SEC ID N.º 17]
```

La secuencia de aminoácidos de Val^8-GLP-1-IgG1 está representada por SEC ID N.º 18.

```
  1 HVEGTFTSDV SSYLEGQAAK EFIAWLVKGR GAEPKSCDKT HTCPPCPAPE LLGGPSVFLF
 61 PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV
121 SVLTVLHQDW LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SREEMTKNQV
181 SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSDGS FFLYSKLTVD KSRWQQGNVF
241 SCSVMHEALH NHYTQKSLSL SPGK [SEC ID N.º 18]
```

La secuencia de aminoácidos de Val[8]-GLP-1-CEx-IgG1 está representada por SEC ID N.º 19.

```
  1 HVEGTFTSDV SSYLEGQAAK EFIAWLVKGR GSSGAPPPSA EPKSCDKTHT CPPCPAPELL
 61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR
181 EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTTP PVLDSDGSFF LYSKLTVDKS
241 RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GK [SEC ID N.º 19]
```

La secuencia de aminoácidos de Val[8]-Glu[22]-GLP-1-IgG1 está representada por SEC ID N.º 20.

```
  1 HVEGTFTSDV SSYLEEQAAK EFIAWLVKGR GAEPKSCDKT HTCPPCPAPE LLGGPSVFLF
 61 PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV
121 SVLTVLHQDW LNGKEYKCKV SNKALPAPIE KTISKAKGQP REPQVYTLPP SREEMTKNQV
181 SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSDGS FFLYSKLTVD KSRWQQGNVF
241 SCSVMHEALH NHYTQKSLSL SPGK [SEC ID N.º 20]
```

La secuencia de aminoácidos de Val[8]-Glu[22]-GLP-1-CEx-IgG1 está representada por SEC ID N.º 21.

```
  1 HVEGTFTSDV SSYLEEQAAK EFIAWLVKGR GSSGAPPPSA EPKSCDKTHT CPPCPAPELL
 61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR
181 EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTTP PVLDSDGSFF LYSKLTVDKS
```

La secuencia de aminoácidos de Gly[8]-Glu[22]GLP-1-C2-IgG1 está representada por SEC ID N.º 22.

```
  1 HGEGTFTSDV SSYLEEQAAK EFIAWLVKGR GSSGASSGAA EPKSCDKTHT CPPCPAPELL
 61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR
181 EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTTP PVLDSDGSFF LYSKLTVDKS
241 RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GK [SEC ID N.º 22]
```

La secuencia de aminoácidos de Gly[8]-Glu[22]-GLP-1-CEx-ligador-IgG1 está representada por SEC ID N.º 23.

```
  1 HGEGTFTSDV SSYLEEQAAK EFIAWLVKGR GSSGAPPPSG GGGSGGGGSG GGGSAEPKSC
 61 DKTHTCPPCP APELLGGPSV FLFPPKPKDT LMISRTPEVT CVVVDVSHED PEVKFNWYVD
121 GVEVHNAKTK PREEQYNSTY RVVSVLTVLH QDWLNGKEYK CKVSNKALPA PIEKTISKAK
181 GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTPPVLDS
241 DGSFFLYSKL TVDKSRWQQG NVFSCSVMHE ALHNHYTQKS LSLSPGK [SEC ID N.º 23]
```

La secuencia de aminoácidos de Gly^{8}-Glu^{22}-GLP-1-CEx-ligador-IgG4 está representada por SEC ID N.º 24.

```
  1 HGEGTFTSDV SSYLEEQAAK EFIAWLVKGR GSSGAPPPSG GGGSGGGGSG GGGSAESKYG
 61 PPCPSCPAPE FLGGPSVFLF PPKPKDTLMI SRTPEVTCVV VDVSQEDPEV QFNWYVDGVE
121 VHNAKTKPRE EQFNSTYRVV SVLTVLHQDW LNGKEYKCKV SNKGLPSSIE KTISKAKGQP
181 REPQVYTLPP SQEEMTKNQV SLTCLVKGFY PSDIAVEWES NGQPENNYKT TPPVLDSDGS
241 FFLYSRLTVD KSRWQEGNVF SCSVMHEALH NHYTQKSLSL SLGK [SEC ID N.º 24]
```

La secuencia de aminoácidos de Gly^{8}-Glu^{22}-GLP-1-CEx-2ligador-IgG1 está representada por SEC ID N.º 25.

```
  1 HGEGTFTSDV SSYLEEQAAK EFIAWLVKGR GSSGAPPPSG GGGSGGGGSG GGGSGGGGSG
 61 GGGSGGGGSA EPKSCDKTHT CPPCPAPELL GGPSVFLFPP KPKDTLMISR TPEVTCVVVD
121 VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN
181 KALPAPIEKT ISKAKGQPRE PQVYTLPPSR EEMTKNQVSL TCLVKGFYPS DIAVEWESNG
241 QPENNYKTTP PVLDSDGSFF LYSKLTVDKS RWQQGNVFSC SVMHEALHNH YTQKSLSLSP
301 GK [SEC ID N.º 25]
```

La secuencia de aminoácidos de Gly^{8}-Glu^{22}-GLP-1-2ligador-IgG1 está representada por SEC ID N.º 26.

```
  1 HGEGTFTSDV SSYLEEQAAK EFIAWLVKGR GGGGGSGGGG SGGGGSGGGG SGGGGSGGGG
 61 SAEPKSCDKT HTCPPCPAPE LLGGPSVFLF PPKPKDTLMI SRTPEVTCVV VDVSHEDPEV
121 KFNWYVDGVE VHNAKTKPRE EQYNSTYRVV SVLTVLHQDW LNGKEYKCKV SNKALPAPIE
181 KTISKAKGQP REPQVYTLPP SREEMTKNQV SLTCLVKGFY PSDIAVEWES NGQPENNYKT
241 TPPVLDSDGS FFLYSKLTVD KSRWQQGNVF SCSVMHEALH NHYTQKSLSL SPGK
[SEC ID N.º 26]
```

La secuencia de aminoácidos de Gly^{8}-Glu^{22}-GLP-1-2CEx-IgG1 está representada por SEC ID N.º 27.

```
  1 HGEGTFTSDV SSYLEEQAAK EFIAWLVKGR GSSGAPPPSS SGAPPPSAEP KSCDKTHTCP
 61 PCPAPELLGG PSVFLFPPKP KDTLMISRTP EVTCVVVDVS HEDPEVKFNW YVDGVEVHNA
121 KTKPREEQYN STYRVVSVLT VLHQDWLNGK EYKCKVSNKA LPAPIEKTIS KAKGQPREPQ
181 VYTLPPSREE MTKNQVSLTC LVKGFYPSDI AVEWESNGQP ENNYKTTPPV LDSDGSFFLY
241 SKLTVDKSRW QQGNVFSCSV MHEALHNHYT QKSLSLSPGK [SEC ID N.º 27]
```

La secuencia de aminoácidos de Gly^{8}-Glu^{22}-Val^{25}-Ile^{33}GLP-1-CEx-ligador-IgG1 está representada por SEC ID N.º 28.

```
  1 HGEGTFTSDV SSYLEEQAVK EFIAWLIKGR GSSGAPPPSG GGGSGGGGSG GGGSAEPKSC
 61 DKTHTCPPCP APELLGGPSV FLFPPKPKDT LMISRTPEVT CVVVDVSHED PEVKFNWYVD
121 GVEVHNAKTK PREEQYNSTY RVVSVLTVLH QDWLNGKEYK CKVSNKALPA PIEKTISKAK
181 GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTPPVLDS
241 DGSFFLYSKL TVDKSRWQQG NVFSCSVMHE ALHNHYTQKS LSLSPGK [SEC ID N.º 28]
```

La secuencia de aminoácidos de Exendina-4-IgG1 está representada por SEC ID N.º 29.

```
  1 HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSA EPKSCDKTHT CPPCPAPELL
 61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR
181 EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTTP PVLDSDGSFF LYSKLTVDKS
241 RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GK [SEC ID N.º 29]
```

La secuencia de aminoácidos de Exendina-4-C2-IgG1 está representada por SEC ID N.º 30.

```
  1 HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGASSGAA EPKSCDKTHT CPPCPAPELL
 61 GGPSVFLFPP KPKDTLMISR TPEVTCVVVD VSHEDPEVKF NWYVDGVEVH NAKTKPREEQ
121 YNSTYRVVSV LTVLHQDWLN GKEYKCKVSN KALPAPIEKT ISKAKGQPRE PQVYTLPPSR
181 EEMTKNQVSL TCLVKGFYPS DIAVEWESNG QPENNYKTTP PVLDSDGSFF LYSKLTVDKS
241 RWQQGNVFSC SVMHEALHNH YTQKSLSLSP GK [SEC ID N.º 30]
```

La secuencia de aminoácidos de Exendina-4-ligador-IgG1 está representada por SEC ID N.º 31.

