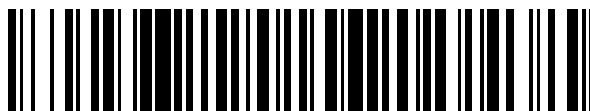


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 725 492**

51 Int. Cl.:

C07K 1/10 (2006.01)

C07K 14/605 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **22.12.2008 PCT/EP2008/068211**

87 Fecha y número de publicación internacional: **09.07.2009 WO09083549**

96 Fecha de presentación y número de la solicitud europea: **22.12.2008 E 08868214 (1)**

97 Fecha y número de publicación de la concesión europea: **20.02.2019 EP 2238153**

54 Título: **Preparación semirrecombinante de análogos del GLP-1**

30 Prioridad:

28.12.2007 EP 07124141

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

24.09.2019

73 Titular/es:

NOVO NORDISK A/S (100.0%)

Novo Allé

2880 Bagsværd, DK

72 Inventor/es:

**LAU, JESPER FÆRGEMAN;
ANDERSEN, ASSER SLOTH;
BLOCH, PAW;
LAU, JESPER;
GARIBAY, PATRICK WILLIAM;
KRUSE, THOMAS;
NØRBY, INGA SIG NIELSEN;
JESSEN, CLAUD ULRICH;
CHRISTENSEN, CASPAR y
NORRILD, JENS CHRISTIAN**

74 Agente/Representante:

PONS ARIÑO, Ángel

ES 2 725 492 T3

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).

DESCRIPCIÓN

Preparación semirrecombinante de análogos del GLP-1

5 Campo de la invención

La presente invención se refiere a un método semirrecombinante para la preparación de análogos del GLP-1 y derivados que contienen aminoácidos no proteinogénicos en la parte N-terminal que combina el uso de técnicas de expresión recombinante y síntesis química de péptidos.

10

Antecedentes de la invención

El péptido 1 similar a glucagón humano (GLP-1) es una hormona de 37 residuos de aminoácidos que se encuentra implicado en la regulación del metabolismo de la glucosa en sangre, secreción gastrointestinal y el metabolismo e ingesta de alimentos. El GLP-1 se origina a partir del preproglucagón que se sintetiza *entre otras* en las células L en el íleon distal, en el páncreas y en el cerebro. El GLP-1 estimula la secreción de insulina de una manera dependiente de glucosa, estimula la biosíntesis de insulina, promueve el rescate de células beta, disminuye la secreción de glucagón, el vaciado gástrico y la ingesta de alimentos. El GLP-1 humano se hidroliza a GLP-1 (7-37) y GLP-1 (7-36)-amida que son ambos agentes insulino-trópicos. Se ha demostrado que las dosis farmacológicas de GLP-1 administradas a pacientes diabéticos tipo 2 aumentan significativamente los niveles de insulina circulante y reducen los niveles en plasma de glucagón. Las acciones de GLP-1 están mediadas por los receptores de GLP-1 en el páncreas, corazón, riñón, sistema nervioso central y tracto gastrointestinal. Por lo tanto, se espera que el GLP-1 se vuelva muy importante en el tratamiento de la diabetes.

25 Sin embargo, el GLP-1 humano nativo se inactiva rápidamente por la degradación por la enzima plasmática dipeptidil peptidasa IV (DPP-IV) a un metabolito de GLP-1(9–36)amida truncado, que sirve como un antagonista del receptor de GLP-1. Esto le confiere al péptido una vida media corta en circulación.

30 Esta vida media corta en circulación es un problema para muchos pacientes diabéticos particularmente en el segmento de diabetes tipo 2 que están sujetos a la llamada "fobia a las agujas", es decir, un temor sustancial a inyectarse ellos mismos. En el segmento de diabetes tipo 2, la mayoría de los pacientes se tratan con agentes hipoglucémicos orales. Dado que se espera que los compuestos de GLP-1 sean un producto farmacéutico inyectable, el temor a las inyecciones puede convertirse en un obstáculo serio para el uso amplio de estos compuestos clínicamente muy prometedores.

35 Por lo tanto, se ha usado una variedad de enfoques y métodos diferentes para modificar la estructura de los compuestos de GLP-1 para proporcionar una mayor duración de acción *in vivo*. Por lo tanto, se realiza un esfuerzo considerable por ejemplo para desarrollar análogos y derivados de compuestos de GLP-1 menos susceptibles a la hidrólisis mediada por DPP-IV con el fin de reducir la velocidad de degradación por DPP-IV. Los documentos núms. 40 WO 2006/097538, WO 2006/097536, WO 2006/037810, WO2006/005667, WO2005/058958, WO 2005/027978, WO 98/08871 y US 2001/0011071 describen varios análogos y derivados del GLP-1, que incluyen análogos de GLP que comprenden aminoácidos no proteinogénicos (es decir, aminoácidos no naturales) que pueden conferir una cierta protección contra la hidrólisis por DPP-IV.

45 Los polipéptidos que contienen solo aminoácidos proteinogénicos (es decir, aminoácidos naturales) tal como GLP-1 nativo, pueden producirse mediante el uso de técnicas recombinantes o mediante síntesis química. Sin embargo, los polipéptidos que contienen, además, aminoácidos no proteinogénicos tales como los análogos del GLP-1 extendidos en el extremo N-terminal, no pueden prepararse actualmente mediante las técnicas de expresión recombinante de una manera práctica y se preparan en general mediante síntesis química. El método más ampliamente usado para la 50 síntesis de péptidos es la síntesis de péptidos en fase sólida donde los aminoácidos adecuados protegidos se incorporan de manera gradual mediante el uso de un polímero como un soporte sólido.

La síntesis de péptidos en fase sólida (SPPS) puede ser muy eficiente en la preparación de algunos péptidos, pero el uso de aminoácidos protegidos en combinación con el uso constante de exceso de reactivos hace que este enfoque sea relativamente costoso. Además, cada prolongación de aminoácidos en una síntesis de polipéptidos en fase sólida requiere un procedimiento de lavado completo. Típicamente, la incorporación de un aminoácido implica hasta 10 etapas de lavado con solventes como NMP, DMF o DCM.

60 Cuando los polipéptidos, tales como agentes insulino-trópicos, análogos del GLP-1, análogos truncados del GLP-1 y derivados del GLP-1 se sintetizan mediante el uso de SPPS, la formación de estructuras secundarias durante la síntesis frecuentemente conduce a una menor eficiencia de las etapas individuales de síntesis. Como una consecuencia, los péptidos más grandes o los péptidos que contienen ciertas secuencias de aminoácidos frecuentemente se producen con baja pureza y rendimientos. Las impurezas son frecuentemente péptidos de eliminación donde faltan uno o más aminoácidos en la secuencia final. Estas impurezas pueden ser muy difíciles de 65 separar del péptido deseado, y resulta en un producto contaminado con péptidos de eliminación.

Es el objetivo de la presente invención proporcionar un método eficiente y económico para la preparación de análogos del GLP-1 y derivados que estén protegidos de DPP-IV por tener aminoácidos no proteínogénicos en la parte N-terminal. El método combina la ventaja de la producción rentable de moléculas precursoras del GLP-1 truncadas, mediante el uso de técnicas recombinantes junto con la síntesis química de extensiones N-terminales que comprenden aminoácidos no proteínogénicos. Se logra una reducción significativa del costo de producción de los análogos o derivados del GLP-1. Los análogos y derivados del GLP-1 menos costosos son altamente convenientes para maximizar el número de pacientes cuyo tratamiento está disponible, así como también para explotar las ventajas de rutas alternativas de suministro que tienen menor biodisponibilidad que la inyección subcutánea, por ejemplo, suministro transdérmico y pulmonar.

Breve descripción de la invención

La presente invención se refiere a un método para producir un derivado de un análogo del GLP-1 que comprende uno o más aminoácidos no proteínogénicos en la parte N-terminal, dicho método que comprende las etapas de:

(i) cultivar una célula huésped que comprende una secuencia de nucleótidos que codifica una molécula precursora de dicho análogo del GLP-1 en condiciones adecuadas para la expresión de dicha molécula precursora;

(ii) separar la molécula precursora expresada del caldo de cultivo;

(iii) acilar el grupo épsilon amino de un residuo de lisina en la molécula precursora expresada con un agente de acilación, que se activa opcionalmente, para obtener un derivado de la molécula precursora, y opcionalmente aislar dicho derivado de la molécula precursora;

(iv) acoplar una extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteínogénicos al derivado de la molécula precursora expresado; y

(v) aislar el derivado resultante de un análogo del GLP-1 por medios adecuados;

en donde

(a) dicha etapa de acilación (iii) tiene lugar en una mezcla de solventes acuosos con un pH entre 9 y 13;

(b) dicha extensión N-terminal tiene una longitud de 2 o 4 aminoácidos, y

(c) dicho derivado resultante de un análogo del GLP-1 es el péptido N-épsilon26-[2-(2-[2-(2-[2-((S)-4-Carboxi-4-(17-carboxi-heptadecanoilamino)-butirilamino]etoxi)etoxi)acetilamino]-etoxi)etoxi)-acetil][Aib8,Arg34] GLP-1-(7-37)

Breve descripción de los dibujos

La Figura 1 muestra el plásmido de levadura Psa273. El plásmido contiene un casete de expresión que comprende un fragmento *EcoRI-XbaI* insertado en el plásmido entre el promotor de la transcripción y el terminador de la transcripción del gen TPI de *S. cerevisiae*. El casete de expresión codifica un producto de fusión compuesto por el líder pre-pro MF α 1*, un sitio de escisión Lys-Arg para el procesamiento dibásico de la endopeptidasa Kex2, y el precursor del GLP-1 (R34)GLP-1(11-37).

La Figura 2 muestra los fragmentos de ADN *NcoI-XbaI* que codifican los precursores del GLP-1. La secuencia de aminoácidos del precursor del GLP-1 codificada por los fragmentos de ADN se indica en negritas. a) (R34)GLP-1(11-37); b) Ext1-(E22,R26,R34)GLP-1(8-37)K38-Ext2 y c) Ext1-(E22,R26,R34,K37)GLP-1(8-37)-Ext2; d)(R34)GLP-1(9-37).

La Figura 3 muestra los fragmentos de ADN *NcoI-XbaI* que codifican los precursores del GLP-1 y las secuencias líderes optimizadas. (34R)GLP-1(11-37) se muestra como un ejemplo no limitante. Las letras mayúsculas indican las diferentes secuencias. Las secuencias de aminoácidos optimizadas codificadas por los fragmentos de ADN se indican en negritas. 'KR' indica un constructo con el sitio de escisión LysArg mínimo Kex2p.

La Figura 4 muestra la optimización del líder (34R)GLP-1(11-37) como un ejemplo. Optimización de P1 a P6: (A-F). Incorporación de prolina: (H, I). Incorporación de aminoácidos cargados en P7 a P11: (M-Z, K5-7, L2-4). Constructos para el procesamiento corriente abajo con DAPasa: (G, J, K, L). Los rendimientos se dan en porcentaje con relación a un (34R)GLP-1(11-37) con el líder MF α 1* y un sitio simple de escisión LysArg Kex2p.

Descripción detallada de la invención

En su aspecto más amplio la presente invención describe un método para producir análogos del GLP-1 y derivados que son DPP-IV protegidos al comprender uno o más aminoácidos no proteínogénicos en la parte N-terminal.

Por lo tanto, en un aspecto la presente invención describe un método para fabricar análogos del GLP-1 y derivados que comprenden uno o más aminoácidos no proteinogénicos, dicho método comprende las siguientes etapas:

(i) cultivar una célula huésped que comprende una secuencia de nucleótidos que codifica una molécula precursora de dicho análogo o derivado del GLP-1 en condiciones adecuadas para la expresión de dicha molécula precursora,

(ii) separar la molécula precursora expresada del caldo de cultivo,

(iii) acoplar una extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos a la molécula precursora expresada,

(iv) aislar el análogo o derivado del GLP-1 resultante por medios adecuados, como se conoce en la técnica.

Los presentes inventores han desarrollado un método económico y eficiente para preparar análogos del GLP-1 y derivados que comprenden uno o más aminoácidos no proteinogénicos en la parte N-terminal. Por lo tanto, se ha descubierto que las moléculas precursoras del GLP-1 producidas de manera recombinante que contienen solamente aminoácidos proteinogénicos pueden extenderse eficientemente en el extremo N-terminal mediante el uso de un aminoácido no proteinogénico o un péptido que contiene uno o más aminoácidos no proteinogénicos, por ejemplo, mediante la reacción del precursor del análogo del GLP-1 con un derivado de ácido carboxílico, que está opcionalmente activado, del fragmento de extensión.

Más específicamente, la presente invención proporciona un método semirrecombinante para la preparación de análogos del GLP-1 y derivados que contienen uno o más aminoácidos no proteinogénicos en el extremo N-terminal, incluyendo las posiciones 7, 8, 9 y/o 10. Adicionalmente, se proporciona un método útil para derivar el ϵ -N de un residuo de lisina en la molécula precursora del análogo del GLP-1 mediante la acilación. Por lo tanto, la presente invención combina la producción rentable de análogos del GLP-1 y derivados mediante el uso de técnicas recombinantes con la flexibilidad de la síntesis química.

En la presente descripción, los términos siguientes tienen los significados indicados:

El término "polipéptido" y "péptido" como se usa en la presente descripción significan un compuesto formado por al menos dos aminoácidos constituyentes conectados por enlaces peptídicos. Los aminoácidos constituyentes pueden ser del grupo de los aminoácidos codificados por el código genético y pueden ser aminoácidos naturales que no se codifican por el código genético, así como también aminoácidos sintéticos. Los aminoácidos naturales que no se codifican por el código genético son, por ejemplo, γ -carboxiglutamato, ornitina, fosfoserina, D-alanina y D-glutamina. Los aminoácidos sintéticos comprenden los aminoácidos fabricados mediante síntesis química, es decir, los D-isómeros de los aminoácidos codificados por el código genético tales como D-alanina y D-leucina, Aib (ácido α -aminoisobutírico), Abu (ácido α -aminobutírico), Tle (terc-butilglicina), β -alanina, ácido 3-aminometil benzoico, y ácido antranílico.

Los 22 aminoácidos proteinogénicos son: Alanina, Arginina, Asparagina, Ácido Aspártico, Cisteína, Cistina, Glutamina, Ácido Glutámico, Glicina, Histidina, Hidroxiprolina, Isoleucina, Leucina, Lisina, Metionina, Fenilalanina, Prolina, Serina, Treonina, Triptófano, Tirosina y Valina.

Por lo tanto, un aminoácido no proteinogénico es una porción que puede incorporarse en un péptido mediante enlaces peptídicos, pero no es un aminoácido proteinogénico. Los ejemplos de aminoácidos no proteinogénicos son, pero sin limitarse a, γ -carboxiglutamato, ornitina, fosfoserina, los D-aminoácidos tales como D-alanina y D-glutamina. Los aminoácidos sintéticos no proteinogénicos comprenden aminoácidos fabricados mediante síntesis química, es decir, isómeros D de los aminoácidos codificados por el código genético tal como D-alanina y D-leucina, Aib (ácido α -aminoisobutírico), Abu (ácido α -aminobutírico), Tle (terc-butilglicina), ácido 3-aminometil benzoico, ácido antranílico, des-amino-Histidina, los análogos beta de aminoácidos tales como β -alanina etc., D-histidina, desamino-histidina, 2-amino-histidina, β -hidroxi-histidina, homohistidina, N^{α} -acetil-histidina, α -fluorometil-histidina, α -metil-histidina, 3-piridilalanina, 2-piridilalanina o 4-piridilalanina, ácido carboxílico (1-aminociclopropil), ácido carboxílico (1-aminociclobutil), ácido carboxílico (1-aminociclopentil), ácido carboxílico (1-aminociclohexilo), ácido carboxílico (1-aminocicloheptil), o ácido carboxílico (1-aminociclooctil), α -metil prolina, 1-metil histidina, 3-metil histidina, y ácido 4,5,6,7-tetrahidro-1H-imidazo[4,5-C]piridina-6-carboxílico β -(1,2,4-triazol-1-il)-alanina.