```
  1 HGEGTFTSDL SKQMEEEAVR LFIEWLKNGG PSSGAPPPSG GGGSGGGGSG GGGSAEPKSC
 61 DKTHTCPPCP APELLGGPSV FLFPPKPKDT LMISRTPEVT CVVVDVSHED PEVKFNWYVD
121 GVEVHNAKTK PREEQYNSTY RVVSVLTVLH QDWLNGKEYK CKVSNKALPA PIEKTISKAK
181 GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTPPVLDS
241 DGSFFLYSKL TVDKSRWQQG NVFSCSVMHE ALHNHYTQKS LSLSPGK [SEC ID N.º 31]
```

TABLA 5a

Actividad in vitro de análogos de GLP-1

Compuesto de GLP–1	Activación del receptor de GLP–1
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

TABLA 5a (continuación)

Compuesto de GLP–1	Activación del receptor de GLP–1
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

TABLA 5a (continuación)

Compuesto de GLP–1	Activación del receptor de GLP–1
Val^{8}–Gly^{33}–GLP–1(7–37)OH	0,01
Val^{8}–Glu^{31}–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^{22}–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^{39}–GLP–1(7–37)OH	0,08
Val^{8}–Lys^{33}–Asn^{39}–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

TABLA 5b

Actividad in vitro de análogos de GLP-1

Compuesto de GLP–1	Activaciones del receptor 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^{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

TABLA 5b (continuación)

Compuesto de GLP–1	Activaciones del receptor GLP–1
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^{11}–GLP–1(7–37)OH	1,69; 1,79
Val^{8}–Trp^{16}–Glu^{11}–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^{16}–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–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

TABLA 6

Actividad in vitro de análogos de GLP/exendina

Secuencia peptídico	Actividad *in vitro* (% de Val^8–GLP–1(7–37)OH)
HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGP–NH2	6,21
HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGA PP PS–NH2	6,75; 3,25
HVEGTFTSDLSKQMEEEAVRLFIAWLVKGRG	2,86
HVEGTFTSDVSSYLEEEAVRLFIAWLVKGRG	1,47
HVEGTFTSDLSKQMEGQAAKEFIAWLVKGRG	0,11
HVEGTFTSDVSKQMEGQAAKEFIAWLVKGRG	0,04
HGEGTFTSDLSKQMEGQAAKEFIEWLKNGGP–NH2	1,44
HGEGTFTSDLSKQMEEEAAKEFIEWLKNGGP–NH2	2,80
HGEGTFTSDVSSYLEEEAVRLFIEWLKNGGP–NH2	5,40
HGEGTFTSDLSSYLEEEAVRLFIEWLKNGGP–NH2	5,07
HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRPSSGA PPPS–NH2	3,30
HAEGTFTSDVSKQLEEEAAKEFIAWLVKGRG	2,15
HVEGTFTSDVSSYLEGQAAKEFIEWLKNGGP–NH2	2,36
HGEGTFTSDLSKQMEEEAVRLFIAWLVKGRG	3,25
HVEGTFTSDVSSYLEEEAAKEFIAWLVKGRG	1,00
HVEGTFTSDVSSYLEGQAAKEFIAWLKNGRG	0,20
HVEGTFTSDVSSYLEGQAAKEFIAWLVKGRG	1,00
HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG	2,12

Ejemplo 7

Farmacocinética in vivo de Val^8-GLP-1-IgG1 y Val^8-GLP-1-ASH

Se realizó un estudio de farmacocinética de Val^8-GLP-1-IgG1 y Val^8-GLP-1-ASH en monos cynomologus. Se administraron a los monos 5,6 nmoles/kg bien de Val^8-GLP-1-IgG1 purificada o de Val^8-GLP-1-ASH purificada. Los compuestos fueron administrados por medio de bolos intravenosos. Se recogieron muestras de sangre antes de la dosis y a las 0,083; 0,25; 0,5; 1, 4, 8, 12, 24, 48, 72, 96, 120, 144, 168 y 216 horas de la administración en tubos que contenían EDTA. Se determinaron las concentraciones de Val^8-GLP-1 inmunorreactivo en plasma usando un radioinmunoanálisis que utilizaba un antisuero policlonal cuya especificidad primaria era para la región N-terminal (7-16) de Val^8-GLP-1(7-37). La figura 9 representa la concentración en plasma de Val^8-GLP-1-Fc y de Val^8-GLP-1-ligador-ASH tras una sola dosis intravenosa a dos monos cynomologus. La proteína de fusión Fc resultó tener una vida media de aproximadamente 45 horas y la fusión de albúmina, una vida media de aproximadamente 87 horas.

Ejemplo 8

Farmacodinámicas in vivo de exendina-4-IgG1

Se estudiaron dos perros sabuesos macho normales canulados crónicamente tras una noche en ayunas. Se abrieron puertos de acceso vasculares arteriales y venosos, se insertó un catéter percutáneamente en una vena cefálica y se aseguró. Se introdujeron los animales en jaulas y se unieron sus catéteres a un sistema de pivote/soga. Se inyectó una solución que contenía la proteína de fusión exendina-4-IgG1 (PM: 11,8 μM) intravenosamente (1,0 nmoles/kg) a través del catéter de la vena cefálica. Entonces se aclaró el catéter con 10 ml de solución salina. Dos horas después, se inició un clamp hiperglicémico (150 mg/dl) y se continuó durante tres horas. Se extrajeron muestras sanguíneas arteriales durante este período de 5 horas para la determinación de las concentraciones de la proteína de fusión, la glucosa y la insulina en plasma.

Se compararon los resultados de este estudio con los de un estudio previo similar en el que ambos animales habían recibido un bolo de solución salina s.c., y habían sido estudiados 3 horas más tarde usando un clamp hiperglicémico (150 mg/dl) de 3 horas.

En ambos conjuntos de estudios, se determinaron las concentraciones de glucosa en plasma usando un analizador de glucosa de Beckman. Las concentraciones de insulina en plasma fueron determinadas por empleados de Linco Research, Inc. usando un equipo RIA desarrollado en sus laboratorios. Los datos se ilustran en las figuras 10 y 11.

Ejemplo 9

Farmacocinéticas in vivo de Gly^{8}-Glu^{22}-GLP-1-CEx-ligador-IgG1

Dos grupos de tres perros sabuesos macho normales recibieron 0,1 mg/kg de Gly^{8}-Glu^{22}-GLP-1-CEx-ligador-IgG1 mediante una administración subcutánea (s.c.) o intravenosa (i.v.). Se determinó la inmunorreactividad de las concentraciones en plasma de Gly^{8}-Glu^{22}-GLP-1-CEx-ligador-IgG1 mediante radioinmunoanálisis en las muestras recogidas 30 minutos antes de la dosis hasta 216 horas posteriores a la dosis tanto para los grupos i.v. como para los grupos s.c. Estas concentraciones se usaron posteriormente para determinar los parámetros farmacocinéticos presentados. La vida media de eliminación media de Gly^{8}-Glu^{22}-GLP-1-CEx-ligador-IgG1 administrado i.v. resultó ser de aproximadamente 55 horas y el aclaramiento corporal total de 1,5 ml/h/kg. La vida media de eliminación media de Gly^{8}-Glu^{22}-GLP-1-CEx-ligador-IgG1 administrado s.c. resultó ser de aproximadamente 38 horas.

REIVINDICACIONES

1. Una proteína de fusión heteróloga que comprende un primer polipéptido con un terminal N y un terminal C fusionados con un segundo polipéptido con un terminal N y un terminal C, siendo el primer polipéptido un compuesto de GLP-1 y seleccionándose el segundo polipéptido entre:

a) albúmina humana;

b) análogos de albúmina humana; y

c) fragmentos de albúmina humana,

y en el que el terminal C del primer polipéptido está fusionado con el terminal N del segundo polipéptido, y en el que la proteína de fusión es biológicamente activa y tiene una vida media en plasma mayor que el compuesto de GLP-1 solo.

2. La proteína de fusión heteróloga de la reivindicación 1, en la que el terminal C del primer polipéptido está fusionado con el terminal N del segundo polipéptido mediante un ligador peptídico.

3. La proteína de fusión heteróloga de la reivindicación 2, en la que el ligador peptídico se selecciona entre

a) un péptido rico en glicina;

b) un péptido que tiene una secuencia $[Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser]_n$ en la que n es 1, 2, 3, 4. 5 ó 6; y

c) un péptido que tiene una secuencia $[Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser]_3$.

4. La proteína de fusión heteróloga de una cualquiera de las reivindicaciones anteriores, en la que el compuesto de GLP-1 comprende la secuencia de fórmula 1 [SEC ID N.º 2]

```
 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
Ile-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-
40  41  42  43  44  45
Xaa-Xaa-Xaa-Xaa-Xaa-Xaa
```

Fórmula I (**SEC ID N.º 2**)

en la que:

el Xaa de la posición 8 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

el Xaa de la posición 9 es Glu, Asp o Lys;

el Xaa de la posición 11 es Thr, Ala, Gly, Ser, Leu, Ile, Val, Glu, Asp o Lys;

el Xaa de la posición 14 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp o Lys;

el Xaa de la posición 16 es Val, Ala, Gly, Ser, Thr, Leu, Ile, Tyr, Glu, Asp, Trp o Lys;

el Xaa de la posición 17 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp o Lys;

el Xaa de la posición 18 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, Trp, Tyr o Lys;

el Xaa de la posición 19 es Tyr, Phe, Trp, Glu, Asp, Gln o Lys;

el Xaa de la posición 20 es Leu, Ala, Gly, Ser, Thr, Ile, Val, Glu, Asp, Met, Trp, Tyr o Lys;

el Xaa de la posición 21 es Glu, Asp o Lys;

el Xaa de la posición 22 es Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

el Xaa de la posición 23 es Gln, Asn, Arg, Glu, Asp o Lys;

el Xaa de la posición 24 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Arg, Glu, Asp o Lys;

el Xaa de la posición 25 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

el Xaa de la posición 26 es Lys, Arg, Gln, Glu, Asp o His;

el Xaa de la posición 27 es Leu, Glu, Asp o Lys;

el Xaa de la posición 30 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

el Xaa de la posición 31 es Trp, Phe, Tyr, Glu, Asp o Lys;

el Xaa de la posición 32 es Leu, Gly, Ala, Ser, Thr, Ile, Val, Glu, Asp o Lys;

el Xaa de la posición 33 es Val, Gly, Ala, Ser, Thr, Leu, Ile, Glu, Asp o Lys;

el Xaa de la posición 34 es Asn, Lys, Arg, Glu, Asp o His;

el Xaa de la posición 35 es Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

el Xaa de la posición 36 es Gly, Arg, Lys, Glu, Asp o His;

el Xaa de la posición 37 es Pro, Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys, o está eliminado;

el Xaa de la posición 38 es Ser, Arg, Lys, Glu, Asp o His, o está eliminado;

el Xaa de la posición 39 es Ser, Arg, Lys, Glu, Asp o His, o está eliminado;

el Xaa de la posición 40 es Gly, Asp, Glu o Lys, o está eliminado;

el Xaa de la posición 41 es Ala, Phe, Trp, Tyr, Glu, Asp o Lys, o está eliminado;

el Xaa de la posición 42 es Ser, Pro, Lys, Glu o Asp, o está eliminado;

el Xaa de la posición 43 es Ser, Pro, Glu, Asp o Lys, o está eliminado;

el Xaa de la posición 44 es Gly, Pro, Glu, Asp o Lys, o está eliminado; y

el Xaa de la posición 45 es Ala, Ser, Val, Glu, Asp o Lys, o está eliminado;

con la condición de que cuando el aminoácido de la posición 37, 38, 39, 40, 41, 42, 43 ó 44 esté eliminado, entonces todos los aminoácidos secuencia abajo de ese aminoácido también estarán eliminados.

5. La proteína de fusión heteróloga de una cualquiera de las reivindicaciones anteriores, en la que el compuesto de GLP-1 no tiene más de 6 aminoácidos que son diferentes del correspondiente aminoácido de GLP-1-(7-37)OH, GLP-1(7-36)OH o exendina-4.

6. La proteína de fusión heteróloga de la reivindicación 5, en la que el compuesto de GLP-1 no tiene más de 5 aminoácidos que son diferentes del correspondiente aminoácido de GLP-1-(7-37)OH, GLP-1(7-36)OH o exendina-4.

7. La proteína de fusión heteróloga de la reivindicación 6, en la que el compuesto de GLP-1 no tiene más de 4 aminoácidos que son diferentes del correspondiente aminoácido de GLP-1-(7-37)OH, GLP-1(7-36)OH o exendina-4.

8. La proteína de fusión heteróloga de la reivindicación 7, en la que el compuesto de GLP-1 no tiene más de 3 aminoácidos que son diferentes del correspondiente aminoácido de GLP-1-(7-37)OH, GLP-1(7-36)OH o exendina-4.

9. La proteína de fusión heteróloga de la reivindicación 8, en la que el compuesto de GLP-1 no tiene más de 2 aminoácidos que son diferentes del correspondiente aminoácido de GLP-1-(7-37)OH, GLP-1(7-36)OH o exendina-4.

10. La proteína de fusión heteróloga de una cualquiera de las reivindicaciones anteriores, en la que el compuesto de GLP-1 tiene glicina o valina en la posición 8.

11. La proteína de fusión heteróloga de una cualquiera de las reivindicaciones 1 a 10, en la que el segundo polipéptido es albúmina humana.

12. La proteína de fusión heteróloga de la reivindicación 11, en la que el segundo polipéptido tiene la secuencia de SEC ID N.º 34.

13. La proteína de fusión heteróloga de una cualquiera de las reivindicaciones 1 a 10, en la que el segundo polipéptido es un fragmento N-terminal de albúmina.

14. Una proteína de fusión heteróloga de una cualquiera de las reivindicaciones 1 a 13 para su uso como un medicamento.

15. Una proteína de fusión heteróloga de una cualquiera de las reivindicaciones 1 a 13 para su uso en el tratamiento de la diabetes mellitus no dependiente de la insulina.

16. Una proteína de fusión heteróloga de una cualquiera de las reivindicaciones 1 a 13 para su uso en el tratamiento de la obesidad.

17. Una formulación farmacéutica adaptada para el tratamiento de pacientes con diabetes mellitus no dependiente de la insulina que comprende una proteína de fusión heteróloga de una cualquiera de las reivindicaciones 1 a 13.

18. Una formulación farmacéutica adaptada para el tratamiento de pacientes con obesidad que comprende una proteína de fusión heteróloga de una cualquiera de las reivindicaciones 1 a 13.

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        [ SEC ID N.º 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
```