El término "protegido de DPP-IV" como se usa en la presente descripción por referencia a un polipéptido significa un polipéptido que se ha modificado químicamente para hacer que dicho compuesto sea resistente a la peptidasa plasmática dipeptidil peptidasa-4 (DPP-IV). Se conoce que la enzima DPP-IV en plasma está implicada en la degradación de varias hormonas peptídicas, por ejemplo GLP-1, GLP-2, Exendina-4 etcétera. Por lo tanto, se hace un esfuerzo considerable para desarrollar análogos y derivados de los polipéptidos susceptibles a la hidrólisis por DPP-IV con el fin de reducir la velocidad de degradación por DPP-IV. En una modalidad un péptido protegido de DPP-IV es más resistente a DPP-IV que el GLP-1(7-37).

La resistencia de un péptido a la degradación por la DPP-IV se determina mediante el ensayo de degradación siguiente:

5 Se incuban alícuotas del péptido (5 nmol) a 37 °C con 1 µl de DPP-IV purificada, correspondiente a una actividad enzimática de 5 mU durante 10-180 minutos en 100 µl de tampón trietilamina-HCl 0,1 M, pH 7,4. Las reacciones enzimáticas se terminan mediante la adición de 5 µl de ácido trifluoroacético al 10 %, y los productos de degradación peptídica se separan y cuantifican mediante el uso de análisis por HPLC. Un método para realizar este análisis es: Las mezclas se aplican sobre una columna Vydac C18 de poro ancho (poros de 30 nm, partículas de 5 µm) de 250×4,6 mm y se eluyen a una velocidad de flujo de 1 ml/min con gradientes lineales, en etapas, de acetonitrilo en 10 0,1 % de ácido trifluoroacético (0 % de acetonitrilo durante 3 minutos, 0-24 % de acetonitrilo durante 17 minutos, 24-48 % de acetonitrilo durante 1 minuto) de acuerdo con Siegel y otros, Regul. Pept. 1999;79:93-102 y Mentlein y otros Eur. J. Biochem. 1993;214:829-35. Los péptidos y sus productos de degradación pueden monitorearse mediante su absorbancia a 220 nm (enlaces peptídicos) o a 280 nm (aminoácidos aromáticos), y se cuantifican mediante la 15 integración de las áreas de sus picos con relación a las de los estándares. La velocidad de hidrólisis de un péptido por la dipeptidil peptidasa DPP-IV se estima a tiempos de incubación que resulten en menos del 10 % de hidrólisis del péptido.

En su sentido más amplio, el término "análogo del GLP-1" como se usa en la presente descripción significa un 20 análogo de una molécula de la familia de péptidos del glucagón o un análogo de la familia de exendinas. La familia de péptidos del glucagón está codificada por el gen de proglucagón y abarca tres péptidos pequeños con un alto grado de homología, es decir glucagón (1-29), GLP-1 (1-37) y GLP-2 (1-33). Las exendinas son péptidos expresados en lagartos, y como el GLP-1, son insulino-trópicas. Ejemplos de exendinas son exendina-3 y exendina-4.

El término "análogo" como se usa en la presente descripción con referencia a un polipéptido significa un péptido 25 modificado en donde uno o más residuos de aminoácidos del péptido se han sustituido por otros residuos de aminoácidos y/o en donde se han eliminado uno o más residuos de aminoácido del péptido y/o en donde se ha agregado uno o más residuos de aminoácido al péptido. Dicha adición o delección de residuos de aminoácido puede tener lugar en el N-terminal del péptido y/o en el C-terminal del péptido. Por lo tanto un análogo del GLP-1 de acuerdo con la invención es un análogo del GLP-1 (1-37) de acuerdo con la invención en donde uno o más residuos 30 de aminoácidos del péptido se han sustituido por otros residuos de aminoácidos y/o en donde uno o más residuos de aminoácidos se han eliminado del péptido y/o en donde uno o más residuos de aminoácidos se han agregado al péptido.

Para los propósitos de la presente, cualquier sustitución, delección y/o adición de aminoácidos se refiere a la 35 secuencia de GLP-1(7-37) humano que se incluye en la presente descripción como la SEQ ID NO: 1. Sin embargo, la numeración de los residuos de aminoácidos en el listado de secuencias siempre comienza con el núm. 1, mientras que para el presente propósito queremos, después de la práctica establecida en la técnica, comenzar con el residuo de aminoácido núm. 7 y asignarle el número 7 a este. Por lo tanto, generalmente, cualquier referencia en la presente descripción al número de posición de la secuencia de GLP-1(7-37) es a la secuencia que comienza con His en la 40 posición 7 y termina con Gly en la posición 37.

A menudo se usa un sistema simple para describir análogos del GLP-1: por ejemplo [Arg³⁴]GLP-1(7-37)Lys designa un análogo del GLP-1(7-37) en donde la lisina de origen natural en la posición 34 se ha sustituido con arginina y en donde una lisina se ha agregado al extremo C-terminal. Otro ejemplo [Arg³⁴]GLP-1(11-37) designa un análogo del 45 GLP-1 en donde la lisina de origen natural en la posición 34 se ha sustituido con arginina y los aminoácidos en las posiciones 7, 8, 9 y 10 están ausentes.

La expresión "una posición equivalente a" cuando se usa en la presente descripción para caracterizar una secuencia 50 de GLP-1(7-37) modificada se refiere a la posición correspondiente en la secuencia de GLP-1(7-37) natural (que tiene la SEQ ID NO: 1). Las posiciones correspondientes se deducen fácilmente, por ejemplo, por escritura a mano y observación simples. En la alternativa, puede usarse un programa estándar de alineamiento de proteínas o péptidos, tal como "align", que es un alineamiento de Needleman-Wunsch. El algoritmo se describe en Needleman, S.B. y Wunsch, C.D., (1970), Journal of Molecular Biology, 48: 443-453, y el programa de alineamiento de Myers y W. Miller en "Alignments Optimum in Linear Space" CABIOS (aplicaciones informáticas en las biociencias) (1988) 4:11- 55 17. Para el alineamiento, puede usarse la matriz de puntuación predeterminada BLOSUM50 y la matriz de identidad predeterminada, y la penalización para el primer residuo en una interrupción puede fijarse en -12 y las penalizaciones para residuos adicionales en una interrupción en -2.

El término "derivado" como se usa en la presente descripción en relación con un péptido significa un péptido modificado 60 químicamente o su análogo, en donde al menos un sustituyente se ha unido al péptido no modificado o su análogo, es decir, un péptido que se ha modificado covalentemente. Las modificaciones típicas son amidas, carbohidratos, grupos alquilo, grupos acilo, ésteres y similares. Un ejemplo de un derivado del GLP-1(7-37) es N²⁶-(γ-Glu(N^α-hexadecanoil))- [Arg³⁴, Lys²⁶]GLP-1(7-37).

65 Debe entenderse que todos los aminoácidos para los que no se indica el isómero óptico son el isómero L.

En las modalidades de la invención un máximo de 17 aminoácidos en el análogo del GLP-1 se han modificado (sustituido, eliminado, agregado o cualquiera de sus combinaciones) con relación al GLP-1(7-37) humano. En modalidades de la invención se han modificado un máximo de 15 aminoácidos en el análogo del GLP-1. En modalidades de la invención se han modificado un máximo de 10 aminoácidos en el análogo del GLP-1. En modalidades de la invención se han modificado un máximo de 8 aminoácidos en el análogo del GLP-1. En modalidades de la invención se han modificado un máximo de 7 aminoácidos en el análogo del GLP-1. En modalidades de la invención se han modificado un máximo de 6 aminoácidos en el análogo del GLP-1. En modalidades de la invención se han modificado un máximo de 5 aminoácidos en el análogo del GLP-1. En modalidades de la invención se han modificado un máximo de 4 aminoácidos en el análogo del GLP-1. En modalidades de la invención se han modificado un máximo de 3 aminoácidos en el análogo del GLP-1. En modalidades de la invención se han modificado un máximo de 2 aminoácidos en el análogo del GLP-1. En modalidades de la invención se ha modificado 1 aminoácido en el análogo del GLP-1.

En una modalidad el análogo del GLP-1 es un agente insulínico.

En una modalidad el análogo del GLP-1 es un agonista del GLP-1.

El término "agonista del GLP-1" como se usa en la presente descripción se refiere a cualquier péptido similar al glucagón, que activa total o parcialmente el receptor del GLP-1 humano. En una modalidad preferida el "agonista del GLP-1" es cualquier péptido similar al glucagón que se une a un receptor de GLP-1, preferentemente, con una constante de afinidad (KD) o una potencia (EC₅₀) de menos de 1 µM, por ejemplo más abajo de 100 nM según se midió por métodos conocidos en la técnica (ver, por ejemplo, el documento WO 98/08871) y exhibe actividad insulínica, donde la actividad insulínica puede medirse mediante ensayos *in vivo* o *in vitro* conocidos por los expertos en la técnica. Por ejemplo, el agonista del GLP-1 puede administrarse a un animal y la concentración de insulina se mide en el tiempo.

En un aspecto, el análogo del GLP-1 que se va a incluir en las composiciones farmacéuticas de la presente invención es un análogo del GLP-1, en donde el análogo comprende al menos un péptido no proteinogénico.

La propiedad insulínica de un compuesto puede determinarse mediante ensayos *in vivo* o *in vitro* bien conocidos en la técnica. Por ejemplo, el compuesto puede administrarse a un animal y monitorear la concentración de insulina en el tiempo. Los análogos insulínicos de exendina-4(1-39) se refieren a las moléculas respectivas en donde uno o más de los residuos de aminoácidos se han intercambiado con otros residuos de aminoácidos y/o de los cuales uno o más residuos de aminoácidos se han eliminado y/o de los cuales uno o más residuos de aminoácidos se han agregado con la condición de que dicho análogo sea insulínico o sea un fármaco de un compuesto insulínico. Un ejemplo de un análogo insulínico de exendina-4(1-39) es Ser²Asp³-exendina-4(1-39) en donde los residuos de aminoácidos en la posición 2 y 3 se han sustituido con serina y ácido aspártico, respectivamente (este análogo particular también se conoce en la técnica como exendina-3). Los derivados insulínicos de exendina-4(1-39) y los análogos de esta son los que los expertos en la técnica consideran derivados de estos péptidos, es decir, que tienen al menos un sustituyente que no está presente en la molécula peptídica original con la condición de que dicho derivado sea insulínico o sea un fármaco de un compuesto insulínico. Los ejemplos de sustituyentes son amidas, carbohidratos, grupos alquilo, ésteres y sustituyentes lipofílicos. Un ejemplo de derivados insulínicos de exendina-4(1-39) y análogos de esta es Tyr³¹-exendina-4(1-31)-amida.

Debe entenderse que todos los aminoácidos para los que no se indica el isómero óptico son el isómero L.

El término "agente insulínico" como se usa en la presente descripción, significa un compuesto que es un agonista del receptor de GLP-1 humano, es decir, un compuesto que estimula la formación de AMPc en un medio adecuado que contiene el receptor de GLP-1 humano (uno de tales medios se describe más abajo). La potencia de un agente insulínico se determina mediante el cálculo del valor de EC₅₀ de la curva dosis-respuesta como se describe a continuación.

Las células de riñón de hámster bebé (BHK) que expresan el receptor del GLP-1 humano clonado (BHK- 467-12A) se cultivaron en medio DMEM con la adición de 100 IU/ml de penicilina, 100 µg/ml de estreptomina, suero fetal bovino al 5 % y 0,5 mg/ml de Geneticina G-418 (Life Technologies). Las células se lavaron dos veces en solución salina tamponada con fosfato y se recogieron con Verseno. Las membranas plasmáticas se prepararon a partir de las células mediante homogenización con un Ultraturax en tampón 1 (20 mM de HEPES-Na, 10 mM de EDTA, pH 7,4). El homogeneizado se centrifugó a 48 000 x g durante 15 min a 4 °C. El sedimento se suspendió por homogenización en el tampón 2 (20 mM de HEPES-Na, 0,1 mM de EDTA, pH 7,4), luego se centrifugó a 48 000 x g durante 15 min a 4 °C. El procedimiento de lavado se repitió una vez más. El sedimento final se suspendió en el tampón 2 y se usó inmediatamente para los ensayos o se almacenó a -80 °C.

El ensayo funcional del receptor se realizó mediante la medición del AMP cíclico (AMPc) como una respuesta a la estimulación por el agente insulínico. El AMPc formado se cuantificó mediante el kit de AMPc AlphaScreen (Perkin Elmer Life Sciences). Se realizaron incubaciones en placas de microtitulación de 96 pocillos de área media

en un volumen total de 50 µl de tampón 3 (50 mM de Tris-HCl, 5 mM de HEPES, 10 mM de MgCl₂, pH 7,4) y con las siguientes adiciones: 1 mM de ATP, 1 µM de GTP, 0,5 mM de 3-isobutil-1-metilxantina (IBMX), 0,01 % de Tween-20, 0,1 % de BSA, 6 µg de preparación de membrana, 15 µg/ml de perlas aceptoras, 20 µg/ml de perlas donantes preincubadas previamente con 6 nM de biotiniil-AMPc. Los compuestos a evaluar para la actividad agonista se disolvieron y diluyeron en el tampón 3. GTP se preparó en el momento para cada experimento. La placa se incubó en la oscuridad con agitación lenta durante tres horas a temperatura ambiente, seguido por el recuento en el instrumento Fusion (Perkin Elmer Life Sciences). Las curvas de concentración-respuesta se representaron gráficamente para los compuestos individuales y los valores estimados de EC₅₀ mediante el uso de un modelo logístico de cuatro parámetros con Prism versión 4.0 (GraphPad, Carlsbad, CA).

El término "semirrecombinante" como se usa en la presente descripción se refiere a un método que comprende las etapas de preparar un primer fragmento peptídico tal como un primer fragmento de un análogo del GLP-1 mediante un proceso recombinante, es decir, a partir de un cultivo de células huésped, y posteriormente acoplar un segundo fragmento de un péptido al primer fragmento del péptido, en donde el segundo fragmento se prepara químicamente, es decir en solución o en una resina mediante el uso de química de péptidos en fase sólida. El acoplamiento de los fragmentos puede realizarse ya sea química o enzimáticamente.

El término "separar la molécula precursora expresada" como se usa en la presente descripción significa cualquier tipo de separación, que incluye el aislamiento del compuesto producido recombinantemente e incluye, además, desintegrar o permeabilizar una célula huésped.

En consecuencia, la presente invención se refiere a un método para producir un análogo o derivado del GLP-1 que comprende uno o más aminoácidos no proteinogénicos en la parte N-terminal, dicho método que comprende las etapas de:

(i) cultivar una célula huésped que comprende una secuencia de nucleótidos que codifica una molécula precursora de dicho análogo o derivado del GLP-1 en condiciones adecuadas para la expresión de dicha molécula precursora,

(ii) separar la molécula precursora expresada del caldo de cultivo,

(iii) acoplar una extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos a la molécula precursora expresada,

(iv) aislar el análogo o derivado del GLP-1 resultante por medios adecuados, como se conoce en la técnica.

En el presente contexto, el término "molécula precursora del análogo del GLP-1", debe entenderse como un péptido derivado de los péptidos GLP-1, GLP-2, Exendina-3 o Exendina-4, es decir, una secuencia análoga de GLP-1, antes de la adición de una extensión de aminoácidos N-terminal. Los ejemplos incluyen, pero sin limitarse a GLP-1 (11-37) humano o un análogo de estos, GLP-1 humano (9-37) o un análogo de estos y GLP-1 (8-37) o un análogo de estos.

En consecuencia, en un aspecto la molécula precursora del análogo del GLP-1 de acuerdo con la invención, comprende un péptido GLP-1 derivado de la secuencia GLP-1 (7-37) como se muestra en la SEQ ID NO:1.

La Exendina-4 es un péptido de residuos de 39 aminoácidos aislado a partir del veneno de los lagartos monstruos de Gila (*Sospecha de Heloderma* y *Heloderma horridum*). Este péptido comparte 52 % de homología con GLP-1 (7-37) en la región de superposición. La exendina-4 es un agonista potente del receptor del GLP-1 y se ha demostrado además que estimula la liberación de insulina y que por consiguiente disminuye el nivel de glucosa en sangre cuando se inyecta en perros.

Por lo tanto, debido a la homología del 52 % con GLP-1 (7-37) en la región de superposición, la molécula precursora del análogo del GLP-1 de acuerdo con la invención, puede derivarse en ciertas modalidades de un péptido exendina-4 como se muestra en la SEQ ID NO:2.

En consecuencia, las moléculas precursoras del análogo del GLP-1 pueden en un aspecto adicional comprender una secuencia de aminoácidos de la fórmula general:

Thr-Phe-Thr-Ser-Asp-Xaa₁₆-Ser-Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Xaa₂₃-Ala-Xaa₂₅-
Xaa₂₆-Xaa₂₇-Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Xaa₃₄-Xaa₃₅-Xaa₃₆-Xaa₃₇-Xaa₃₈-Xaa₃₉-
Xaa₄₀-Xaa₄₁-Xaa₄₂-Xaa₄₃-Xaa₄₄-Xaa₄₅ (SEQ ID NO:3)

en donde

Xaa₁₆ es Val o Leu;

Xaa₁₈ es Ser, Lys o Arg;

- Xaa₁₉ es Tyr o Gln;
- Xaa₂₀ es Leu o Met;
- 5 Xaa₂₂ es Gly o Glu;
- Xaa₂₃ es Gln, Glu, Lys o Arg;
- Xaa₂₅ es Ala o Val;
- 10 Xaa₂₆ es Lys, Glu o Arg;
- Xaa₂₇ es Glu o Leu;
- 15 Xaa₃₀ es Ala, Glu o Arg;
- Xaa₃₃ es Val, Lys, o Arg;
- Xaa₃₄ es Lys, Glu, Asn, His o Arg;
- 20 Xaa₃₅ es Gly;
- Xaa₃₆ es Arg, Gly o Lys;
- 25 Xaa₃₇ es Gly, Ala, Glu, Pro, Lys, o está ausente;
- Xaa₃₈ es Lys, Ser o está ausente.
- Xaa₃₉ es Ser, Lys, o está ausente;
- 30 Xaa₄₀ es Gly o ausente.
- Xaa₄₁ es Ala o está ausente;
- 35 Xaa₄₂ es Pro o está ausente;
- Xaa₄₃ es Pro o está ausente;
- Xaa₄₄ es Pro o está ausente;
- 40 Xaa₄₅ es Ser o está ausente;
- siempre y cuando si Xaa₃₈, Xaa₃₉, Xaa₄₀, Xaa₄₁, Xaa₄₂, Xaa₄₃, Xaa₄₄ o Xaa₄₅ está ausente, entonces cada residuo de aminoácido corriente abajo también está ausente.
- 45 En otros aspectos, la molécula precursora del análogo del GLP-1 se selecciona de la lista de moléculas precursoras que comprende la secuencia de aminoácidos de la fórmula general:
- 50 Xaa₉- Xaa₁₀-Thr-Phe-Thr-Ser-Asp-Xaa₁₆-Ser-Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Xaa₂₃-
Ala-Xaa₂₅-Xaa₂₆-Xaa₂₇-Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Xaa₃₄-Xaa₃₅-Xaa₃₆-Xaa₃₇-
Xaa₃₈-Xaa₃₉-Xaa₄₀-Xaa₄₁-Xaa₄₂-Xaa₄₃-Xaa₄₄-Xaa₄₅ (SEQ ID NO:4)
- 55 en donde
- Xaa₉ es Glu o Asp
- Xaa₁₀ es Gly o Ala
- 60 Xaa₁₆ es Val o Leu;
- Xaa₁₈ es Ser, Lys o Arg;
- 65 Xaa₁₉ es Tyr o Gln;

Xaa₂₀ es Leu o Met;

Xaa₂₂ es Gly o Glu;

5 Xaa₂₃ es Gln, Glu, Lys o Arg;
Xaa₂₅ es Ala o Val;

Xaa₂₆ es Lys, Glu o Arg;

10 Xaa₂₇ es Glu o Leu;

Xaa₃₀ es Ala, Glu o Arg;

Xaa₃₃ es Val, Lys, o Arg;

15 Xaa₃₄ es Lys, Glu, Asn, His o Arg;

Xaa₃₅ es Gly;

20 Xaa₃₆ es Arg, Gly o Lys;

Xaa₃₇ es Gly, Ala, Glu, Pro, Lys, o está ausente;

Xaa₃₈ es Lys, Ser o está ausente.

25 Xaa₃₉ es Ser, Lys, o está ausente;

Xaa₄₀ es Gly o ausente.

30 Xaa₄₁ es Ala o está ausente;

Xaa₄₂ es Pro o está ausente;

Xaa₄₃ es Pro o está ausente;

35 Xaa₄₄ es Pro o está ausente;

Xaa₄₅ es Ser o está ausente;

40 siempre y cuando si Xaa₃₈, Xaa₃₉, Xaa₄₀, Xaa₄₁, Xaa₄₂, Xaa₄₃, Xaa₄₄ o Xaa₄₅ está ausente, entonces cada residuo de aminoácido corriente abajo también está ausente.

En un aspecto adicional, la molécula precursora del análogo del GLP-1 se selecciona de las moléculas precursoras que comprenden una secuencia de aminoácidos de la fórmula general:

45
 Xaa₈- Xaa₉- Xaa₁₀-Thr-Phe-Thr-Ser-Asp-Xaa₁₆-Ser-Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-
 Xaa₂₃-Ala-Xaa₂₅-Xaa₂₆-Xaa₂₇-Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Xaa₃₄-Xaa₃₅-Xaa₃₆-
 50 Xaa₃₇-Xaa₃₈-Xaa₃₉-Xaa₄₀-Xaa₄₁-Xaa₄₂-Xaa₄₃-Xaa₄₄-Xaa₄₅ (SEQ ID NO:5)

en donde

55 Xaa₈ es Ala, Gly, Val, Leu, Ile o Lys;

Xaa₉ es Gly o Asp;

Xaa₁₀ es Gly o Ala;

60 Xaa₁₆ es Val o Leu;

Xaa₁₈ es Ser, Lys o Arg;

65 Xaa₁₉ es Tyr o Gln;

Xaa₂₀ es Leu o Met;

Xaa₂₂ es Gly o Glu;

5 Xaa₂₃ es Gln, Glu, Lys o Arg;

Xaa₂₅ es Ala o Val;

Xaa₂₆ es Lys, Glu o Arg;

10 Xaa₂₇ es Glu o Leu;

Xaa₃₀ es Ala, Glu o Arg;

15 Xaa₃₃ es Val, Lys, o Arg;

Xaa₃₄ es Lys, Glu, Asn, His o Arg;

Xaa₃₅ es Gly;

20 Xaa₃₆ es Arg, Gly o Lys;

Xaa₃₇ es Gly, Ala, Glu, Pro, Lys, o está ausente;

25 Xaa₃₈ es Lys, Ser o está ausente,

Xaa₃₉ es Ser, Lys, o está ausente;

Xaa₄₀ es Gly o ausente.

30 Xaa₄₁ es Ala o está ausente;

Xaa₄₂ es Pro o está ausente;

35 Xaa₄₃ es Pro o está ausente;

Xaa₄₄ es Pro o está ausente;

Xaa₄₅ es Ser o está ausente;

40 siempre y cuando si Xaa₃₈, Xaa₃₉, Xaa₄₀, Xaa₄₁, Xaa₄₂, Xaa₄₃, Xaa₄₄ o Xaa₄₅ está ausente, entonces cada residuo de aminoácido corriente abajo también está ausente.

45 Las moléculas precursoras de los análogos del GLP-1 de la presente invención pueden prepararse por medio de tecnología de ADN recombinante mediante el uso de métodos y principios generales conocidos por la persona experta en la técnica, mediante el cultivo de una célula huésped adecuada que comprende una secuencia de nucleótidos que codifica una molécula precursora del análogo del GLP-1 en condiciones adecuadas para la expresión de la molécula precursora.

50 La secuencia de ácido nucleico que codifica la molécula precursora del análogo del GLP-1 puede prepararse sintéticamente por métodos estándar establecidos bien conocidos en la técnica, por ejemplo al sintetizar oligonucleótidos en un sintetizador de ADN automático, y posteriormente purificar, alinear, ligar y clonar las secuencias de ácido nucleico en vectores adecuados.

55 La clonación de las secuencias de ácidos nucleicos que codifican la molécula precursora del análogo del GLP-1 puede efectuarse, por ejemplo, mediante el uso de la bien conocida reacción en cadena de la polimerasa (PCR).

60 La secuencia de ácido nucleico que codifica la molécula precursora del análogo del GLP-1 se inserta en un vector de expresión recombinante que puede ser cualquier vector que pueda someterse convenientemente a los procedimientos de ADN recombinante. La elección del vector dependerá a menudo de la célula huésped en la que se va a introducir. Por lo tanto, el vector puede ser un vector de replicación autónoma, es decir, un vector que existe como una entidad extracromosómica, cuya replicación es independiente de la replicación cromosómica, por ejemplo, un plásmido. Alternativamente, el vector puede ser uno que, cuando se introduce en una célula huésped, se integra en el genoma de la célula huésped y se replica junto con el(los) cromosoma(s) donde se ha integrado.

65

Para los productos extracelulares los componentes proteicos del sobrenadante se aíslan mediante filtración, cromatografía en columna o precipitación, por ejemplo, microfiltración, ultrafiltración, precipitación isoeléctrica, purificación mediante una variedad de procedimientos cromatográficos, por ejemplo, cromatografía de intercambio iónico, cromatografía de interacción hidrofóbica, cromatografía de filtración en gel, cromatografía de afinidad, o similares, dependiendo del tipo de polipéptido en cuestión. Para los productos intracelulares o periplasmáticos las células aisladas del medio de cultivo se desintegran o permeabilizan y se extraen para recuperar el polipéptido del producto o precursor de este.

La secuencia de ADN que codifica la molécula precursora del análogo del GLP-1 puede ser adecuadamente de origen genómico o de ADNc, por ejemplo, obtenida mediante la preparación de una biblioteca genómica o de ADNc y mediante la selección de secuencias de ADN que codifican todo o parte del péptido por hibridación mediante el uso de sondas de oligonucleótidos sintéticas de acuerdo con técnicas estándar (ver, por ejemplo, Sambrook, J, Fritsch, EF y Maniatis, T, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Nueva York, 1989). La secuencia de ADN que codifica el polipéptido puede prepararse, además, sintéticamente mediante métodos estándar establecidos, por ejemplo, el método de fosfoamidita descrito por Beaucage y Caruthers, *Tetrahedron Letters* 22 (1981), 1859-1869, o el método descrito por Matthes y otros, *EMBO Journal* 3 (1984), 801-805. La secuencia de ADN puede prepararse, además, por la reacción en cadena de la polimerasa mediante el uso de cebadores específicos, por ejemplo como se describió en la patente de Estados Unidos núm. 4,683,202 o Saiki y otros, *Science* 239 (1988), 487-491.

La secuencia de ADN que codifica la molécula precursora del análogo del GLP-1 puede insertarse en cualquier vector que pueda someterse convenientemente a procedimientos de ADN recombinante, y la elección del vector dependerá a menudo de la célula huésped en la que se va a introducir. De este modo, el vector puede ser un vector de replicación autónoma, es decir, un vector que existe como una entidad extracromosómica, cuya replicación es independiente de la replicación cromosómica, por ejemplo, un plásmido. Alternativamente, el vector puede ser uno que, cuando se introduce en una célula huésped, se integra en el genoma de la célula huésped y se replica junto con el(los) cromosoma(s) donde se ha integrado.

El vector es preferentemente un vector de expresión en el que la secuencia de ADN que codifica la molécula precursora del análogo del GLP-1 se une operativamente a segmentos adicionales necesarios para la transcripción del ADN, tal como un promotor. El promotor puede ser cualquier secuencia de ADN que muestre actividad transcripcional en la célula huésped de elección y puede derivarse de genes que codifican proteínas ya sea homólogas o heterólogas a la célula huésped. Los ejemplos de promotores adecuados para dirigir la transcripción del ADN que codifica el péptido de la invención en una diversidad de células huésped son bien conocidos en la técnica (consultar, por ejemplo, Sambrook y otros, más arriba).

La secuencia de ADN que codifica la molécula precursora del análogo del GLP-1 puede, además, si es necesario, conectarse operativamente a un terminador adecuado, señales de poliadenilación, secuencias potenciadoras de la transcripción y secuencias potenciadoras de la traducción. El vector recombinante de la invención puede comprender, además, una secuencia de ADN que permite que el vector se replique en la célula huésped en cuestión.

El vector puede comprender, además, un marcador seleccionable, por ejemplo, un gen cuyo producto complemente un defecto en la célula huésped o uno que confiera resistencia a un fármaco, por ejemplo, ampicilina, kanamicina, tetraciclina, cloranfenicol, neomicina, higromicina o metotrexato. Para la fabricación a gran escala el marcador seleccionable preferentemente no es la resistencia a antibióticos, por ejemplo, los genes de resistencia a antibióticos en el vector se escinden preferentemente cuando el vector se usa para la fabricación a gran escala. Los métodos para eliminar genes de resistencia a antibióticos de vectores se conocen en la técnica, ver por ejemplo US 6,358,705 que se incorpora en la presente descripción como referencia.

Para dirigir la molécula precursora del análogo del GLP-1 de la presente invención hacia la ruta secretora de las células huésped, puede proporcionarse una secuencia señal secretora (también conocida como una secuencia líder, secuencia prepro o secuencia pre) en el vector recombinante. La secuencia señal secretora se une a la secuencia de ADN que codifica el péptido en el marco de lectura correcto. Las secuencias señal secretoras se colocan comúnmente 5' con respecto a la secuencia de ADN que codifica el péptido. La secuencia señal secretora puede ser la asociada normalmente con el péptido o puede ser de un gen que codifica otra proteína secretada.

Los procedimientos usados para ligar las secuencias de ADN que codifican la molécula precursora del análogo del GLP-1, el promotor y opcionalmente el terminador y/o la secuencia señal secretora, respectivamente, y para insertarlos en vectores adecuados que contienen la información necesaria para la replicación, son bien conocidos por los expertos en la técnica (consultar, por ejemplo, Sambrook y otros, más arriba).

La célula huésped en la que la secuencia de ADN o el vector recombinante se introduce puede ser cualquier célula que sea capaz de producir la molécula precursora del análogo del GLP-1 e incluye, pero sin limitarse a, células huésped de mamífero, células huésped de aves, células huésped de insectos, células huésped vegetales, células huésped bacterianas, células huésped fúngicas y células huésped de levaduras. Los ejemplos de células huésped

adecuadas bien conocidas y usadas en la técnica incluyen, pero sin limitarse a *Saccharomyces cerevisiae*, *Pichia pastoris* y otras levaduras, *E. coli*, *Bacillus subtilis*, y *Pseudomonas fluorescens* o líneas celulares de mamíferos BHK o CHO.

5 Los métodos para introducir el ácido nucleico exógeno en los huéspedes se conocen bien en la técnica, y variarán con la célula huésped usada. La célula huésped, cuando se cultiva en condiciones apropiadas, sintetiza una molécula precursora del GLP-1 que puede recolectarse posteriormente del medio de cultivo (si la célula huésped lo secreta al medio) o directamente de la célula huésped que lo produce (si no lo secreta).

10 En una modalidad, las moléculas precursoras del GLP-1 se producen en células de levadura. Los huéspedes de expresión de levadura se conocen bien en la técnica e incluyen, pero sin limitarse a, *Saccharomyces cerevisiae*, *Candida albicans* y *C. Maltosa*, *Hansenula polymorfa*, *Kluyveromyces fragilis* y *K. lactis*, *Pichia guillerimoni*, *Pichia metanolica* y *Pichia pastoris*, *Schizosaccharomyces pombe*, y *Yarrowia lipolytica*.

15 En una modalidad, las moléculas precursoras del GLP-1 se producen en sistemas bacterianos. Los huéspedes de expresión bacteriana se conocen bien en la técnica, e incluyen pero sin limitarse a *E. coli*, *Bacillus subtilis*, *Pseudomonas fluorescens*, *Lactococcus lactis*, *Streptococcus cremoris*, y *Streptococcus nutridans*.