Continuación de la Fig. 2

```
                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.º 34]

Fig. 3

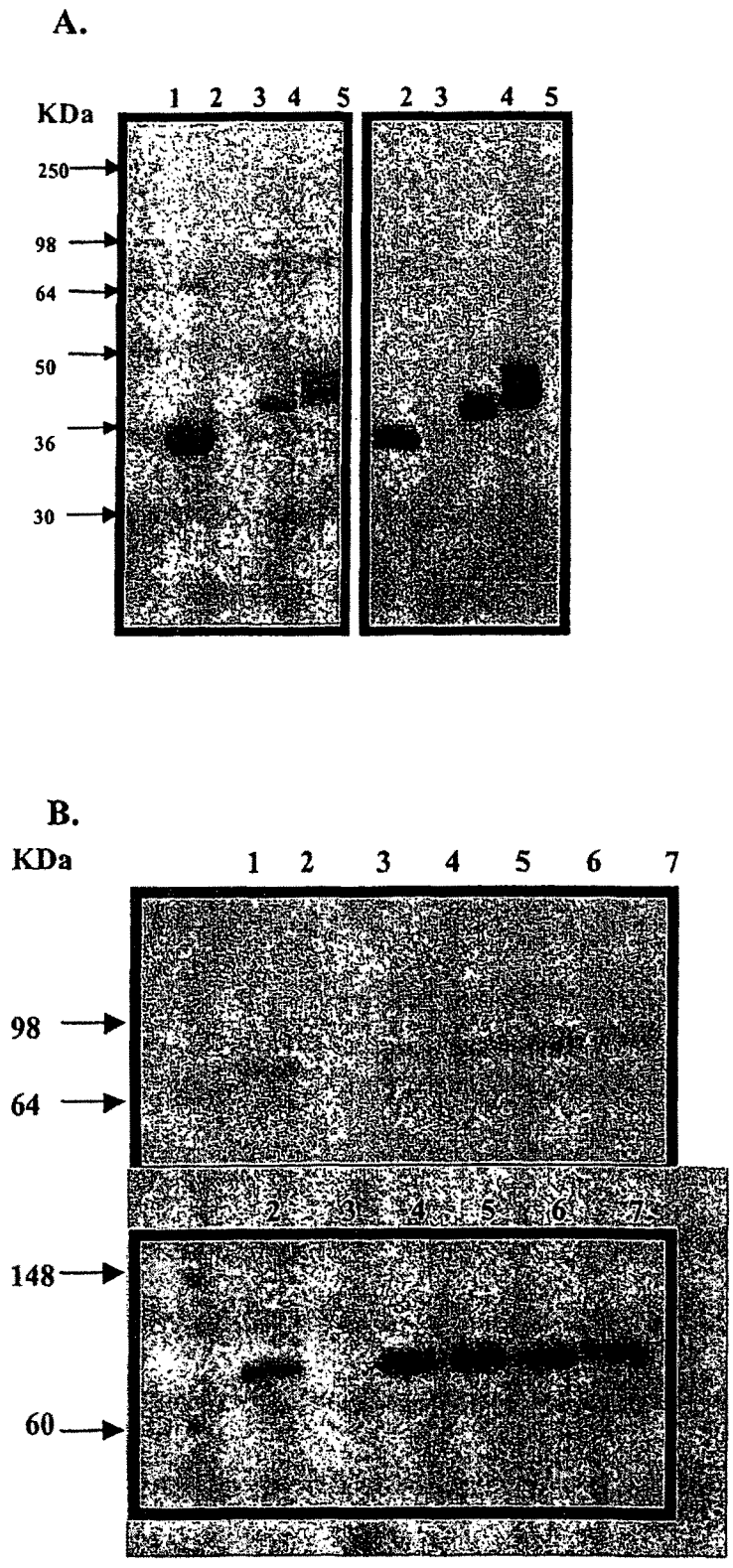

A.
KDa
1 2 3 4 5 2 3 4 5
250
98
64
50
36
30
B.
KDa
1 2 3 4 5 6 7
98
64
2 3 4 5 6 7
148
60

Fig. 4

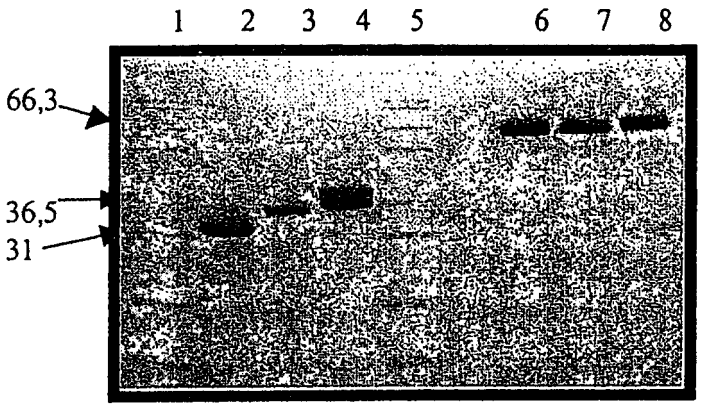


Fig. 5

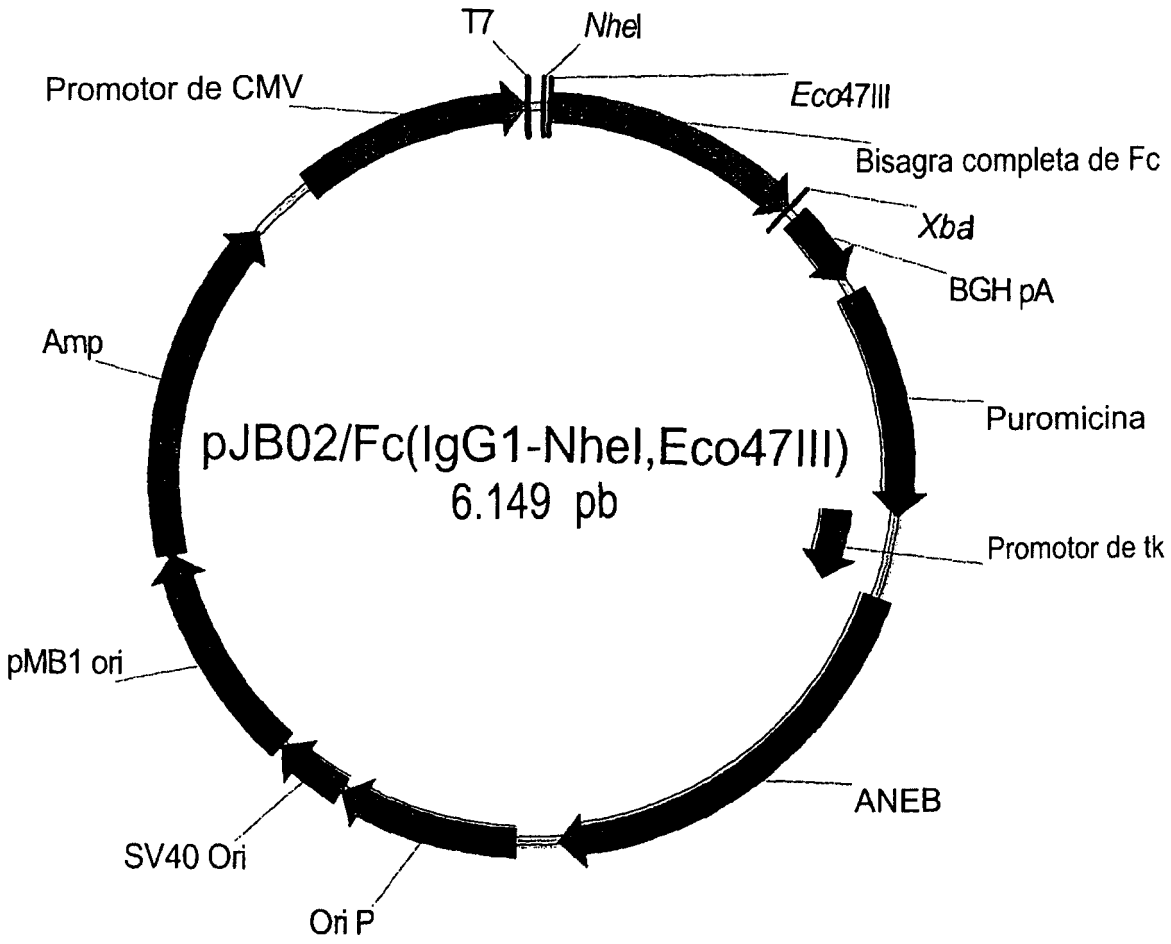

T7
NheI
Promotor de CMV
Eco47III
Bisagra completa de Fc
XbaI
BGH pA
Amp
pJB02/Fc(IgG1-NheI,Eco47III)
6.149 pb
Puromicina
Promotor de tk
pMB1 ori
ANEB
SV40 Ori
Ori P

Fig. 6

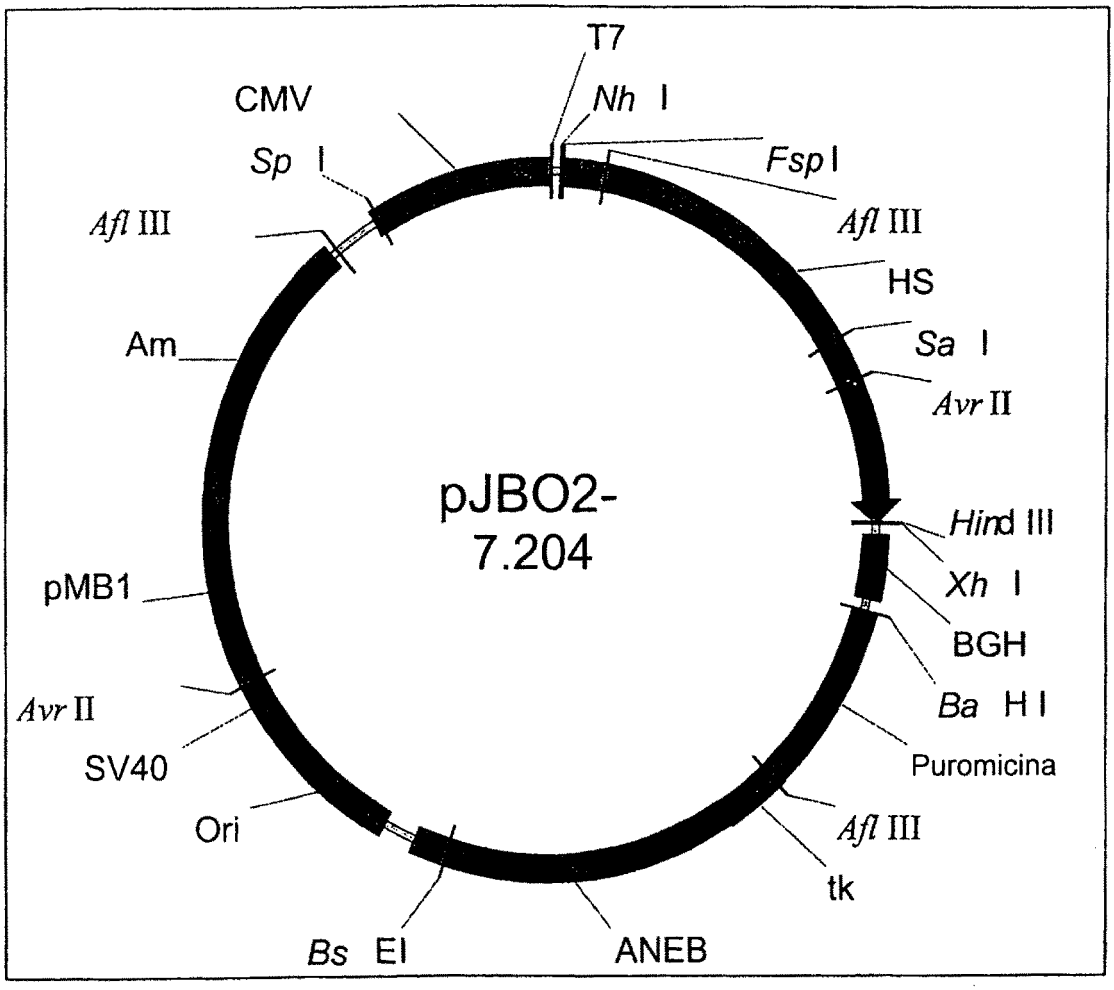

T7
Nh I
Fsp I
Afl III
HS
Sa I
Avr II
Hind III
Xh I
BGH
Ba H I
Puromicina
Afl III
tk
ANEB
Bs EI
Ori
SV40
Avr II
pMB1
Am
Afl III
Sp I
CMV
pJBO2-
7.204

Fig. 7

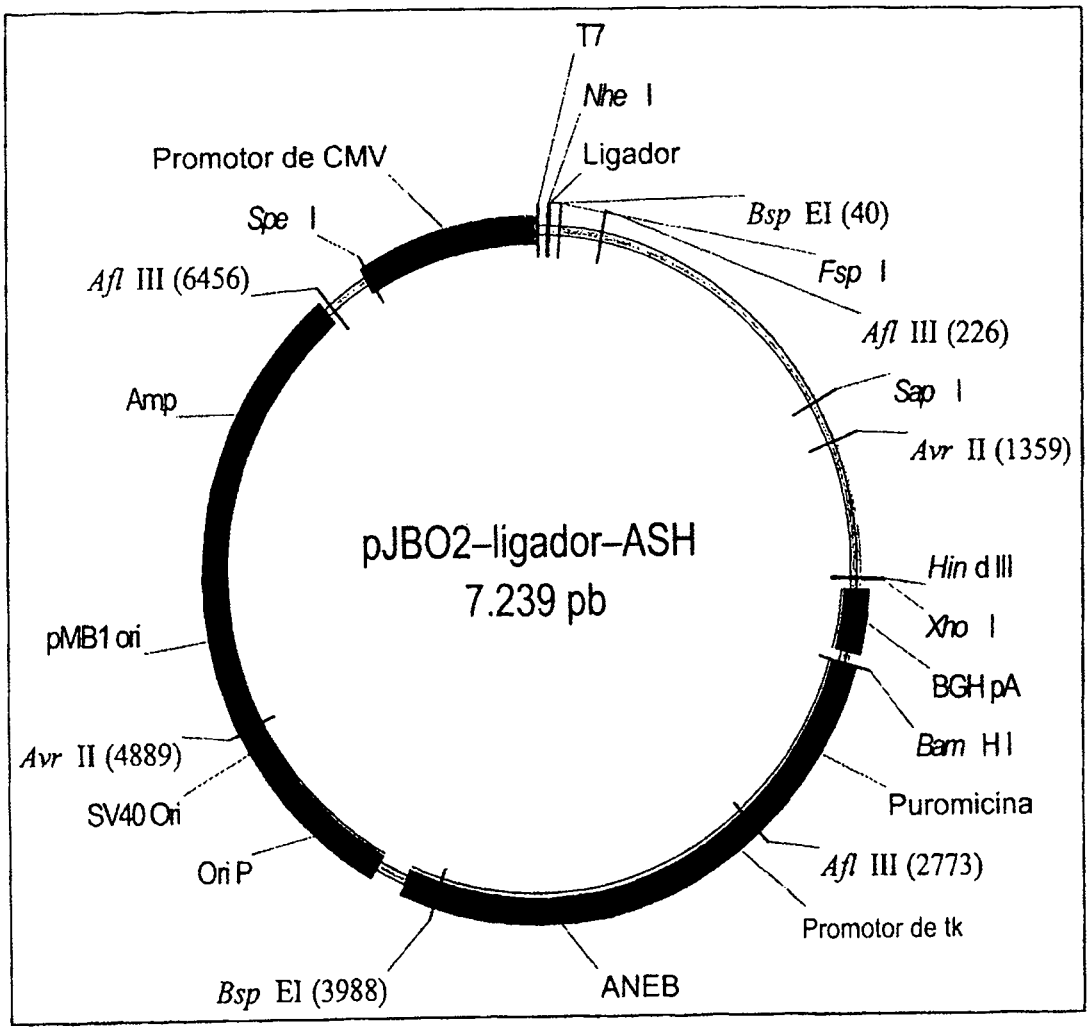

T7
Nhe I
Ligador
Promotor de CMV
Spe I
Bsp EI (40)
Fsp I
Afl III (6456)
Afl III (226)
Amp
Sap I
Avr II (1359)
pJBO2–ligador–ASH
7.239 pb
Hin d III
Xho I
pMB1 ori
BGH pA
Bam H I
Avr II (4889)
Puromicina
SV40 Ori
Afl III (2773)
Ori P
Promotor de tk
Bsp EI (3988)
ANEB

Fig. 8

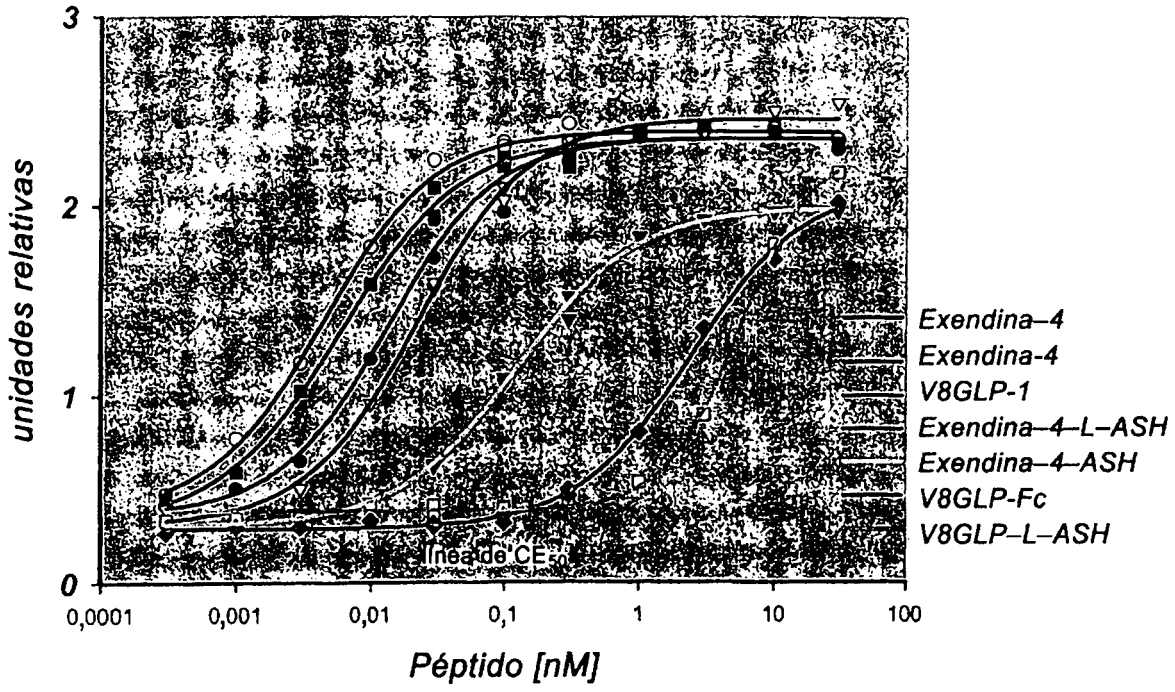

unidades relativas
3
2
1
0
0,0001
0,001
0,01
0,1
1
10
100
Péptido [nM]
Línea de CE50
Exendina–4
Exendina-4
V8GLP-1
Exendina–4–L–ASH
Exendina–4–ASH
V8GLP-Fc
V8GLP–L–ASH

Fig. 9

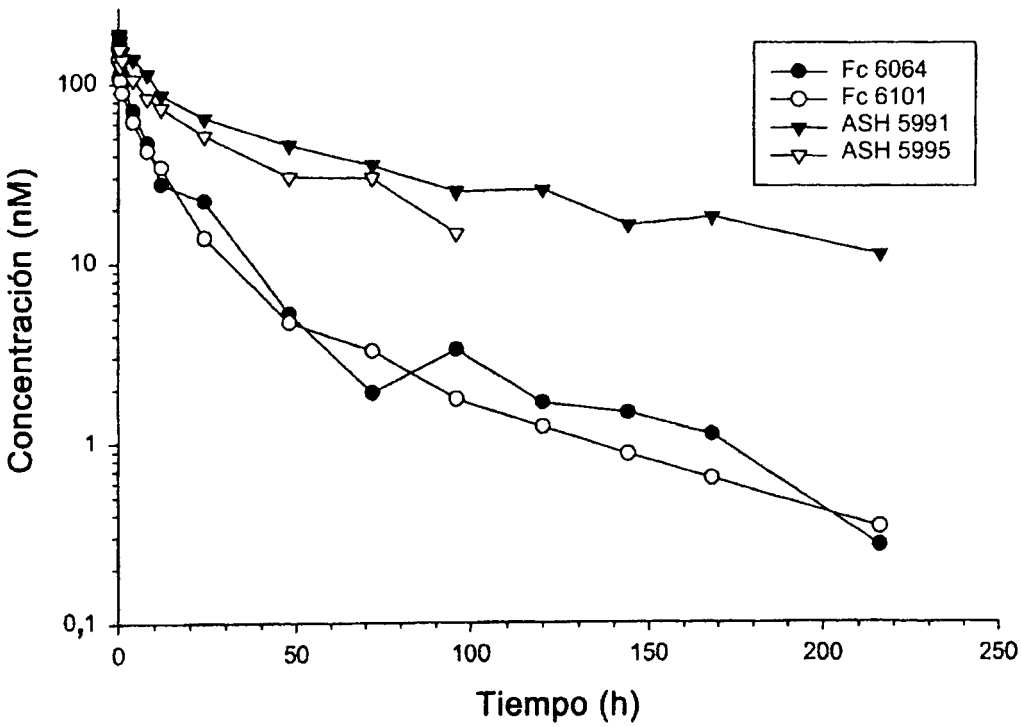

Fc 6064
Fc 6101
ASH 5991
ASH 5995
Concentración (nM)
100
10
1
0,1
0
50
100
150
200
250
Tiempo (h)

Fig. 10

A.

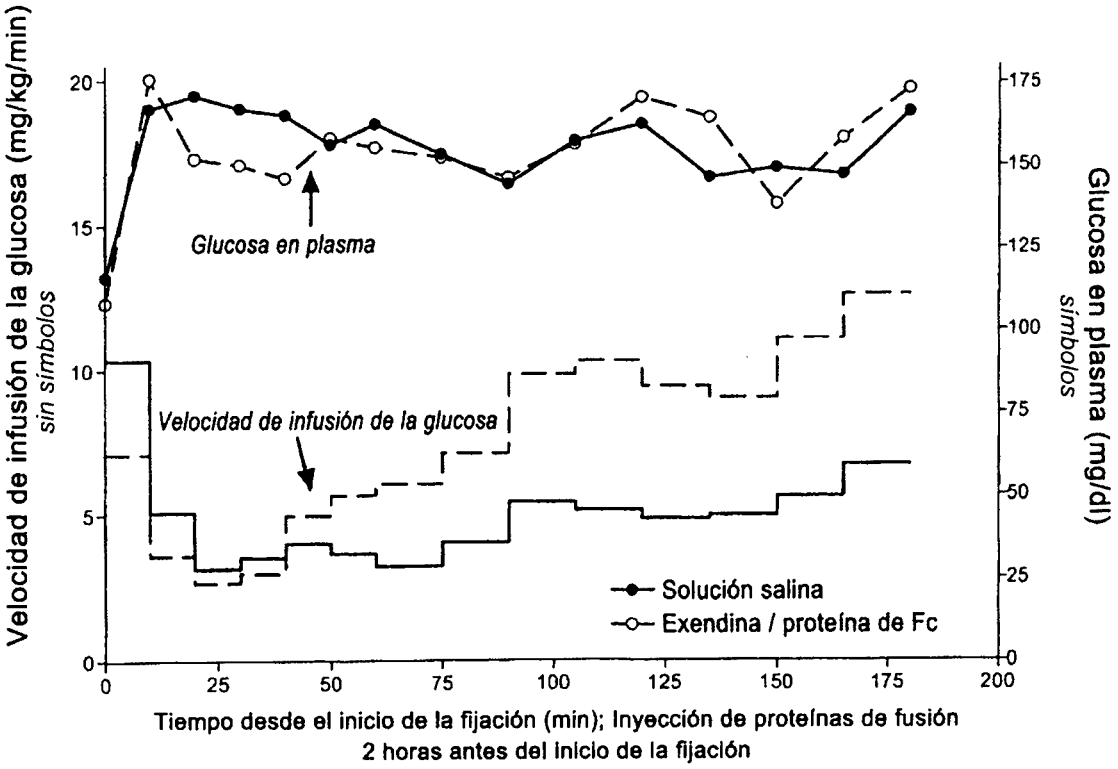

Velocidad de infusión de la glucosa (mg/kg/min)
sin símbolos
Glucosa en plasma (mg/dl)
símbolos
Glucosa en plasma
Velocidad de infusión de la glucosa
Solución salina
Exendina / proteína de Fc
Tiempo desde el inicio de la fijación (min); Inyección de proteínas de fusión
2 horas antes del inicio de la fijación

Fig. 10

B.

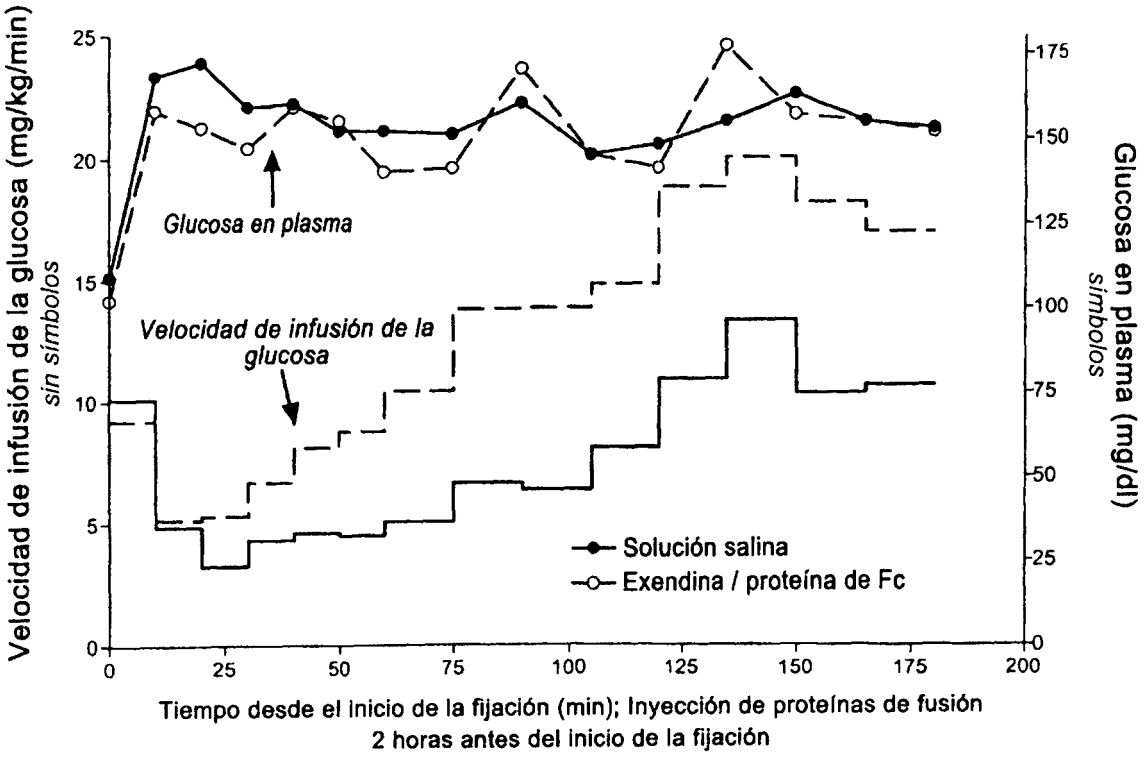

Velocidad de infusión de la glucosa (mg/kg/min)
sin símbolos
Glucosa en plasma (mg/dl)
símbolos
Glucosa en plasma
Velocidad de infusión de la glucosa
Solución salina
Exendina / proteína de Fc
Tiempo desde el inicio de la fijación (min); Inyección de proteínas de fusión
2 horas antes del inicio de la fijación

Fig. 11

A.

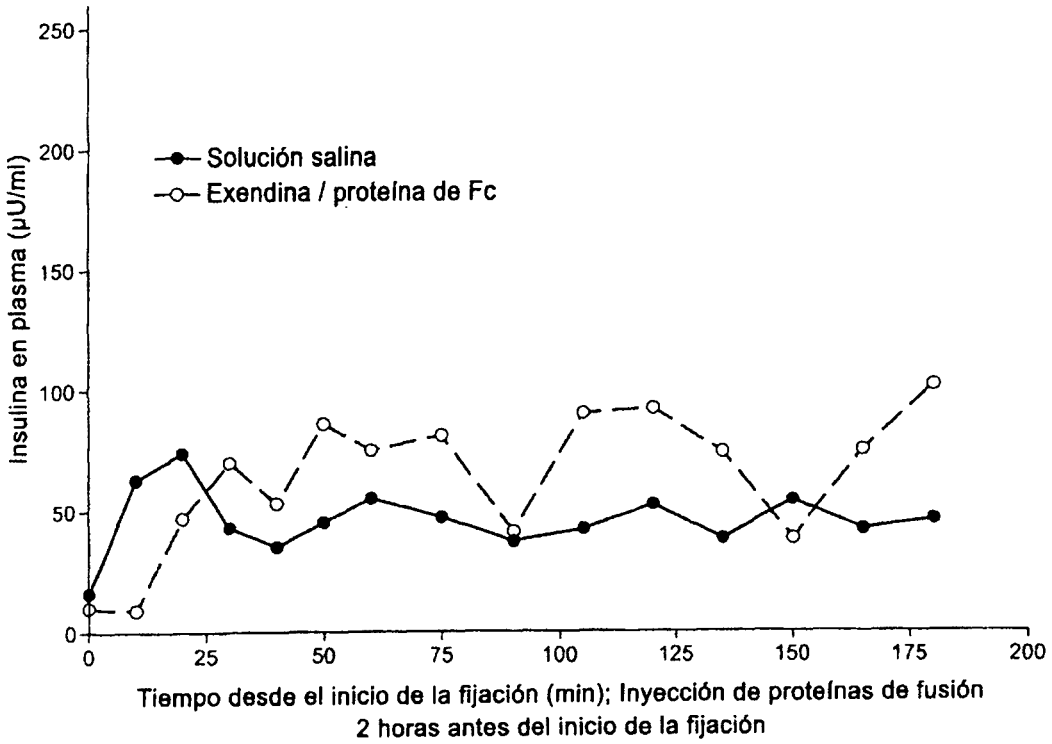

250
200
150
100
50
0
Insulina en plasma (µU/ml)
Solución salina
Exendina / proteína de Fc
0
25
50
75
100
125
150
175
200
Tiempo desde el inicio de la fijación (min); Inyección de proteínas de fusión
2 horas antes del inicio de la fijación

Fig. 11

B.

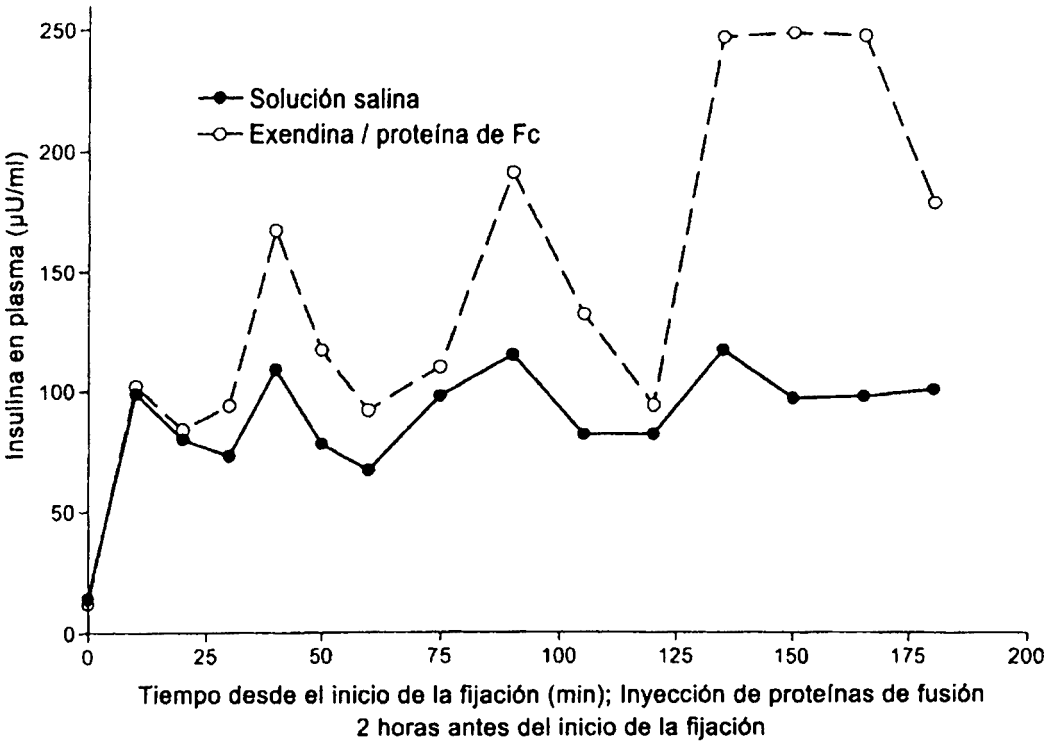

Solución salina
Exendina / proteína de Fc
Insulina en plasma (µU/ml)
250
200
150
100
50
0
0
25
50
75
100
125
150
175
200
Tiempo desde el inicio de la fijación (min); Inyección de proteínas de fusión
2 horas antes del inicio de la fijación

Fig. 12

1
GAGCCCAAATCTTGTGACAAAACTCACACATGCCCACCGTGCCCAGCACC
50
TGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCCAAAACCCAAGG
100
ACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGAC
150
GTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGT
200
GGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCA
250
CGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAAT
300
GGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCAT
350
CGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGT
400
ACACCCTGCCCCCATCCCGGGAGGAGATGACCAAGAACCAGGTCAGCCTG
450
ACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGA
500
GAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCCCGTGCTGG
550
ACTCCGACGGCTCCTTCTTCCTCTATAGCAAGCTCACCGTGGACAAGAGC
600
AGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCT
650
GCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAT
700
AGT [SEC ID N.º 33]

Fig. 13

1
GATGCGCACAAGAGTGAGGTTGCTCATCGGTTTAAAGATTTGGGAGAAGA
50
AAATTTCAAAGCCTTGGTGTTGATTGCCTTTGCTCAGTATCTTCAGCAGT
100
GTCCATTTGAAGATCATGTAAAATTAGTGAATGAAGTAACTGAATTTGCA
150
AAAACATGTGTTGCTGATGAGTCAGCTGAAAATTGTGACAAATCACTTCA
200
TACCCTTTTTGGAGACAAATTATGCACAGTTGCAACTCTTCGTGAAACCT
250
ATGGTGAAATGGCTGACTGCTGTGCAAAACAAGAACCTGAGAGAAATGAA
300
TGCTTCTTGCAACACAAAGATGACAACCCAAACCTCCCCCGATTGGTGAG
350
ACCAGAGGTTGATGTGATGTGCACTGCTTTTCATGACAATGAAGAGACAT
400
TTTTGAAAAAATACTTATATGAAATTGCCAGAAGACATCCTTACTTTTAT
450
GCCCCGGAACTCCTTTTCTTTGCTAAAAGGTATAAAGCTGCTTTTACAGA
500
ATGTTGCCAAGCTGCTGATAAAGCTGCCTGCCTGTTGCCAAAGCTCGATG
550
AACTTCGGGATGAAGGGAAGGCTTCGTCTGCCAAACAGAGACTCAAGTGT
600
GCCAGTCTCCAAAAATTTGGAGAAAGAGCTTTCAAAGCATGGGCAGTAGC
650
TCGCCTGAGCCAGAGATTTCCCAAAGCTGAGTTTGCAGAAGTTTCCAAGT
700
TAGTGACAGATCTTACCAAAGTCCACACGGAATGCTGCCATGGAGATCTG
750
CTTGAATGTGCTGATGACAGGGCGGACCTTGCCAAGTATATCTGTGAAAA
800
TCAAGATTCGATCTCCAGTAAACTGAAGGAATGCTGTGAAAAACCTCTGT
850
TGGAAAAATCCCACTGCATTGCCGAAGTGGAAAATGATGAGATGCCTGCT
900
GACTTGCCTTCATTAGCTGCTGATTTTGTTGAAAGTAAGGATGTTTGCAA
950
AAACTATGCTGAGGCAAAGGATGTCTTCCTGGGCATGTTTTTGTATGAAT
1000
ATGCAAGAAGGCATCCTGATTACTCTGTCGTGCTGCTGCTGAGACTTGCC

Continuación de la Fig. 13

1050
AAGACATATGAAACCACTCTAGAGAAGTGCTGTGCCGCTGCAGATCCTCA
1100
TGAATGCTATGCCAAAGTGTTCGATGAATTTAAACCTCTTGTGGAAGAGC
1150
CTCAGAATTTAATCAAACAAAATTGTGAGCTTTTTGAGCAGCTTGGAGAG
1200
TACAAATTCCAGAATGCGCTATTAGTTCGTTACACCAAGAAAGTACCCCA
1250
AGTGTCAACTCCAACTCTTGTAGAGGTCTCAAGAAACCTAGGAAAAGTGG
1300
GCAGCAAATGTTGTAAACATCCTGAAGCAAAAAGAATGCCCTGTGCAGAA
1350
GACTATCTATCCGTGGTCCTGAACCAGTTATGTGTGTTGCATGAGAAAAC
1400
GCCAGTAAGTGACAGAGTCACCAAATGCTGCACAGAATCCTTGGTGAACA
1450
GGCGACCATGCTTTTCAGCTCTGGAAGTCGATGAAACATACGTTCCCAAA
1500
GAGTTTAATGCTGAAACATTCACCTTCCATGCAGATATATGCACACTTTC
1550
TGAGAAGGAGAGACAAATCAAGAAACAAACTGCACTTGTTGAGCTCGTGA
1600
AACACAAGCCCAAGGCAACAAAAGAGCAACTGAAAGCTGTTATGGATGAT
1650
TTCGCAGCTTTTGTAGAGAAGTGCTGCAAGGCTGACGATAAGGAGACCTG
1700
CTTTGCCGAGGAGGGTAAAAAACTTGTTGCTGCAAGTCAAGCTGCCTTAG
1750
GCTTATAATGAC **[SEC ID N.º 35]**

LISTA DE SECUENCIAS

<110> Eli Lilly y Compañía

<120> PROTEÍNAS DE FUSIÓN GLP-1

<130> X13991A

<150> US 60/251.954
<151> 12-06-2000

<160> 35

<170> PatentIn 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> Constructo sintético

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> El Xaa de la posición 2 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp, o Lys;

<220>
<221> MISC_FEATURE
<222> (3)..(3)
<223> El Xaa de la posición 3 es Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (5)..(5)
<223> El Xaa de la posición 5 es Thr, Ala, Gly, Ser, Leu, Ile, Val, Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (8)..(8)

<223> El Xaa de la posición 8 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (10)..(10)
<223> El Xaa de la posición 10 es Val, Ala, Gly, Ser, Thr, Leu, Ile, Tyr, Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (11)..(11)
<223> El Xaa de la posición 11 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (12)..(12)
<223> El Xaa de la posición 12 es Ser, Ala, Gly, Thr, Leu, Ile, Val, Glu, Asp, Lys, Trp o Tyr;

<220>
<221> MISC_FEATURE
<222> (13)..(13)
<223> El Xaa de la posición 13 es Tyr, Phe, Trp, Glu, Asp, Gln o Lys;

<220>
<221> MISC_FEATURE
<222> (14)..(14)
<223> El Xaa de la posición 14 es Leu, Ala, Gly, Ser, Thr, Ile, Val, Glu, Asp, Met, Lys, Trp o Tyr;

<220>
<221> MISC_FEATURE
<222> (15)..(15)