20 En una modalidad, las moléculas precursoras del GLP-1 se expresan en hongos filamentosos. Los huéspedes de hongos filamentosos se conocen bien en la técnica, e incluyen pero sin limitarse a *Neurospora*, *penicillium*, *tolipocladio*, y *Aspergillus* huéspedes como *Aspergillus nidulans*, *Aspergillus oryzea* y *Aspergillus niger*.

25 En una modalidad, las moléculas precursoras del GLP-1 se producen en células de insectos. Se han desarrollado sistemas de expresión para *Aedes aegypti*, *Autographa californica*, *Bombyx mori*, *Drosophila melangaster*, *Spodoptera frugiperda*, y *Trichoplusia ni*. Cada una de estas líneas celulares se conoce por y está disponible para los expertos en la técnica de expresión de proteínas.

30 En una modalidad, las moléculas precursoras del GLP-1 se producen en otras células eucariotas superiores. Estas incluyen, pero sin limitarse a, células de aves y células vegetales tal como *Chlamydomonas reinhardtii*.

35 En una modalidad, las moléculas precursoras del GLP-1 se producen en líneas celulares huésped de mamífero. Los ejemplos de estas incluyen, pero sin limitarse a riñón de hámster bebé (BHK), ovario de hámster chino (CHO) y células COS. Los ejemplos más específicos incluyen, pero sin limitarse a la línea de riñón de mono CV1 transformada por SV40 (COS-7, ATCC CRL 1651); línea de riñón embrionaria humana (293); células hepáticas humanas (Hep G2, HB 8065).

La célula huésped adecuada se selecciona por la persona experta en la técnica.

40 El medio usado para cultivar las células huéspedes puede ser cualquier medio convencional adecuado para hacer crecer las células huésped, tales como medios mínimos o complejos que contengan los suplementos apropiados. Los medios adecuados están disponibles a través de proveedores comerciales o pueden prepararse de acuerdo con fórmulas publicadas (por ejemplo, en catálogos de la American Type Culture Collection).

45 La molécula precursora del análogo del GLP-1 producida por las células puede entonces separarse del caldo de cultivo o medio de cultivo mediante procedimientos convencionales que incluyen separar las células huéspedes del medio mediante centrifugación o filtración.

50 Como se mencionó anteriormente, es un aspecto de la presente invención acoplar una extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos a la molécula precursora del análogo del GLP-1 expresada, para producir el compuesto más resistente a la peptidasa plasmática dipeptidil peptidasa 4 (DPP-IV).

55 Tales aminoácidos no proteinogénicos pueden seleccionarse de γ -carboxiglutamato, ornitina, fosfoserina, D-aminoácidos tales como D-alanina y D-glutamina, D-alanina y D-leucina, Aib (ácido α -aminobutírico), Abu (ácido α -aminobutírico), Tle (tert-butilglicina), 3-ácido aminometil benzoico, ácido antranílico, des-amino-Histidina, β -alanina, D-histidina, desamino-histidina, 2-amino-histidina, β -hidroxi-histidina, homohistidina, N^{ac}-acetil-histidina, α -fluorometil-histidina, α -metil-histidina, 3-piridilalanina, 2-piridilalanina, 4-piridilalanina, ácido (1-aminociclopropil) carboxílico, ácido (1-aminociclobutil) carboxílico, ácido (1-aminociclopropil) carboxílico, ácido (1-aminociclohexil) carboxílico, ácido (1-aminocicloheptil) carboxílico, ácido (1-aminociclooctil) carboxílico, α -metil prolina, 1-metil histidina, 3-metil histidina, y ácido 4,5,6,7-tetrahidro-1H-imidazo[4,5-C]piridina-6-carboxílico β -(1,2,4-triazol-1-il)-alanina. Por lo tanto, la desamino-histidina, aunque no contiene un grupo amino, se refiere en la presente descripción como aminoácido no proteinogénico, para facilitar su clasificación y nomenclatura.

60 En un aspecto, la longitud de la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos es 1 aminoácido, 2 aminoácidos, 3 aminoácidos o 4 aminoácidos.

65

En un aspecto de la invención, la longitud de la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteínogénicos es 2 aminoácidos o 4 aminoácidos.

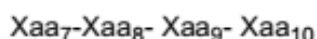
5 En un aspecto, la longitud de la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteínogénicos es 1 aminoácido o 2 aminoácidos.

En un aspecto de la invención, la longitud de la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteínogénicos es 2 aminoácidos.

10 Se contempla que no todos los aminoácidos de la extensión de aminoácidos N-terminal necesariamente son aminoácidos no proteínogénicos.

En otros aspectos de la invención, la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteínogénicos tiene la fórmula general

15



20 en donde

Xaa₇ se selecciona de L-histidina, D-histidina, desamino-histidina, 2-amino-histidina, β-hidroxi-histidina, homohistidina, N^α-acetil-histidina, α-fluorometil-histidina, α-metil-histidina, 3-piridilalanina, 2-piridilalanina y 4-piridilalanina, 1-metil histidina, 3-metil histidina, y ácido 4,5,6,7-tetrahidro-1H-imidazo[4,5-C]piridina-6-carboxílico β-(1,2,4-triazol-1-il)-alanina;

25

Xaa₈ se selecciona de Ala, Gly, Val, Leu, Ile, Lys, Aib, ácido (1-aminociclopropil) carboxílico, ácido (1-aminociclobutil) carboxílico, ácido (1-aminociclopentil) carboxílico, ácido (1-aminociclohexil) carboxílico, ácido (1-aminocicloheptil) carboxílico, o ácido (1-aminociclooctil) carboxílico y alfa-metil prolina

30

Xaa₉ se selecciona de Glu, Asp, γ,γ-dimetil Glu, β,β-dimetil Glu y β,β-dimetil Asp, y

Xaa₁₀ se selecciona de Gly, Aib, ácido (1-aminociclopropil) carboxílico, ácido (1-aminociclobutil) carboxílico, ácido (1-aminociclopentil) carboxílico, ácido (1-aminociclohexil) carboxílico, ácido (1-aminocicloheptil) carboxílico, y ácido (1-aminociclooctil) carboxílico.

35

El ácido carboxílico libre de la secuencia Xaa₇-Xaa₈-Xaa₉-Xaa₁₀ puede prepararse eficientemente en una solución o en una resina tal como resina de 2-clorotritlorcloruro, por ejemplo, mediante una estrategia basada en Fmoc mediante el uso de bloques de construcción protegidos adecuadamente mediante el uso de las condiciones de reacción adecuadas descritas en las referencias (Organic Synthesis on solid Phase, Florencio Zaragoza Dörwald, Wiley-VCH Verlag GmbH, D-69469 Weinheim, 2000), (Novabiochem Catalog, Merck Biosciences 2006/2007) y (Fmoc Solid Phase Peptide Synthesis, Editado por W.C. Chan y P.D. White, Oxford University Press, 2000, ISBN 0-19-963724-5). Los aminoácidos en la secuencia pueden estar protegidos o no protegidos. El péptido de extensión N-terminal de la secuencia Xaa₇-Xaa₈-Xaa₉-Xaa₁₀ puede activarse a un agente de acilación activado. El término agente de acilación "activado" significa un agente de acilación que se ha activado mediante el uso de técnicas generales como por ejemplo, las descritas en "Amide bond formation and peptide coupling" [Tetrahedron 61(46), 10827-10852 (2005)].

40

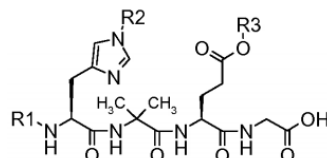
45

Los ejemplos de ésteres activados incluyen, pero sin limitarse a cloruro de ácido, bromuros de ácidos, fluoruros de ácidos, anhídrido simétrico, anhídrido mezclado, ácidos carboxílicos activados mediante el uso de carbodiimidias comunes tales como, pero sin limitarse a, diisopropilcarbodiimida (DIPCDI), N,N'-diciclohexilcarbodiimida (DCC), 1-etil-3-(3'-dimetilaminopropil)carbodiimida hidrocloreuro (EDC). Además, se incluyen, pero sin limitarse a, ácidos carboxílicos que usan los carbodiimidias antes mencionados y un aditivo tal como pero sin limitarse N-hidroxibenzotriazol (HOBt), 1-hidroxi-7-azabenzotriazol, 6-cloro-N-hidroxibenzotriazol (HOAt), 3-hidroxi-3,4-dihidro-4-oxo-1,2,3-benzotriazina (DhbToH). También se incluyen, pero sin limitarse a, ácidos carboxílicos activados con una sal de uronio o una sal de fosfonio, tal como, pero sin limitarse a, hexafluorofosfato de O-Benzotriazol-N,N,N',N'-tetrametiluronio (HBTU), tetrafluoroborato de O-(Benzotriazol-1-il)-N,N,N',N'-tetrametiluronio (TBTU), hexafluorofosfato de 2-(6-cloro-1H-benzotriazol-1-il)-1,1,3,3-tetrametilaminio (HCTU), tetrafluoroborato de 2-(6-cloro-1H-benzotriazol-1-il)-1,1,3,3-tetrametilaminio (TCTU), hexafluorofosfato de 2-(1H-7-Azabenzotriazol-1-il)-1,1,3,3-tetrametil uronio (HATU), hexafluorofosfato de 1-benzotriazolioxitris(dimetilamino)fosfonio (BOP), hexafluorofosfato de benzotriazol-1-il-oxitripirrolidinofosfonio (PYBOP). Otros ésteres activados incluyen, pero sin limitarse a ésteres de N-hidroxisuccinimida (éster NHS), pentafluorofenol éster (Pfp-éster), 2,4-dinitrofenil éster, 4-nitrofenil éster, 3-hidroxi-3,4-dihidro-4-oxo-1,2,3-benzotriazina (HODhbt) y Carbonildiimidazol (CDI), N-etil-5-fenilsoxazolium-3'-sulfonato (NEPIS), preferentemente un éster de N-hidroxisuccinimida o un éster HOBt o un derivado de este mediante el uso de las condiciones de reacción descritas en las referencias (Organic Synthesis on solid Phase, Florencio Zaragoza Dörwald, Wiley-VCH Verlag GmbH, D-69469 Weinheim, 2000), (Catálogo Novabiochem, Merck

65

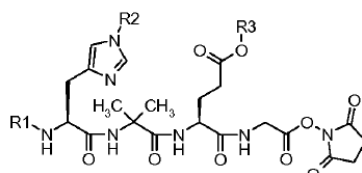
Biosciences 2006/2007) y (Fmoc Solid Phase Peptide Synthesis, Editado por W.C. Chan y P.D. White, Oxford University Press, 2000, ISBN 0-19-963724-5).

Un ejemplo de un péptido de extensión N-terminal que tiene la fórmula general Xaa₇-Xaa₈- Xaa₉-Xaa₁₀ es;



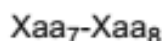
en donde R1 y R2 son grupos protectores elegidos individualmente tales como, pero sin limitarse a: Carbamatos que incluyen pero sin limitarse a 9H-fluoren-9-ilmetoxicarbonilo (Fmoc), terc-butoxicarbonilo (Boc) y bencilcarbarmato (Cbz); Amidas que incluyen pero sin limitarse a benzamida; imidas que incluyen pero sin limitarse a N-ftalimida; y alquilo que incluyen pero sin limitarse a 1-adamantilo (1-Adoc) y trifenilmetilo (Trt). En una modalidad preferida R1 se selecciona del grupo que consiste en 9H-fluoren-9-ilmetoxicarbonilo (Fmoc), terc-butoxicarbonilo (Boc) y bencilcarbarmato (Cbz), y R2 se selecciona del grupo que consiste en terc-butoxicarbonilo (Boc) y trifenilmetilo (Trt). R3 es un grupo protector seleccionado individualmente tal como pero sin limitarse a: Ésteres que incluyen pero sin limitarse a metilo, etilo, terc-butilo, bencilo y 9H-fluoren-9-ilmetilo (Fm). En una modalidad preferida R3 es terc-butil o bencilo.

Un ejemplo de un péptido de extensión N-terminal de la secuencia Xaa₇-Xaa₈- Xaa₉-Xaa₁₀ que se activa como el éster de N-hidroxisuccinimida correspondiente es;



en donde R1, R2, y R3 son grupos protectores seleccionados individualmente tal como se describió anteriormente.

En un aspecto adicional de la invención, la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos tiene la fórmula general



en donde

Xaa₇ se selecciona de L-histidina, D-histidina, desamino-histidina, 2-amino-histidina, β-hidroxi-histidina, homohistidina, N^α-acetil-histidina, α-fluorometil-histidina, α-metil-histidina, 3-piridilalanina, 2-piridilalanina y 4-piridilalanina; 1-metil histidina, 3-metil histidina, ácido 4,5,6,7-tetrahidro-1H-imidazo[4,5-C]piridina-6-carboxílico β-(1,2,4-triazol-1-il)-alanina; y

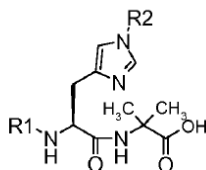
Xaa₈ se selecciona de Gly, Aib, ácido (1-aminociclopropil) carboxílico, ácido (1-aminociclobutil) carboxílico, ácido (1-aminociclopentil) carboxílico, ácido (1-aminociclohexil) carboxílico, ácido (1-aminocicloheptil) carboxílico, o ácido (1-aminociclooctil) carboxílico y alfa metil prolina

El ácido carboxílico libre de la secuencia Xaa₇-Xaa₈ puede prepararse eficientemente en una resina tal como resina de 2-clorotritlorcloruro, por ejemplo, mediante una estrategia basada en Fmoc mediante el uso de bloques de construcción protegidos adecuadamente mediante el uso de condiciones de reacción apropiadas como se describió anteriormente o alternativamente puede prepararse en solución. Los aminoácidos en la secuencia pueden estar protegidos o no protegidos.

El péptido de extensión N-terminal de la secuencia Xaa₇-Xaa₈ puede activarse a agentes de acilación activados como se describió anteriormente.

Un ejemplo de un péptido de extensión N-terminal con la fórmula Xaa₇-Xaa₈ es:

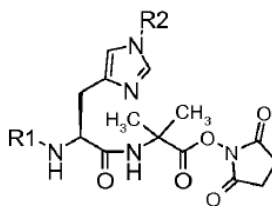
5



10 en donde R1 y R2 son grupos protectores seleccionados individualmente como se describió anteriormente.

Un ejemplo de un péptido de extensión N-terminal con la secuencia Xaa₇-Xaa₈ que se activa como el éster de N-hidroxisuccinimida correspondiente es;

15



20

25 en donde R1 y R2 son grupos protectores seleccionados individualmente como se describió anteriormente.

En un aspecto adicional de la invención, el análogo del GLP-1 producido de acuerdo con la invención, es un análogo del GLP-1 donde la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no

30

proteínogénicos tiene la fórmula general

Xaa₇-Xaa₈- Xaa₉- Xaa₁₀ como se definió anteriormente;

y la molécula precursora de dicho análogo del GLP-1 tiene la fórmula general

35 Thr-Phe-Thr-Ser-Asp-Xaa₁₆-Ser-Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Xaa₂₃-Ala-Xaa₂₅-Xaa₂₆-Xaa₂₇-Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Xaa₃₄-Xaa₃₅-Xaa₃₆-Xaa₃₇-Xaa₃₈-Xaa₃₉-Xaa₄₀-Xaa₄₁-Xaa₄₂-Xaa₄₃-Xaa₄₄-Xaa₄₅ como se definió anteriormente.

40

En un aspecto adicional de la invención, el análogo del GLP-1 producido de acuerdo con la invención, es un análogo del GLP-1 donde la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no

40 Xaa₇-Xaa₈ como se definió anteriormente;

y la molécula precursora de dicho análogo del GLP-1 tiene la fórmula general

45

Xaa₉-Xaa₁₀-Thr-Phe-Thr-Ser-Asp-Xaa₁₆-Ser-Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Xaa₂₃-Ala-Xaa₂₅-Xaa₂₆-Xaa₂₇-Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Xaa₃₄-Xaa₃₅-Xaa₃₆-Xaa₃₇-Xaa₃₈-Xaa₃₉-Xaa₄₀-Xaa₄₁-Xaa₄₂-Xaa₄₃-Xaa₄₄-Xaa₄₅ como se definió anteriormente.

50

El acoplamiento de análogos de aminoácidos adicionales en un fragmento polipeptídico no protegido tal como la molécula precursora del análogo del GLP-1 puede no ser una tarea trivial, debido a la gran cantidad de reacciones secundarias posibles que pueden tener lugar en las condiciones de reacción usadas. Un reactivo electrolítico como por ejemplo un éster activo tal como un éster de N-hidroxisuccinimida, en general, reaccionará con los nucleófilos presentes en un polipéptido. Tales nucleófilos pueden ser tioles, aminas, imidazoles, grupos hidroxilo fenólicos, alcoholes, y guanidinas. La velocidad de acilación en los grupos funcionales dependerá de la reactividad inherente de los grupos individuales, pero también de la estructura primaria, secundaria, terciaria e incluso cuaternaria del polipéptido. Por lo tanto, a menudo será difícil predecir la reactividad relativa de diferentes porciones de aminoácidos en un polipéptido. En la práctica las reacciones en un polipéptido no protegido dan lugar a mezclas de productos. Sin embargo, la reactividad relativa de los grupos individuales en un polipéptido puede, en algunos casos, controlarse parcialmente por las condiciones en las que se realiza la reacción. Por lo tanto, un cambio en el pH en una mezcla de reacción puede conducir a una reactividad alterada de algunos grupos. Igualmente, un cambio en el solvente de uno no prótico a uno prótico puede resultar en una reactividad alterada de algunos grupos.

55

60

En el método de acuerdo con la invención, el acoplamiento de la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteínogénicos a la molécula precursora del GLP-1 puede, en ciertas

65

modalidades útiles, tener lugar en un solvente, que incluye un solvente polar orgánico seleccionado de 1-metil-pirrolidin-2-ona, acetonitrilo, tetrahidrofurano, dimetilsulfóxido, *N,N*-dimetilformamida y *N*-metil-formamida.

5 En otras modalidades útiles, el acoplamiento de la extensión de aminoácidos N-terminal puede tener lugar en una mezcla de solventes acuoso, tal como una mezcla de solventes acuosos que puede comprender un solvente polar orgánico, tal como un solvente seleccionado de 1-metil-pirrolidin-2-ona, acetonitrilo, tetrahidrofurano, dimetilsulfóxido, *N,N*-dimetilformamida y *N*-metil-formamida.

10 Además, se ha descubierto que el acoplamiento de la extensión de aminoácidos N-terminal de acuerdo con la invención, puede en modalidades útiles tener lugar en una mezcla de solventes acuosos, en donde el pH en la mezcla de reacción está entre 6 y 12, tal como entre 7 y 10. En modalidades adicionales el pH en la mezcla de reacción está en el intervalo de 7-8, 8-9, 9-10 o 10-12.

15 Como se mencionó previamente, los aminoácidos en la extensión de aminoácidos N-terminal que comprenden uno o más aminoácidos no proteínogénicos pueden comprender uno o más grupos de protección. Por lo tanto, se contempla que el método de acuerdo con la invención, puede comprender además una o más etapas de eliminar uno o más de estos grupos de protección de la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteínogénicos. Tales etapas de desprotección se describen bien en la literatura.

20 El método de acuerdo con la invención puede incluir una etapa adicional de acilación del grupo épsilon amino de al menos un residuo de lisina en la molécula precursora expresada o el análogo del GLP-1 obtenido del acoplamiento de la molécula de extensión N-terminal a la molécula precursora expresada, con un agente de acilación.

25 En ciertas modalidades, puede ser ventajoso llevar a cabo esta etapa antes del acoplamiento de la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteínogénicos a la molécula precursora del análogo del GLP-1. Por lo tanto, la molécula precursora del análogo del GLP-1 usada en el método de acuerdo con la invención puede acilarse en ciertas modalidades con un agente de acilación, activado opcionalmente, en, por ejemplo, un grupo épsilon amino de un residuo de lisina.

30 Por lo tanto, en un aspecto la presente invención se refiere a un método para producir un análogo o derivado del GLP-1 que comprende uno o más aminoácidos no proteínogénicos en la parte N-terminal, dicho método que comprende las etapas de:

35 (i) cultivar una célula huésped que comprende una secuencia de nucleótidos que codifica una molécula precursora de dicho análogo del GLP-1 en condiciones adecuadas para la expresión de dicha molécula precursora,

(ii) separar la molécula precursora expresada del caldo de cultivo,

40 (iii) opcionalmente acilar el grupo épsilon amino de al menos un residuo de lisina en la molécula precursora expresada con un agente de acilación, que se activa opcionalmente, para obtener un derivado de la molécula precursora, y opcionalmente aislar dicho derivado de la molécula precursora,

45 (iv) acoplar una extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteínogénicos a la molécula precursora expresada o derivado de la molécula precursora,

(v) aislar el análogo o derivado del GLP-1 resultante por medios adecuados.

50 El método de acuerdo con la invención puede incluir opcionalmente además una etapa de proteger selectivamente el grupo épsilon amino de al menos un residuo de lisina con un agente protector de amina como se describió anteriormente.

55 En ciertas modalidades, puede ser ventajoso llevar a cabo esta etapa de protección antes del acoplamiento de la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteínogénicos a la molécula precursora del análogo del GLP-1. Por lo tanto, el derivado de la molécula precursora del análogo del GLP-1 con la lisina protegida resultante usado en el método de acuerdo con dicha modalidad de la invención puede acoplarse a la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteínogénicos. Después de dicho acoplamiento de la extensión de aminoácidos N-terminal a la molécula precursora del análogo del GLP-1 con la lisina protegida, el(los) grupo(s) de protección en los grupos épsilon amina de cualquier lisina(s) puede eliminarse mediante el uso de técnicas convencionales y el grupo amino épsilon resultante de una lisina puede opcionalmente acilarse con un agente de acilación, activado opcionalmente.

60

65 Por lo tanto, en un aspecto la presente invención se refiere a un método para producir un análogo del GLP-1 que comprende uno o más aminoácidos no proteínogénicos en la parte N-terminal, dicho método que comprende las etapas de:

(i) cultivar una célula huésped que comprende una secuencia de nucleótidos que codifica una molécula precursora de dicho péptido insulínico en condiciones adecuadas para la expresión de dicha molécula precursora,

(ii) separar la molécula precursora expresada del caldo de cultivo,

(iii) proteger el grupo épsilon amino de al menos un residuo de lisina en la molécula precursora expresada con un agente protector para obtener una molécula precursora protegida, y opcionalmente aislar dicha molécula precursora protegida,

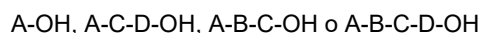
(iv) acoplar una extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos a la molécula precursora protegida para obtener un análogo del GLP-1 protegido y opcionalmente aislar dicho análogo del GLP-1 protegido,

(v) desprotección de los grupos de protección épsilon amino, y aislar opcionalmente el análogo del GLP-1 resultante,

(vi) opcionalmente acilar el grupo épsilon amino de al menos un residuo de lisina en el análogo del GLP-1 con un agente de acilación, que se activa opcionalmente, para obtener un derivado de la molécula precursora,

(vii) aislar el análogo o derivado del GLP-1 resultante por medios adecuados, como se conoce en la técnica.

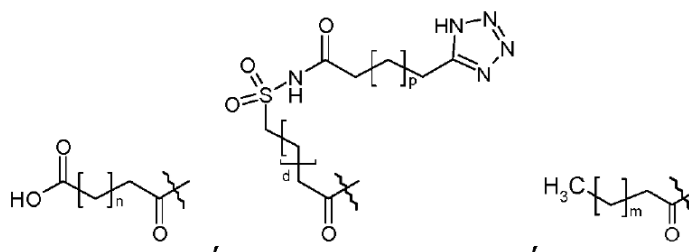
En un aspecto, el agente de acilación usado en la acilación del grupo épsilon amino de un residuo de lisina es un análogo de ácido carboxílico de la fórmula general:



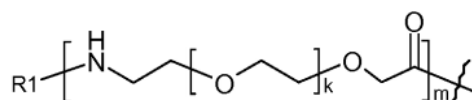
en donde un ácido carboxílico de terminal libre, tal como el grupo $-OH$ de A-OH, A-C-D-OH, A-B-C-OH o A-B-C-D-OH, opcionalmente se activa y en donde cualquier otro ácido carboxílico libre contenido opcionalmente se protege con un grupo o grupos protectores adecuados, por ejemplo, seleccionados de pero sin limitarse a terc-butilo, bencilo o alilo y

en donde

A se selecciona del grupo que consiste en;

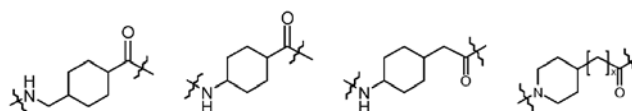


y

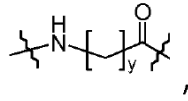


en donde n se selecciona del grupo que consiste en 14, 15, 16, 17, 18 y 19, p se selecciona del grupo que consiste en 10, 11, 12, 13 y 14, y d se selecciona del grupo que consiste en 0, 1, 2, 3, 4 y 5, y m se selecciona del grupo que consiste en 11, 12, 13, 14, 15, 16, 17, k se selecciona del grupo que consiste en 0, 1, 2, 3, 4, 5, 11 y 27, y m se selecciona del grupo que consiste en 0, 1, 2, 3, 4, 5 y 6, y R1 se selecciona del grupo de 9H-fluoren-9-ilmtoxycarbonilo (Fmoc), terc-butoxicarbonilo (Boc), bencilcarbamatato (Cbz)-

B se selecciona del grupo que consiste en



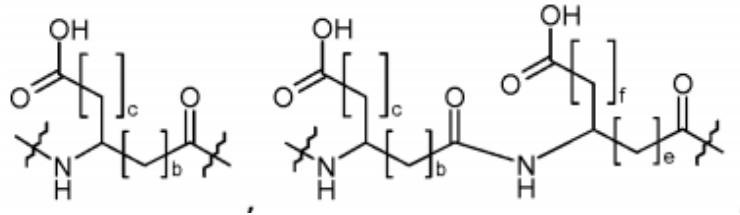
5



en donde x se selecciona del grupo que consiste en 0, 1, 2, 3 y 4, e y se selecciona del grupo que consiste en 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 y 12,

C se selecciona del grupo que consiste en

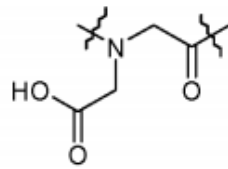
10



15

20 y

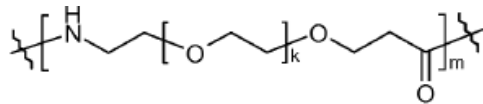
25



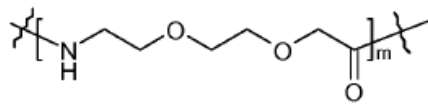
30

en donde b y e se seleccionan cada uno independientemente del grupo que consiste en 0, 1 y 2, y c y f se seleccionan cada uno independientemente del grupo que consiste en 0, 1 y 2 siempre y cuando b es 1 o 2 cuando c es 0, o b es 0 cuando c es 1 o 2, y e es 1 o 2 cuando f es 0, o e es 0 cuando f es 1 o 2, y D se selecciona del grupo que consiste en

35

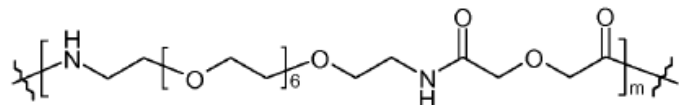


40

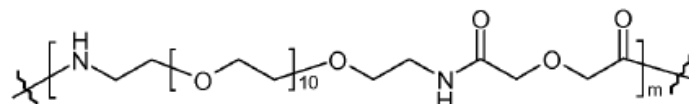


45

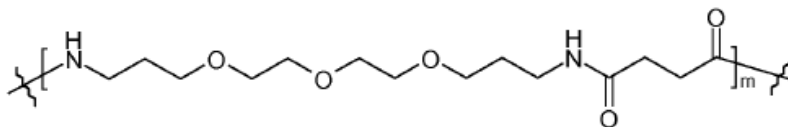
50



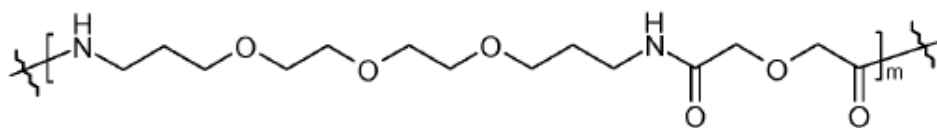
55



60



65



en donde k se selecciona del grupo que consiste en 0, 1, 2, 3, 4, 5, 11 y 27 y m se selecciona del grupo que consiste en 0, 1, 2, 3, 4, 5, y 6.

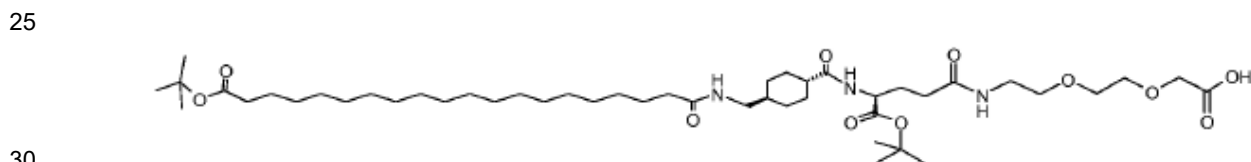
5 En un aspecto los fragmentos A-OH, A-C-D-OH, A-B-C-OH o A-B-C-D-OH usados en la acilación de la ε-N de una porción de lisina son derivados de ácido carboxílico que se activan a agentes de acilación activados como se describió anteriormente.

10 En un aspecto un grupo químico funcional libre, tal como un ácido carboxílico contenido en el fragmento A o C de los fragmentos A-OH, A-C-D-OH, A-B-C-OH o A-B-C-D-OH se protege con un grupo protector apropiado seleccionado de, pero sin limitarse a terc-butilo, bencilo o alilo.

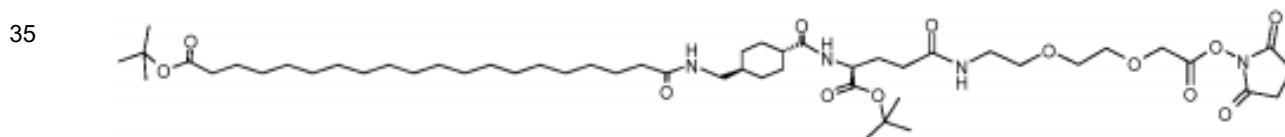
15 Cuando se usa en la presente descripción el término “contenido” en relación con un péptido y/o un agente de acilación alrededor de un grupo químico funcional, tal como un ácido carboxílico, significará un grupo químico funcional presente en el péptido o agente de acilación.

20 Los derivados de ácido carboxílico pueden prepararse eficientemente en una resina tal como resina de 2-clorotritlorcloruro mediante una estrategia basada en Fmoc mediante el uso de bloques de construcción protegidos adecuadamente mediante el uso de las condiciones de reacción descritas en (Fmoc Solid Phase Peptide Synthesis, Editado por W.C. Chan y P.D. White, Oxford University Press, 2000, ISBN 0-19-963724-5).