<223> El Xaa de la posición 15 es Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (16)..(16)
<223> El Xaa de la posición 16 es Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (17)..(17)
<223> El Xaa de la posición 17 es Gln, Asn, Arg, Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (18)..(18)
<223> El Xaa de la posición 18 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Arg, Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (19)..(19)
<223> El Xaa de la posición 19 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (20)..(20)
<223> El Xaa de la posición 20 es Lys, Arg, Gln, Glu, Asp o His;

<220>
<221> MISC_FEATURE
<222> (21)..(21)
<223> El Xaa de la posición 21 es Leu, Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (24)..(24)
<223> El Xaa de la posición 24 es Ala, Gly, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (25)..(25)
<223> El Xaa de la posición 25 es Trp, Phe, Tyr, Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (26)..(26)
<223> El Xaa de la posición 26 es Leu, Gly, Ala, Ser, Thr, Ile, Val, Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (27)..(27)
<223> El Xaa de la posición 27 es Val, Gly, Ala, Ser, Thr, Leu, Ile, Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (28)..(28)
<223> El Xaa de la posición 28 es Asn, Lys, Arg, Glu, Asp o His;

<220>
<221> MISC_FEATURE
<222> (29)..(29)
<223> El Xaa de la posición 29 es Gly, Ala, Ser, Thr, Leu, Ile, Val, Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (30)..(30)
<223> El Xaa de la posición 30 es Gly, Arg, Lys, Glu, Asp o His;

<220>
<221> MISC_FEATURE
<222> (31)..(31)
<223> El Xaa de la posición 31 es Pro, Gly, Ala, Ser, Thr, Leu, Ile, val, Glu, Asp o Lys, o está eliminado;

<220> —

<221> MISC_FEATURE
<222> (32)..(32)
<223> El Xaa de la posición 32 es Ser, Arg, Lys, Glu, Asp o His, o está eliminado;

<220>
<221> MISC_FEATURE
<222> (33)..(33)
<223> El Xaa de la posición 33 es Ser, Arg, Lys, Glu, Asp o His, o está eliminado;

<220>
<221> MISC_FEATURE
<222> (34)..(34)
<223> El Xaa de la posición 34 es Gly, Asp, Glu o Lys, o está eliminado;

<220>
<221> MISC_FEATURE
<222> (35)..(35)
<223> El Xaa de la posición 35 es Ala, Phe, Trp, Tyr, Glu, Asp o Lys, o está eliminado;

<220>
<221> MISC_FEATURE
<222> (36)..(36)
<223> El Xaa de la posición 36 es Ser, Pro, Lys, Glu o Asp, o está eliminado;

<220>
<221> MISC_FEATURE
<222> (37)..(37)
<223> El Xaa de la posición 37 es Ser, Pro, Glu, Asp o Lys, o está eliminado;

<220>
<221> MISC_FEATURE
<222> (38)..(38)
<223> El Xaa de la posición 38 es Gly, Pro, Glu, Asp o Lys, o está eliminado;

<220>
<221> MISC_FEATURE
<222> (39)..(39)
<223> El Xaa de la posición 39 es Ala, Ser, Val, Glu, Asp o Lys, o está eliminado;

<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> Constructo sintético

<220>
<221> MISC_FEATURE
<222> (1)..(1)
<223> El Xaa de la posición 1 es L-histidina, D-histidina o está eliminado;

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> El Xaa de la posición 2 es Gly, Ala, Val, Leu, Ile, Ser o Thr;

<220>
<221> MISC_FEATURE
<222> (3)..(3)
<223> El Xaa de la posición 3 es Thr, Ser, Arg, Lys, Trp, Phe, Tyr, Glu o His;

<220>
<221> MISC_FEATURE
<222> (5)..(5)
<223> El Xaa de la posición 5 es Asp, Glu, Arg, Thr, Ala, Lys o His;

<220>
<221> MISC_FEATURE
<222> (6)..(6)
<223> El Xaa de la posición 6 es His, Trp, Phe o Tyr;

<220>
<221> MISC_FEATURE
<222> (10)..(10)
<223> El Xaa de la posición 10 es Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Tyr, Glu o Ala;

<220>
<221> MISC_FEATURE
<222> (12)..(12)