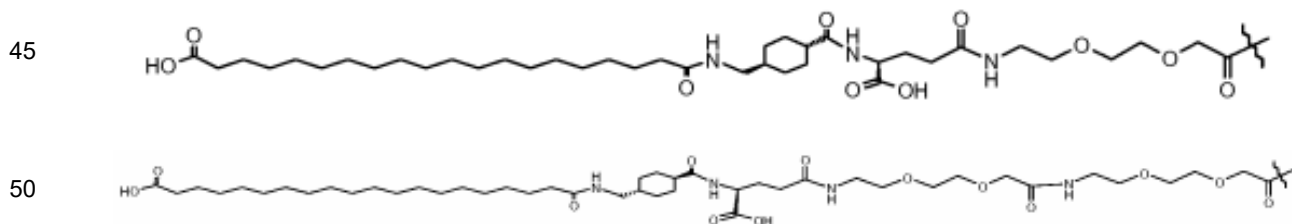
Un ejemplo de un fragmento A-B-C-D-OH es:

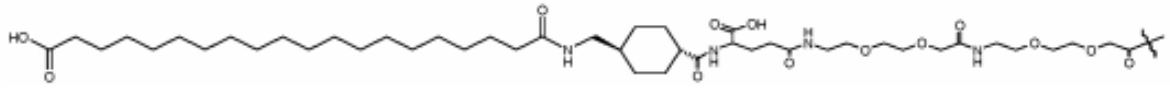


El análogo activado correspondiente podría ser el éster de N-hidroxisuccinimida;

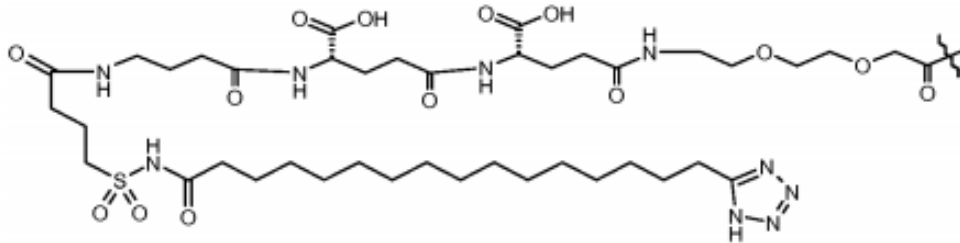


Ejemplos de fragmentos A-B-C-D-:



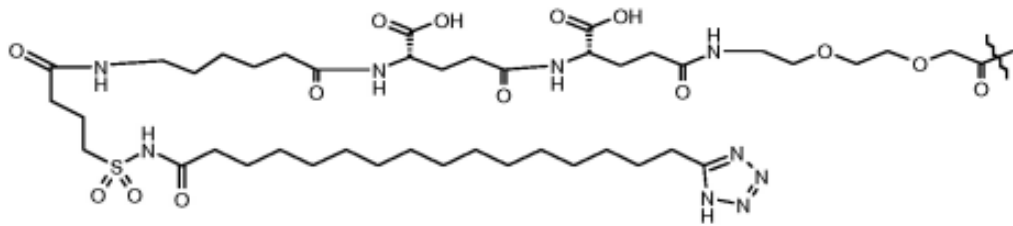


5



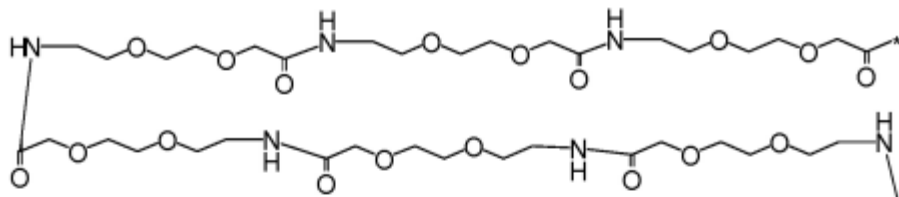
10

15



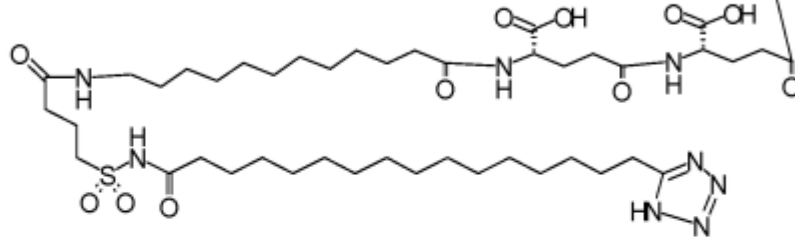
20

25



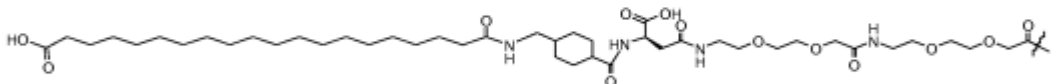
30

35

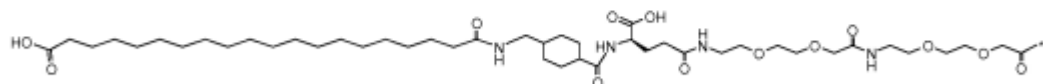


40

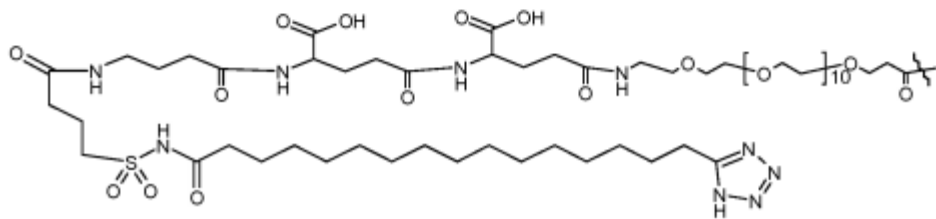
45



50



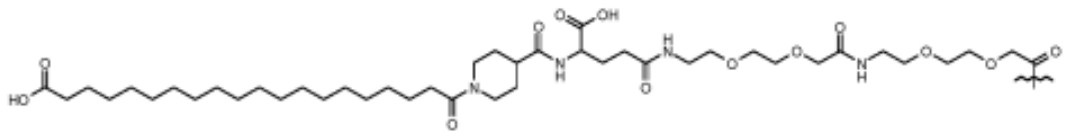
55



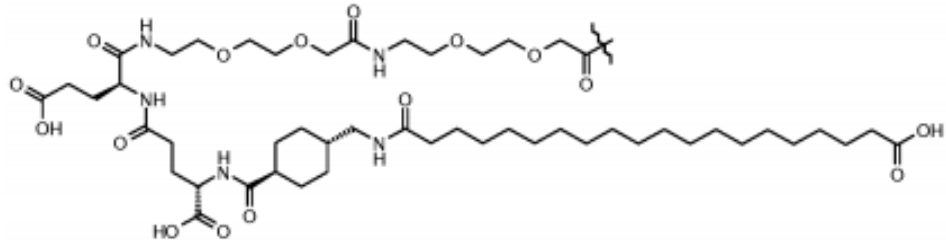
60

65

5



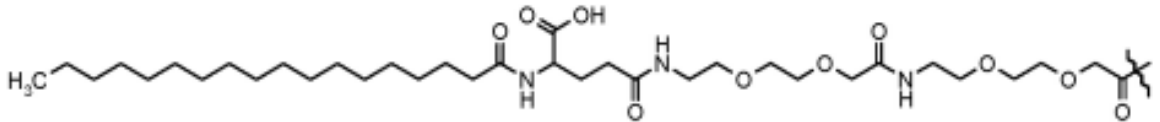
10



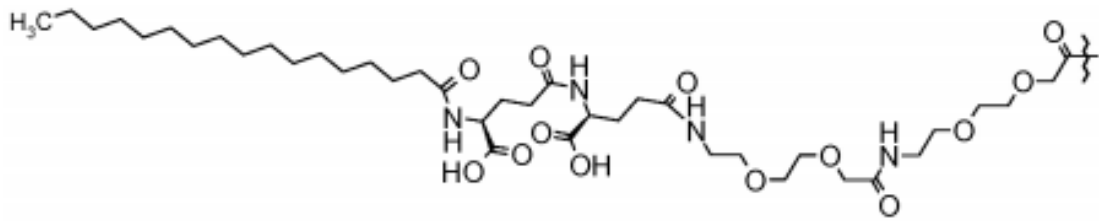
15

Ejemplos de fragmentos A-C-D-:

20



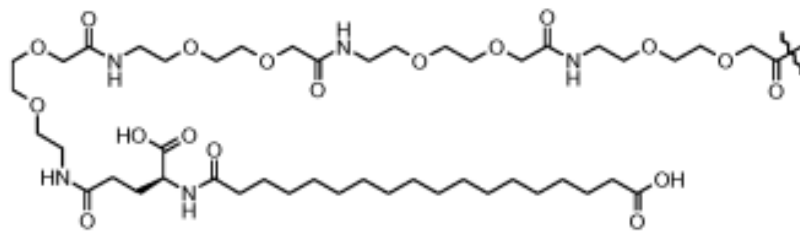
25



30

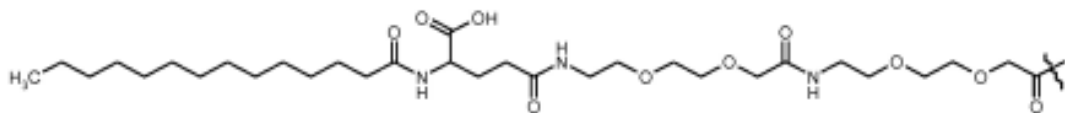
35

40

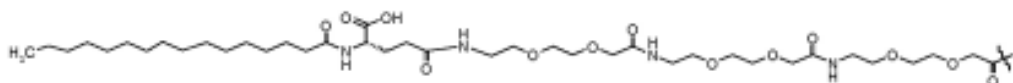


45

50

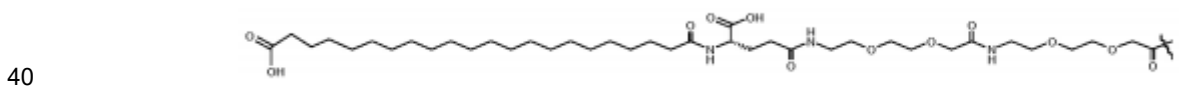
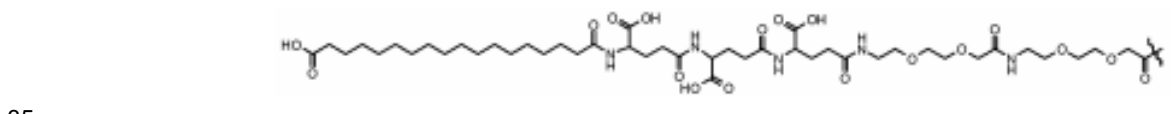
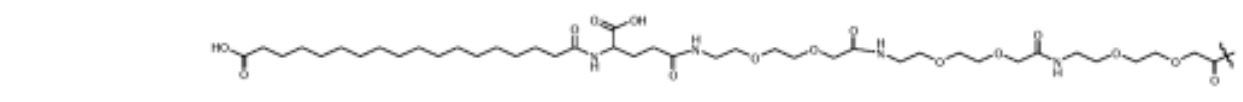
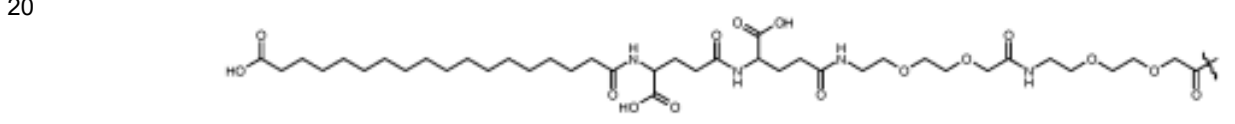
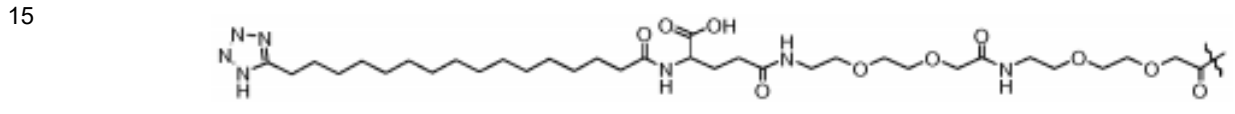
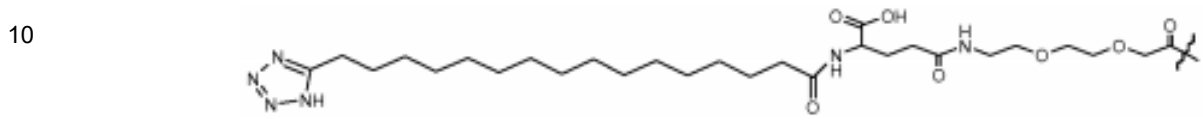
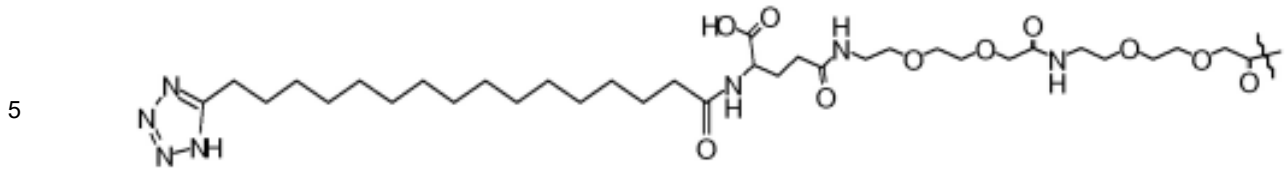


55

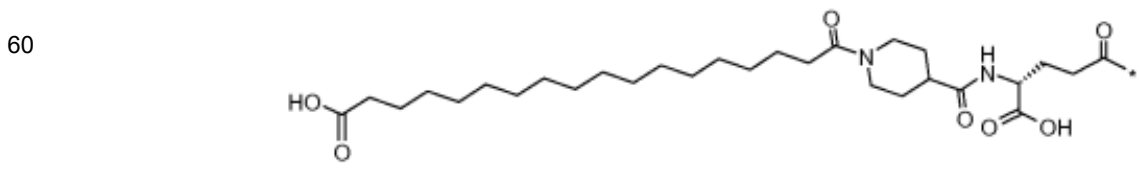
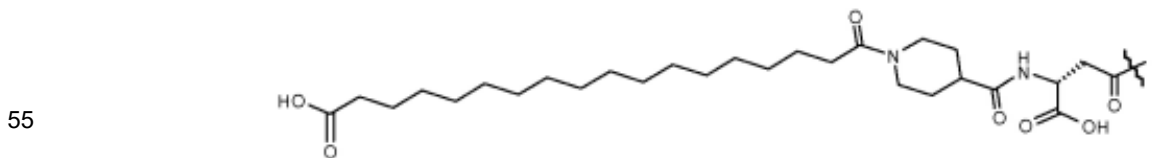
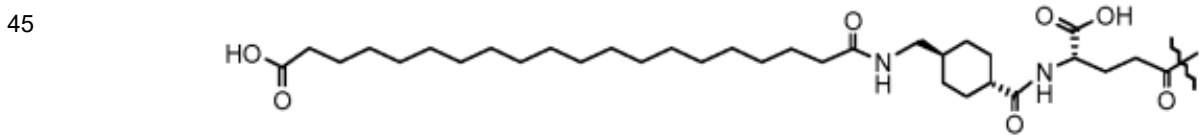


60

65



Ejemplos de fragmentos A-B-C-:



60

65

En todos los fragmentos específicos mencionados A-OH, A-C-D-OH, A-B-C-OH o A-B-C-D-OH cualquier ácido carboxílico libre contenido, puede protegerse opcionalmente con un grupo protector apropiado seleccionado de, pero sin limitarse a terc-butilo, bencilo o alilo.

5 En todos los fragmentos específicos mencionados A-OH, A-C-D-OH, A-B-C-OH o A-B-C-D-OH los centros quirales se especifiquen o no pueden ser el enantiómero R o S independiente de otros centros quirales en el mismo fragmento.

10 En una modalidad, cuando uno o más centros quirales están presentes, los fragmentos pueden estar en forma de una mezcla de enantiómeros.

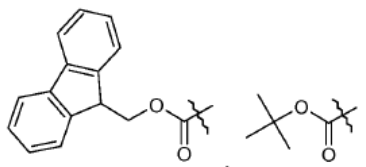
En otra modalidad, cuando uno o más centros quirales están presentes, los fragmentos pueden estar en forma de un enantiómero puro, en donde cada uno de los centros quirales son R o S.

15 En otra modalidad, la quiralidad es como se representa en los fragmentos de la presente invención mencionados específicamente.