<223> El Xaa de la posición 12 es His, Pro, Asp, Glu, Arg, Ser, Ala o Lys;

<220>
<221> MISC_FEATURE
<222> (16)..(16)
<223> El Xaa de la posición 13 es Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;

<220>
<221> MISC_FEATURE
<222> (17)..(17)
<223> El Xaa de la posición 17 es His, Asp, Lys, Glu, Gln o Arg;

<220>
<221> MISC_FEATURE

<222> (18)..(18)
<223> El Xaa de la posición 18 es Glu, Arg, Ala o Lys;

<220>
<221> MISC_FEATURE
<222> (20)..(20)
<223> El Xaa de la posición 20 es Trp, Tyr, Phe, Asp, Lys, Glu o His;

<220>
<221> MISC_FEATURE
<222> (21)..(21)
<223> El Xaa de la posición 21 es Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;

<220>
<221> MISC_FEATURE
<222> (24)..(24)
<223> El Xaa de la posición 24 es Ala, Glu, Asp, Ser o His;

<220>
<221> MISC_FEATURE
<222> (25)..(25)
<223> El Xaa de la posición 25 es Asp, Glu, Ser, Thr, Arg, Trp o Lys;

<220>
<221> MISC_FEATURE
<222> (27)..(27)
<223> El Xaa de la posición 27 es Asp, Arg, Val, Lys, Ala, Gly o Glu;

<220>
<221> MISC_FEATURE
<222> (28)..(28)
<223> El Xaa de la posición 28 es Glu, Lys o Asp;

<220>
<221> MISC_FEATURE
<222> (29)..(29)
<223> El Xaa de la posición 29 es Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o Glu;

<220>
<221> MISC_FEATURE
<222> (30)..(30)
<223> El Xaa de la posición 30 es Thr, Ser, Asp, Trp, Tyr, Phe, Arg, Glu o His;

<220>
<221> MISC_FEATURE
<222> (31)..(31)
<223> El Xaa de la posición 31 es Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, o está eliminado;

<220>
<221> MISC_FEATURE
<222> (32)..(32)

<223> El Xaa de la posición 31 es Pro, o está eliminado;

<400> 3

Xaa Xaa Xaa Gly Xaa Xaa Thr Ser Asp Xaa Ser Xaa Tyr Leu Glu Xaa
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> Constructo sintético

<220>
<221> MISC_FEATURE
<222> (1)..(1)
<223> El Xaa de la posición 1 es L-histidina, D-histidina o está eliminado;

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> El Xaa de la posición 2 es Gly, Ala, Val, Leu, Ile, Ser o Thr;

<220>
<221> MISC_FEATURE
<222> (5)..(5)
<223> El Xaa de la posición 5 es Asp, Glu, Arg, Thr, Ala, Lys o His;

<220>
<221> MISC_FEATURE
<222> (6)..(6)
<223> El Xaa de la posición 6 es His, Trp, Phe o Tyr;

<220>
<221> MISC_FEATURE
<222> (10)..(10)
<223> El Xaa de la posición 10 es Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Glu o Ala;

<220>
<221> MISC_FEATURE
<222> (16)..(16)
<223> El Xaa de la posición 16 es Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;

<220>
<221> MISC_FEATURE
<222> (17)..(17)
<223> El Xaa de la posición 17 es His, Asp, Lys, Glu o Gln;

<220>
<221> MISC_FEATURE
<222> (18)..(18)
<223> El Xaa de la posición 18 es Glu, His, Ala o Lys;

<220>
<221> MISC_FEATURE
<222> (19)..(19)
<223> El Xaa de la posición 19 es Asp, Lys, Glu o His;

<220>
<221> MISC_FEATURE
<222> (21)..(21)
<223> El Xaa de la posición 21 es Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;

<220>
<221> MISC_FEATURE
<222> (24)..(24)
<223> El Xaa de la posición 24 es Ala, Glu, Asp, Ser o His;

<220>
<221> MISC_FEATURE
<222> (27)..(27)
<223> El Xaa de la posición 27 es Asp, Arg, Val, Lys, Ala, Gly o Glu;

<220>
<221> MISC_FEATURE
<222> (28)..(28)
<223> El Xaa de la posición 28 es Glu, Lys o Asp;

<220>
<221> MISC_FEATURE
<222> (29)..(29)
<223> El Xaa de la posición 29 es Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o Glu;

<220>
<221> MISC_FEATURE
<222> (30)..(30)
<223> El Xaa de la posición 30 es Arg, Glu o His;

<220>
<221> MISC_FEATURE
<222> (31)..(31)
<223> El Xaa de la posición 31 es Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe His, Gly, o está eliminado;

<220>
<221> MISC_FEATURE
<222> (32)..(32)
<223> El Xaa de la posición 32 es Pro, o está eliminado;

<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> Constructo sintético

<220>
<221> MISC_FEATURE
<222> (1)..(1)
<223> El Xaa de la posición 1 es L-histidina, D-histidina o está eliminado;

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> El Xaa de la posición 2 es Gly, Ala, Val, Leu, Ile, Ser, Met o Thr;

<220>
<221> MISC_FEATURE
<222> (6)..(6)
<223> El Xaa de la posición 6 es His, Trp, Phe o Tyr;

<220>
<221> MISC_FEATURE
<222> (10)..(10)
<223> El Xaa de la posición 10 es Leu, Ser, Thr, Trp, His, Phe, Asp, Val, Tyr, Glu o Ala;

<220>
<221> MISC_FEATURE
<222> (16)..(16)
<223> El Xaa de la posición 16 es Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;

<220>
<221> MISC_FEATURE
<222> (17)..(17)
<223> El Xaa de la posición 17 es His, Asp, Lys, Glu o Gln;

<220>
<221> MISC_FEATURE
<222> (20)..(20)
<223> El Xaa de la posición 20 es Asp, Lys, Glu o His;

<220>
<221> MISC_FEATURE

<222> (24)..(24)
<223> El Xaa de la posición 24 es Ala, Glu, Asp, Ser o His;

<220>
<221> MISC_FEATURE
<222> (29)..(29)
<223> El Xaa de la posición 29 es Thr, Ser, Lys, Arg, Trp, Tyr, Phe, Asp, Gly, Pro, His o Glu;

<220>
<221> MISC_FEATURE
<222> (31)..(31)
<223> El Xaa de la posición 31 es Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, o está eliminado;

<220>
<221> MISC_FEATURE
<222> (32)..(32)
<223> El Xaa de la posición 32 es Pro, o está eliminado;

<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> Secuencia artificial

<220>
<223> Constructo sintético
<220>

<221> MISC_FEATURE
<222> (1)..(1)
<223> El Xaa de la posición 1 es L-histidina, D-histidina o está eliminado;

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> El Xaa de la posición 2 es Gly, Ala, Val, Leu, Ile, Ser o Thr;

<220>
<221> MISC_FEATURE
<222> (16)..(16)
<223> El Xaa de la posición 16 es Gly, Asp, Glu, Gln, Asn, Lys, Arg o Cys;

<220>
<221> MISC_FEATURE
<222> (17)..(17)
<223> El Xaa de la posición 17 es His, Asp, Lys, Glu o Gln;

<220>
<221> MISC_FEATURE
<222> (18)..(18)
<223> El Xaa de la posición 18 es Ala, Glu, His, Phe, Tyr, Trp, Arg o Lys;

<220>
<221> MISC_FEATURE
<222> (24)..(24)
<223> El Xaa de la posición 24 es Ala, Glu, Asp, Ser o His;

<220>
<221> MISC_FEATURE
<222> (31)..(31)
<223> El Xaa de la posición 31 es Lys, Arg, Thr, Ser, Glu, Asp, Trp, Tyr, Phe, His, Gly, Gly-Pro, o está eliminado;

<220>
<221> MISC_FEATURE
<222> (32)..(32)
<223> El Xaa de la posición 32 es Pro, o está eliminado;

<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> Constructo sintético

<220>
<221> MISC_FEATURE
<222> (1)..(1)
<223> El Xaa de la posición 1 es L-histidina, D-histidina o está eliminado;

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> El Xaa de la posición 2 es Ala, Gly, Val, Thr, Ile y alfa-metil-Ala;

<220>
<221> MISC_FEATURE
<222> (15)..(15)
<223> El Xaa de la posición 15 es Glu, Gln, Ala, Thr, Ser y Gly;

<220>
<221> MISC_FEATURE

<222> (21)..(21)
<223> El Xaa de 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> Constructo sintético

<220>
<221> MISC_FEATURE
<222> (19)..(19)
<223> El Xaa de la posición 19 es Lys o Arg;

<220>
<221> MOD_RES
<222> (27)..(27)
<223> ACETILACIÓN

<220>
<221> MISC_FEATURE
<222> (30)..(30)
<223> El Xaa de 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 Glu Gly Gln
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> Constructo sintético

<400> 9