En un aspecto el grupo de protección de amina épsilon de un residuo de aminoácido de lisina puede seleccionarse de, pero sin limitarse a, los siguientes ejemplos:

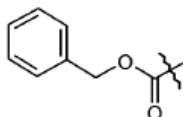
20

25



30 y

35



40 En el método de acuerdo con la invención, la etapa de acilación en ciertas modalidades útiles puede tener lugar en un solvente, que incluye un solvente polar orgánico seleccionado de 1-metil-pirrolidin-2-ona, acetonitrilo, tetrahidrofurano, dimetilsulfóxido, *N,N*-dimetilformamida y *N*-metil-formamida.

45 En una modalidad adicional, la etapa de acilación puede tener lugar en una mezcla de solventes acuosos, tal como una mezcla de solventes acuosos que comprende un solvente polar orgánico. Tal solvente polar orgánico puede seleccionarse ventajosamente de 1-metil-pirrolidin-2-ona, acetonitrilo, tetrahidrofurano, dimetilsulfóxido, *N,N*-dimetilformamida y *N*-metil-formamida.

50 En una modalidad adicional, la etapa de acilación puede tener lugar en solución acuosa.

55 La cantidad de 1-metil-pirrolidin-2-ona, acetonitrilo, tetrahidrofurano, dimetilsulfóxido y *N,N*-dimetilformamida o *N*-metil-formamida puede, en ciertas modalidades, ser de modo que la cantidad de dicho solvente sea menor que 80 % (v/v), tal como menor que 50 % (v/v). En otros aspectos la cantidad de 1-metil-pirrolidin-2-ona, acetonitrilo, tetrahidrofurano, dimetilsulfóxido, *N,N*-dimetilformamida o *N*-metil-formamida está en el intervalo de aproximadamente 40-80 % (v/v), que incluye el intervalo de 40-45, 45-50, 50-60, 60-70 y 70-80 % (v/v). Se contempla además que el agente de acilación puede agregarse a la mezcla de reacción como un sólido.

60 El valor de pH de la mezcla de reacción de solventes acuosos en donde tiene lugar la etapa de acilación, es en modalidades útiles entre 9 y 13, tal como entre 10 y 12, que incluye entre 10,5 y 11,5. En modalidades adicionales el pH en la mezcla de reacción está en el intervalo de 9-10, 10-11, 11-12 o 12-13.

65 El agente de acilación usado de acuerdo con la invención puede comprender uno o más grupos de protección. Por lo tanto, se contempla que el método de acuerdo con la invención puede comprender además una o más etapas de eliminar uno o más de estos grupos de protección a partir del agente de acilación una vez que se ha acoplado a la molécula precursora del análogo del GLP-1. Tales etapas de desprotección se describen bien en la literatura.

La expresión eficiente de proteínas heterólogas en las células huésped, frecuentemente requiere que la proteína se exprese en forma de un propéptido, es decir, un polipéptido que tiene una secuencia señal secretora (también conocida como secuencia líder, secuencia prepro o secuencia pre). Sin embargo, esta secuencia líder no se quiere normalmente en el polipéptido deseado, y por lo tanto el método de acuerdo con la invención puede, en ciertos aspectos, incluir una etapa adicional de eliminar las extensiones propeptídicas en el N-terminal y/o C-terminal de la molécula precursora del análogo del GLP-1.

El método de acuerdo con la invención comprende una etapa de aislar el análogo o derivado del GLP-1 resultante por medios adecuados, por ejemplo métodos de purificación estándar conocidos en la técnica.

También está dentro del alcance de la invención que los análogos del GLP-1 o derivados que resultan del método de acuerdo con la invención en ciertas modalidades pueden derivarse con una porción de unión a albúmina o pegilarse.

Las secuencias líderes de la presente invención pueden estar en la forma optimizada como se enumera en la figura 4.

El término "porción de unión a albúmina" como se usa en la presente descripción significa un residuo que se une no covalentemente a la albúmina sérica humana. El residuo de unión a albúmina unido al polipéptido terapéutico tiene típicamente una afinidad por debajo de 10 μM para la albúmina de suero humano y preferentemente por debajo de 1 μM . Se conoce un intervalo de residuos de unión a la albúmina entre las porciones lipófilas lineales y ramificadas que contienen de 4-40 átomos de carbono que tienen un grupo ácido distal.

Modalidades de acuerdo con la invención

1. Un método para producir un análogo o derivado del GLP-1 que comprende uno o más aminoácidos no proteinogénicos en la parte N-terminal, dicho método que comprende las etapas de

(i) cultivar una célula huésped que comprende una secuencia de nucleótidos que codifica una molécula precursora de dicho análogo del GLP-1 en condiciones adecuadas para la expresión de dicha molécula precursora,

(ii) separar la molécula precursora expresada del caldo de cultivo,

(iii) acoplar una extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos a la molécula precursora expresada,

(iv) aislar el análogo o derivado del GLP-1 resultante por medios adecuados.

2. Un método para producir un análogo o derivado del GLP-1 que comprende uno o más aminoácidos no proteinogénicos en la parte N-terminal, dicho método que comprende las etapas de:

i) cultivar una célula huésped que comprende una secuencia de nucleótidos que codifica una molécula precursora de dicho análogo del GLP-1 en condiciones adecuadas para la expresión de dicha molécula precursora,

ii) separar la molécula precursora expresada del caldo de cultivo,

iii) proteger el(los) grupo(s) de lisina(s) de la molécula precursora,

iv) acoplar una extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos a la molécula precursora protegida, para obtener un análogo del GLP-1 protegido,

v) desproteger el grupo de lisina de dicho análogo del GLP-1 protegido,

vi) aislar el análogo o derivado del GLP-1 resultante por medios adecuados.

3. Un método de acuerdo con las modalidades 1 o 2, en donde la extensión de aminoácidos N-terminal se protege antes de su uso en la etapa de acoplamiento, donde una extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos se acopla a la molécula precursora protegida y en donde el análogo del GLP-1 protegido se desprotege de nuevo después del acoplamiento de la extensión N-terminal.

4. Un método para producir un análogo o derivado del GLP-1 que comprende uno o más aminoácidos no proteinogénicos en la parte N-terminal, dicho método que comprende las etapas de:

i) cultivar una célula huésped que comprende una secuencia de nucleótidos que codifica una molécula precursora de dicho análogo del GLP-1 en condiciones adecuadas para la expresión de dicha molécula precursora,

ii) separar la molécula precursora expresada del caldo de cultivo,

iii) acoplar una extensión de aminoácidos N-terminal protegida que comprende uno o más aminoácidos no proteinogénicos a la molécula precursora expresada,

iv) desproteger cualquier grupo protector del análogo o derivado del GLP-1 protegido,

v) aislar el análogo o derivado del GLP-1 resultante por medios adecuados.

5. Un método para producir un análogo o derivado del GLP-1 que comprende uno o más aminoácidos no proteinogénicos en la parte N-terminal, dicho método que comprende las etapas de:

i) cultivar una célula huésped que comprende una secuencia de nucleótidos que codifica una molécula precursora de dicho análogo del GLP-1 en condiciones adecuadas para la expresión de dicha molécula precursora,

ii) separar la molécula precursora expresada del caldo de cultivo,

iii) proteger el(los) grupo(s) de lisina(s) de la molécula precursora,

iv) acoplar una extensión de aminoácidos N-terminal protegida que comprende uno o más aminoácidos no proteinogénicos a la molécula precursora protegida, para obtener un análogo o derivado del GLP-1 protegido,

v) desproteger cualquier grupo protector del análogo o derivado del GLP-1 protegido,

vi) aislar el análogo o derivado del GLP-1 resultante por medios adecuados.

6. Un método de acuerdo con cualquiera de las modalidades anteriores, en donde la molécula precursora de dicho análogo del GLP-1 se selecciona de la lista de moléculas precursoras que comprende la secuencia de aminoácidos de la fórmula general:

Thr-Phe-Thr-Ser-Asp-Xaa₁₆-Ser-Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Xaa₂₃-Ala-Xaa₂₅-Xaa₂₆-Xaa₂₇-Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Xaa₃₄-Xaa₃₅-Xaa₃₆-Xaa₃₇-Xaa₃₈-Xaa₃₉-Xaa₄₀-Xaa₄₁-Xaa₄₂-Xaa₄₃-Xaa₄₄-Xaa₄₅

en donde

Xaa₁₆ es Val o Leu;

Xaa₁₈ es Ser, Lys o Arg;

Xaa₁₉ es Tyr o Gln;

Xaa₂₀ es Leu o Met;

Xaa₂₂ es Gly o Glu;

Xaa₂₃ es Gln, Glu, Lys o Arg;

Xaa₂₅ es Ala o Val;

Xaa₂₆ es Lys, Glu o Arg;

Xaa₂₇ es Glu o Leu;

Xaa₃₀ es Ala, Glu o Arg;

Xaa₃₃ es Val, Lys, o Arg;

Xaa₃₄ es Lys, Glu, Asn, His o Arg;

Xaa₃₅ es Gly;

Xaa₃₆ es Arg, Gly o Lys;

Xaa₃₇ es Gly, Ala, Glu, Pro, Lys, o está ausente;

Xaa₃₈ es Lys, Ser o está ausente.

Xaa₃₉ es Ser, Lys, o está ausente;

Xaa₄₀ es Gly o ausente.

Xaa₄₁ es Ala o está ausente;

5 Xaa₄₂ es Pro o está ausente;

Xaa₄₃ es Pro o está ausente;

Xaa₄₄ es Pro o está ausente;

10 Xaa₄₅ es Ser o está ausente;

siempre y cuando si Xaa₃₈, Xaa₃₉, Xaa₄₀, Xaa₄₁, Xaa₄₂, Xaa₄₃, Xaa₄₄ o Xaa₄₅ está ausente, entonces cada residuo de aminoácido corriente abajo también está ausente.

15 7. Un método de acuerdo con cualquiera de las modalidades 1-5, en donde la molécula precursora de dicho análogo del GLP-1 se selecciona de la lista de moléculas precursoras que comprende la secuencia de aminoácidos de la fórmula general:

20 Xaa₉- Xaa₁₀-Thr-Phe-Thr-Ser-Asp-Xaa₁₆-Ser-Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Xaa₂₃-Ala-Xaa₂₅-Xaa₂₆-Xaa₂₇-Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Xaa₃₄-Xaa₃₅-Xaa₃₆-Xaa₃₇-Xaa₃₈-Xaa₃₉-Xaa₄₀-Xaa₄₁-Xaa₄₂-Xaa₄₃-Xaa₄₄-Xaa₄₅

en donde

25 Xaa₉ es Glu o Asp

Xaa₁₀ es Gly o Ala

Xaa₁₆ es Val o Leu;

30 Xaa₁₈ es Ser, Lys o Arg;

Xaa₁₉ es Tyr o Gln;

35 Xaa₂₀ es Leu o Met;

Xaa₂₂ es Gly o Glu;

Xaa₂₃ es Gln, Glu, Lys o Arg;

40 Xaa₂₅ es Ala o Val;

Xaa₂₆ es Lys, Glu o Arg;

45 Xaa₂₇ es Glu o Leu;

Xaa₃₀ es Ala, Glu o Arg;

Xaa₃₃ es Val, Lys, o Arg;

50 Xaa₃₄ es Lys, Glu, Asn, His o Arg;

Xaa₃₅ es Gly;

55 Xaa₃₆ es Arg, Gly o Lys;

Xaa₃₇ es Gly, Ala, Glu, Pro, Lys, o está ausente;

Xaa₃₈ es Lys, Ser o está ausente.

60 Xaa₃₉ es Ser, Lys, o está ausente;

Xaa₄₀ es Gly o ausente.

65 Xaa₄₁ es Ala o está ausente;

Xaa₄₂ es Pro o está ausente;

Xaa₄₃ es Pro o está ausente;

5 Xaa₄₄ es Pro o está ausente;

Xaa₄₅ es Ser o está ausente;

10 siempre y cuando si Xaa₃₈, Xaa₃₉, Xaa₄₀, Xaa₄₁, Xaa₄₂, Xaa₄₃, Xaa₄₄ o Xaa₄₅ está ausente, entonces cada residuo de aminoácido corriente abajo también está ausente.

8. Un método de acuerdo con cualquiera de las modalidades 1-5, en donde la molécula precursora de dicho análogo del GLP-1 se selecciona de la lista de moléculas precursoras que comprende la secuencia de aminoácidos de la fórmula general:

15 Xaa₈-Xaa₉-Xaa₁₀-Thr-Phe-Thr-Ser-Asp-Xaa₁₆-Ser-Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Xaa₂₃-Ala-Xaa₂₅-Xaa₂₆-Xaa₂₇-Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Xaa₃₄-Xaa₃₅-Xaa₃₆-Xaa₃₇-Xaa₃₈-Xaa₃₉-Xaa₄₀-Xaa₄₁-Xaa₄₂-Xaa₄₃-Xaa₄₄-Xaa₄₅

en donde

20 Xaa₈ es Ala, Gly, Val, Leu, Ile, Lys

Xaa₉ es Glu, Asp

25 Xaa₁₀ es Gly, Ala

Xaa₁₆ es Val o Leu;

Xaa₁₈ es Ser, Lys o Arg;

30 Xaa₁₉ es Tyr o Gln;

Xaa₂₀ es Leu o Met;

35 Xaa₂₂ es Gly o Glu;

Xaa₂₃ es Gln, Glu, Lys o Arg;

Xaa₂₅ es Ala o Val;

40 Xaa₂₆ es Lys, Glu o Arg;

Xaa₂₇ es Glu o Leu;

45 Xaa₃₀ es Ala, Glu o Arg;

Xaa₃₃ es Val, Lys, o Arg;

Xaa₃₄ es Lys, Glu, Asn, His o Arg;

50 Xaa₃₅ es Gly;

Xaa₃₆ es Arg, Gly o Lys;

55 Xaa₃₇ es Gly, Ala, Glu, Pro, Lys, o está ausente;

Xaa₃₈ es Lys, Ser o está ausente.

Xaa₃₉ es Ser, Lys, o está ausente;

60 Xaa₄₀ es Gly o ausente.

Xaa₄₁ es Ala o está ausente;

65 Xaa₄₂ es Pro o está ausente;

Xaa₄₃ es Pro o está ausente;

Xaa₄₄ es Pro o está ausente;

5 Xaa₄₅ es Ser o está ausente;

siempre y cuando si Xaa₃₈, Xaa₃₉, Xaa₄₀, Xaa₄₁, Xaa₄₂, Xaa₄₃, Xaa₄₄ o Xaa₄₅ está ausente, entonces cada residuo de aminoácido corriente abajo también está ausente.

10 9. Un método de acuerdo con cualquiera de las modalidades anteriores, en donde los aminoácidos no
 15 proteinogénicos en la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no
 proteinogénicos, se seleccionan del grupo que consiste en γ -carboxiglutamato, ornitina, fosfoserina, D-aminoácidos
 tales como D-alanina y D-glutamina, D-alanina y D-leucina, Aib (α -ácido α -aminobutírico), Abu (ácido α -
 20 aminobutírico), Tle (tert-butilglicina), ácido 3-aminometil benzoico, ácido antranílico, des-amino-histidina, β -alanina,
 D-histidina, desamino-histidina, 2-amino-histidina, β -hidroxi-histidina, homohistidina, N^o-acetil-histidina, α -fluorometil-
 histidina, α -metil-histidina, 3-piridilalanina, 2-piridilalanina, 4-piridilalanina, ácido (1-aminociclopropil) carboxílico,
 ácido (1-aminociclobutil) carboxílico, ácido (1-aminociclopentil) carboxílico, ácido (1-aminociclohexil) carboxílico,
 ácido (1-aminocicloheptil) carboxílico, ácido (1-aminociclooctil) carboxílico, α -metil prolina, 1-metil histidina, 3-metil
 histidina, y ácido 4,5,6,7-tetrahidro-1H-imidazo[4,5-C]piridina-6-carboxílico β -(1,2,4-triazol-1-yl)-alanina.

20 10. Un método de acuerdo con cualquiera de las modalidades anteriores, en donde la longitud de la extensión de
 aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos es de 1 aminoácido.

25 11. Un método de acuerdo con cualquiera de las modalidades 1-9, en donde la longitud de la extensión de
 aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos es 2 aminoácidos.

12. Un método de acuerdo con cualquiera de las modalidades 1-9, en donde la longitud de la extensión de
 aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos es 3 aminoácidos.