```
His Ser Asp 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> 10
<211> 39
<212> PRT
<213> Secuencia artificial

<220>
<223> Constructo sintético

<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> Constructo sintético

<220>
<221> MISC_FEATURE
<222> (1)..(1)
<223> El Xaa de la posición 1 es L-histidina, D-histidina o está eliminado;

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> El Xaa de la posición 2 es Gly, Ala o Val;

<220>
<221> MISC_FEATURE
<222> (10)..(10)
<223> El Xaa de la posición 10 es Leu o Val;

<220>
<221> MISC_FEATURE

<222> (12)..(12)
<223> El Xaa de la posición 12 es Lys o Ser;

<220>
<221> MISC_FEATURE
<222> (13)..(13)
<223> El Xaa de la posición 13 es Gln o Tyr;

<220>
<221> MISC_FEATURE
<222> (14)..(14)
<223> El Xaa de la posición 14 es Met o Leu;

<220>
<221> MISC_FEATURE
<222> (16)..(16)
<223> El Xaa de la posición 16 es Glu o Gln;

<220>
<221> MISC_FEATURE
<222> (17)..(17)
<223> El Xaa de la posición 17 es Glu o Gln;

<220>
<221> MISC_FEATURE
<222> (19)..(19)
<223> El Xaa de la posición 19 es Val o Ala;

<220>
<221> MISC_FEATURE
<222> (20)..(20)
<223> El Xaa de la posición 20 es Arg o Lys;

<220>
<221> MISC_FEATURE
<222> (21)..(21)
<223> El Xaa de la posición 21 es Leu o Glu;

<220>
<221> MISC_FEATURE
<222> (24)..(24)
<223> El Xaa de la posición 24 es Glu o Ala;

<220>
<221> MISC_FEATURE
<222> (27)..(27)
<223> El Xaa de la posición 27 es Val o Lys;

<220>
<221> MISC_FEATURE
<222> (28)..(28)

<223> El Xaa de la posición 28 es Asn o Lys;

<220>
<221> MISC_FEATURE
<222> (30)..(30)
<223> El Xaa de la posición 30 es Gly o Arg; y

<220>
<221> MISC_FEATURE
<222> (31)..(31)
<223> El Xaa de la posición 31 es Gly o Pro;

<220>
<221> MISC_FEATURE
<222> (32)..(32)
<223> El Xaa de la posición 32 es Ser o está ausente;

<220>
<221> MISC_FEATURE
<222> (33)..(33)
<223> El Xaa de la posición 33 es Ser o está ausente;

<220>
<221> MISC_FEATURE
<222> (34)..(34)
<223> El Xaa de la posición 34 es Gly o está ausente;

<220>
<221> MISC_FEATURE
<222> (35)..(35)
<223> El Xaa de la posición 35 es Ala o está ausente;

<220>
<221> MISC_FEATURE
<222> (36)..(36)
<223> El Xaa de la posición 36 es Pro o está ausente;

<220>
<221> MISC_FEATURE
<222> (37)..(37)
<223> El Xaa de la posición 37 es Pro o está ausente;

<220>
<221> MISC_FEATURE
<222> (38)..(38)
<223> El Xaa de la posición 38 es Pro o está ausente;

<220>
<221> MISC_FEATURE
<222> (39)..(39)
<223> El Xaa de la posición 39 es Pro o está ausente;

<220>
<221> MISC_FEATURE
<222> (39)..(39)
<223> El Xaa de 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
35

<210> 12
<211> 31
<212> PRT
<213> Secuencia artificial

<220>
<223> Constructo sintético

<220>
<221> MISC_FEATURE
<222> (1)..(1)
<223> El Xaa de la posición 1 es L-histidina, D-histidina o está eliminado;

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> El Xaa de la posición 2 es Ala, Gly, Val, Leu, Ile, Ser o Thr;

<220>
<221> MISC_FEATURE
<222> (6)..(6)
<223> El Xaa de la posición 6 es Phe, Trp o Tyr;

<220>
<221> MISC_FEATURE
<222> (10)..(10)
<223> El Xaa de la posición 10 es Val, Trp, Ile, Leu, Phe o Tyr;

<220>
<221> MISC_FEATURE
<222> (12)..(12)
<223> El Xaa de la posición 12 es Ser, Trp, Tyr, Phe, Lys, Ile, Leu, Val;

<220>
<221> MISC_FEATURE
<222> (13)..(13)
<223> El Xaa de la posición 13 es Tyr, Trp o Phe;

<220>
<221> MISC_FEATURE
<222> (14)..(14)
<223> El Xaa de la posición 14 es Leu, Phe, Tyr o Trp;

<220>
<221> MISC_FEATURE
<222> (16)..(16)
<223> El Xaa de la posición 16 es Gly, Glu, Asp o Lys;

<220>
<221> MISC_FEATURE
<222> (19)..(19)
<223> El Xaa de la posición 19 es Ala, Val, Ile o Leu;

<220>
<221> MISC_FEATURE
<222> (21)..(21)
<223> El Xaa de la posición 21 es Glu, Ile o Ala;

<220>
<221> MISC_FEATURE
<222> (24)..(24)
<223> El Xaa de la posición 24 es Ala o Glu;

<220>
<221> MISC_FEATURE
<222> (27)..(27)
<223> El Xaa de la posición 27 es Val o Ile; y

<220>
<221> MOD_RES
<222> (30)..(30)
<223> AMIDACIÓN

<220>
<221> MISC_FEATURE
<222> (31)..(31)
<223> El Xaa de 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> Constructo sintético

<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> Constructo sintético

<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> Constructo sintético

<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

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> Constructo sintético

<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

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> Constructo sintético

<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> Constructo sintético

<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> Constructo sintético

<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> Constructo sintético

<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> Constructo sintético

<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> Constructo sintético

<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
```

<210> 23

<211> 287

<212> PRT

<213> Secuencia artificial

<220>

<223> Constructo sintético

<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> Constructo sintético

<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> Constructo sintético

<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> Constructo sintético

<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> Constructo sintético

<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> Constructo sintético

<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

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> PRT

<213> Secuencia artificial

<220>

<223> Constructo sintético

<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> Constructo sintético

<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> PRT

<213> Secuencia artificial

<220>

<223> Constructo sintético

<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> ADN

<213> *Homo sapiens*

<400> 33

gagcccaaat cttgtgacaa aactcacaca tgcccaccgt gcccagcacc tgaactcctg 60

gggggaccgt cagtcttcct cttcccccca aaacccaagg acaccctcat gatctcccgg 120

acccctgagg tcacatgcgt ggtggtggac gtgagccacg aagaccctga ggtcaagttc 180

aactggtacg tggacggcgt ggaggtgcat aatgccaaga caaagccgcg ggaggagcag 240

tacaacagca cgtaccgtgt ggtcagcgtc ctcaccgtcc tgcaccagga ctggctgaat 300

ggcaaggagt acaagtgcaa ggtctccaac aaagccctcc cagccccat cgagaaaacc 360

atctccaaag ccaaagggca gccccgagaa ccacaggtgt acaccctgcc cccatcccgg 420

gaggagatga ccaagaacca ggtcagcctg acctgcctgg tcaaaggctt ctatcccagc 480

gacatcgccg tggagtggga gagcaatggg cagccggaga acaactacaa gaccacgcct 540

cccgtgctgg actccgacgg ctccttcttc ctctatagca agctcaccgt ggacaagagc 600

aggtggcagc aggggaacgt 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> ADN
<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 aacctccccc 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 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
gccaagtata tctgtgaaaa tcaagattcg atctccagta aactgaagga atgctgtgaa 840
aaacctctgt tggaaaaatc ccactgcatt gccgaagtgg aaaatgatga gatgcctgct 900

gacttgcctt cattagctgc tgattttgtt gaaagtaagg atgtttgcaa aaactatgct 960

gaggcaaagg atgtcttcct gggcatgttt ttgtatgaat atgcaagaag gcatcctgat 1020

tactctgtcg tgctgctgct gagacttgcc aagacatatg aaaccactct agagaagtgc 1080

tgtgccgctg cagatcctca tgaatgctat gccaaagtgt tcgatgaatt taaacctctt 1140

gtggaagagc ctcagaattt aatcaaacaa aattgtgagc tttttgagca gcttggagag 1200

tacaaattcc agaatgcgct attagttcgt tacaccaaga aagtacccca agtgtcaact 1260

ccaactcttg tagaggtctc aagaaaccta ggaaaagtgg gcagcaaatg ttgtaaacat 1320

cctgaagcaa aaagaatgcc ctgtgcagaa gactatctat ccgtggtcct gaaccagtta 1380

tgtgtgttgc atgagaaaac gccagtaagt gacagagtca ccaaatgctg cacagaatcc 1440

ttggtgaaca ggcgaccatg cttttcagct ctggaagtcg atgaaacata cgttcccaaa 1500

gagtttaatg ctgaaacatt caccttccat gcagatatat gcacactttc tgagaaggag 1560

agacaaatca agaaacaaac tgcacttgtt gagctcgtga aacacaagcc caaggcaaca 1620

aaagagcaac tgaaagctgt tatggatgat ttcgcagctt ttgtagagaa gtgctgcaag 1680

gctgacgata aggagacctg ctttgccgag gagggtaaaa aacttgttgc tgcaagtcaa 1740

gctgccttag gcttataatg ac 1762