30 13. Un método de acuerdo con cualquiera de las modalidades 1-9, en donde la longitud de la extensión de
 aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos es de 4 aminoácidos.

14. Un método de acuerdo con cualquiera de las modalidades anteriores, en donde la extensión de aminoácidos N-
 terminal comprende dos aminoácidos no proteinogénicos.

35 15. Un método de acuerdo con cualquiera de las modalidades 1-13, en donde la extensión de aminoácidos N-
 terminal comprende tres aminoácidos no proteinogénicos.

40 16. Un método de acuerdo con cualquiera de las modalidades 1-13, en donde la extensión de aminoácidos N-
 terminal comprende cuatro aminoácidos no proteinogénicos.

17. Un método de acuerdo con cualquiera de las modalidades anteriores, en donde la extensión de aminoácidos N-
 terminal que comprende uno o más aminoácidos no proteinogénicos tiene la fórmula general

45 Xaa₇-Xaa₈- Xaa₉- Xaa₁₀

en donde

50 Xaa₇ se selecciona de L-histidina, D-histidina, desamino-histidina, 2-amino-histidina, β -hidroxi-histidina,
 homohistidina, N^o-acetil-histidina, α -fluorometil-histidina, α -metil-histidina, 3-piridilalanina, 2-piridilalanina y 4-
 piridilalanina, 1-metil histidina, 3-metil histidina, y ácido 4,5,6,7-tetrahidro-1H-imidazo[4,5-C]piridina-6-carboxílico β -
 (1,2,4-triazol-1-il)-alanina

55 Xaa₈ se selecciona de Ala, Gly, Val, Leu, Ile, Lys, Aib, ácido (1-aminociclopropil) carboxílico, ácido (1-
 aminociclobutil) carboxílico, ácido (1-aminociclopentil) carboxílico, ácido (1-aminociclohexil) carboxílico, ácido (1-
 aminocicloheptil) carboxílico, o ácido (1-aminociclooctil) carboxílico y alfa-metil prolina

Xaa₉ se selecciona de Glu, Asp, γ , γ -dimetil Glu, β , β -dimetil Glu y β , β -dimetil Asp, y

60 Xaa₁₀ se selecciona de Gly, Aib, ácido (1-aminociclopropil) carboxílico, ácido (1-aminociclobutil) carboxílico, ácido (1-
 aminociclopentil) carboxílico, ácido (1-aminociclohexil) carboxílico, ácido (1-aminocicloheptil) carboxílico, y ácido (1-
 aminociclooctil) carboxílico.

65 18. Un método de acuerdo con cualquiera de las modalidades 1-16, en donde la extensión de aminoácidos N-
 terminal que comprende uno o más aminoácidos no proteinogénicos tiene la fórmula general

Xaa₇-Xaa₈

en donde

5 Xaa₇ se selecciona de L-histidina, D-histidina, desamino-histidina, 2-amino-histidina, β-hidroxi-histidina, homohistidina, N^α-acetil-histidina, α-fluorometil-histidina, α-metil-histidina, 3-piridilalanina, 2-piridilalanina y 4-piridilalanina; 1-metil histidina, 3-metil histidina, ácido 4,5,6,7-tetrahidro-1H-imidazo[4,5-C]piridina-6-carboxílico β-(1,2,4-triazol-1-il)-alanina; y

10 Xaa₈ se selecciona de Gly, Aib, ácido (1-aminociclopropil) carboxílico, ácido (1-aminociclobutil) carboxílico, ácido (1-aminociclopentil) carboxílico, ácido (1-aminociclohexil) carboxílico, ácido (1-aminocicloheptil) carboxílico, o ácido (1-aminociclooctil) carboxílico y alfa metil prolina

15 19. Un método de acuerdo con cualquiera de las modalidades 1-16, en donde la extensión de aminoácidos N-terminal que comprende un aminoácido no proteinogénico que tiene la fórmula general

Xaa₇

en donde

20 Xaa₇ se selecciona de D-histidina, desamino-histidina, 2-amino-histidina, β-hidroxi-histidina, homohistidina, N^α-acetil-histidina, α-fluorometil-histidina, α-metil-histidina, 3-piridilalanina, 2-piridilalanina y 4-piridilalanina; 1-metil histidina, 3-metil histidina, y ácido 4,5,6,7-tetrahidro-1H-imidazo[4,5-C]piridina-6-carboxílico β-(1,2,4-triazol-1-il)-alanina.

25 20. Un método de acuerdo con la modalidad 19, en donde la extensión de aminoácidos N-terminal es desamino-histidina y la reacción de acoplamiento se lleva a cabo por una sal de 4-[2-(2,5-dioxo-pirrolidin-1-iloxycarbonil)-etil]-1H-imidazol-1-io, tal como trifluoroacetato, con la molécula precursora expresada.

30 21. Un método de acuerdo con cualquiera de las modalidades 1-5, en donde la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos tiene la fórmula general

Xaa₇-Xaa₈- Xaa₉- Xaa₁₀ como se define en la modalidad 17;

y el análogo del GLP-1 tiene la fórmula general

35 Thr-Phe-Thr-Ser-Asp-Xaa₁₆-Ser-Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Xaa₂₃-Ala-Xaa₂₅-Xaa₂₆-Xaa₂₇-Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Xaa₃₄-Xaa₃₅-Xaa₃₆-Xaa₃₇-Xaa₃₈-Xaa₃₉-Xaa₄₀-Xaa₄₁-Xaa₄₂-Xaa₄₃-Xaa₄₄-Xaa₄₅ como se define en la modalidad 3.

40 22. Un método de acuerdo con cualquiera de las modalidades 1-5, en donde la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos tiene la fórmula general

Xaa₇-Xaa₈ como se define en la modalidad 18;

45 y el análogo del GLP-1 tiene la fórmula general

50 Xaa₉-Xaa₁₀-Thr-Phe-Thr-Ser-Asp-Xaa₁₆-Ser-Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Xaa₂₃-Ala-Xaa₂₅-Xaa₂₆-Xaa₂₇-Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Xaa₃₄-Xaa₃₅-Xaa₃₆-Xaa₃₇-Xaa₃₈-Xaa₃₉-Xaa₄₀-Xaa₄₁-Xaa₄₂-Xaa₄₃-Xaa₄₄-Xaa₄₅ como se define en la modalidad 4.

55 23. Un método de acuerdo con cualquiera de las modalidades 1-5, en donde la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos tiene la fórmula general

Xaa₇ como se define en la modalidad 19;

60 y el análogo del GLP-1 tiene la fórmula general

60 Xaa₈- Xaa₉- Xaa₁₀-Thr-Phe-Thr-Ser-Asp-Xaa₁₆-Ser-Xaa₁₈-Xaa₁₉-Xaa₂₀-Glu-Xaa₂₂-Xaa₂₃-Ala-Xaa₂₅-Xaa₂₆-Xaa₂₇-Phe-Ile-Xaa₃₀-Trp-Leu-Xaa₃₃-Xaa₃₄-Xaa₃₅-Xaa₃₆-Xaa₃₇-Xaa₃₈-Xaa₃₉-Xaa₄₀-Xaa₄₁-Xaa₄₂-Xaa₄₃-Xaa₄₄-Xaa₄₅ como se define en la modalidad 4.

65 24. Un método de acuerdo con cualquiera de las modalidades 17-23, en donde el aminoácido C-terminal no proteinogénico o proteinogénico en las extensiones N-terminal Xaa₇-Xaa₈-Xaa₉-Xaa₁₀ o Xaa₇-Xaa₈, o el aminoácido no proteinogénico Xaa₇ se activa.

- 5 25. Un método de acuerdo con la modalidad 24, en donde el(los) aminoácido(s) activado(s) se selecciona(n) del grupo que consiste en: cloruro de ácido, bromuro de ácido, fluoruro de ácido, anhídrido simétrico, anhídrido mezclado, ácidos carboxílicos activados mediante el uso de carbodiimidas comunes, ácidos carboxílicos mediante el uso de carbodiimidas y un aditivo y ácidos carboxílicos activados con una sal de uronio o una sal de fosfonio.
- 10 26. Un método de acuerdo con la modalidad 24, en donde el(los) aminoácido(s) activado(s) se selecciona(n) del grupo que consiste en: ácidos carboxílicos activados mediante el uso de carbodiimidas comunes, diisopropilcarbodiimida (DIPCDI), N,N'-díciclohexilcarbodiimida (DCC), 1-etil-3-(3'-dimetilaminopropil)carbodiimida hidrocloreto (EDC).
- 15 27. Un método de acuerdo con la modalidad 24, en donde el(los) aminoácido(s) activado(s) se selecciona(n) del grupo que consiste en: ácidos carboxílicos con carbodiimidas y un aditivo seleccionado del grupo que consiste en *N*-hidroxibenzotriazol (HOBt), 1-Hidroxi-7-azabenzotriazol, 6-cloro-*N*-hidroxibenzotriazol (HOAt) y 3-Hidroxi-3,4-dihidro-4-oxo-1,2,3-benzotriazina (DhbToH),
- 20 28. Un método de acuerdo con la modalidad 24, en donde el(los) aminoácido(s) activado(s) se selecciona(n) del grupo que consiste en: ácidos carboxílicos activados con una sal de uronio o una sal de fosfonio, O-Benzotriazol-*N,N,N',N'*-tetrametil-uronio-hexafluoro-fosfato (HBTU), O-(Benzotriazol-1-il)-*N,N,N',N'*-tetrafluoroborato de -tetrametiluronio (TBTU), 2-(1*H*-7-Azabenzotriazol-1-il)-1,1,3,3-tetrametil uronio hexafluorofosfato (HATU), 1-benzotriazolioxitris(dimetilamino)fosfonio hexafluorofosfato (BOP) benzotriazol-1-il-oxitripirrolidinofosfonio hexafluorofosfato (PYBOP), ésteres de *N*-hidroxisuccinimida (éster NHS), pentafluorofenol éster (PfP-éster), 2,4-dinitrofenil éster, 4-nitrofenil éster, 3-hidroxi-3,4-dihidro-4-oxo-1,2,3-benzotriazina (HODhbt) Carbonildiimidazol (CDI), *N*-etil-5-fenilsoxazolium-3'-sulfonato (NEPIS).
- 25 29. Un método de acuerdo con cualquiera de las modalidades anteriores, en donde el acoplamiento de la extensión de aminoácidos *N*-terminal que comprende uno o más aminoácidos no proteinogénicos tiene lugar en un disolvente polar orgánico seleccionado del grupo que consiste en 1-metil-pirrolidin-2-ona, acetonitrilo, tetrahidrofurano, dimetilsulfóxido *N,N*-dimetilformamida y *N*-metil-formamida.
- 30 30. Un método de acuerdo con cualquiera de las modalidades anteriores, en donde el acoplamiento de la extensión de aminoácidos *N*-terminal que comprende uno o más aminoácidos no proteinogénicos tiene lugar en una mezcla de solventes acuosos.
- 35 31. Un método de acuerdo con la modalidad 30, en donde la mezcla de solventes acuosos comprende un solvente polar orgánico seleccionado del grupo que consiste en 1-metil-pirrolidin-2-ona, acetonitrilo, tetrahidrofurano, dimetilsulfóxido *N,N*-dimetilformamida y *N*-metil-formamida.
- 40 32. Un método de acuerdo con las modalidades 30 o 31, en donde el acoplamiento de la extensión de aminoácidos *N*-terminal que comprende uno o más aminoácidos no proteinogénicos tiene lugar en una mezcla de solventes acuosos, en donde el pH en la mezcla de reacción está entre 4 y 12, preferentemente entre 6 y 10.
- 45 33. Un método de acuerdo con cualquiera de las modalidades anteriores, en donde los aminoácidos en la extensión de aminoácidos *N*-terminal que comprende uno o más aminoácidos no proteinogénicos comprenden uno o más grupos de protección.
- 50 34. Un método de acuerdo con la modalidad 33, que comprende además la etapa de eliminar uno o más grupos de protección de dicha extensión de aminoácidos *N*-terminal que comprende uno o más aminoácidos no proteinogénicos.
- 55 35. Un método de acuerdo con la modalidad 1, que comprende además la etapa de acilar el grupo épsilon amino de al menos un residuo de lisina en la molécula precursora expresada con un agente de acilación.
36. Un método de acuerdo con la modalidad 35, en donde el agente de acilación es un análogo de ácido carboxílico de la fórmula general;
- A-OH, A-C-D-OH, A-B-C-OH o A-B-C-D-OH
- que se activa y/o se protege opcionalmente con uno o más grupos de protección,
- 60 en donde
- A se selecciona del grupo que consiste en;
- 65

5

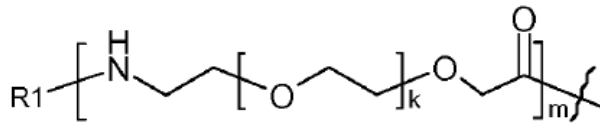
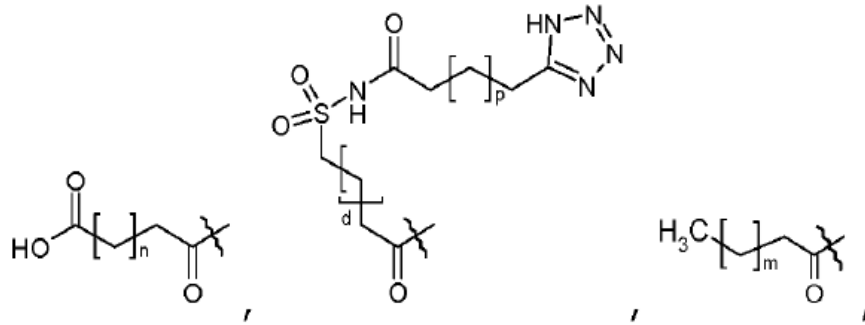
10

15

y

20

25



30

en donde n se selecciona del grupo que consiste en 14, 15, 16, 17, 18 y 19, p se selecciona del grupo que consiste en 10, 11, 12, 13 y 14, y d se selecciona del grupo que consiste en 0, 1, 2, 3, 4 y 5, y m se selecciona del grupo que consiste en 11, 12, 13, 14, 15, 16, 17, k se selecciona del grupo que consiste en 0, 1, 2, 3, 4, 5, 11 y 27, y m se selecciona del grupo que consiste en 0, 1, 2, 3, 4, 5 y 6, y R1 se selecciona del grupo de 9H-fluoren-9-ilmetoxicarbonilo (Fmoc), terc-butoxicarbonilo (Boc), bencilcarbarmata (Cbz).

35

B se selecciona del grupo que consiste en

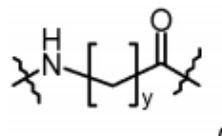
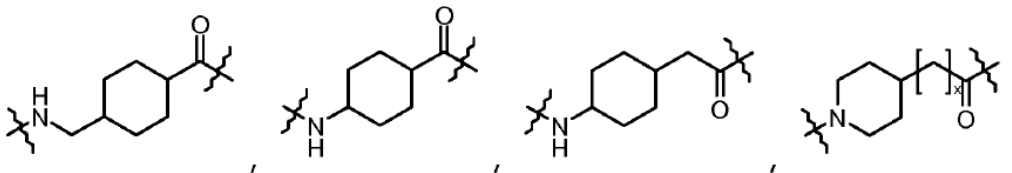
40

45

y

50

55



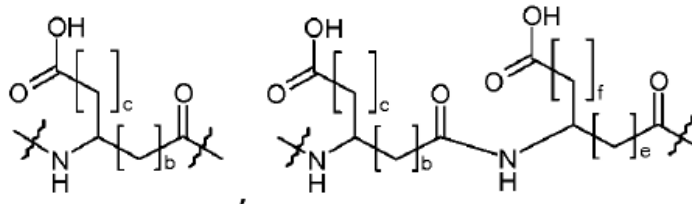
60

en donde x se selecciona del grupo que consiste en 0, 1, 2, 3 y 4, e y se selecciona del grupo que consiste en 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 y 12,

C se selecciona del grupo que consiste en

65

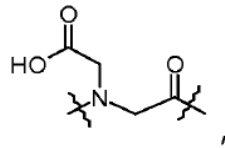
5



10

y

15



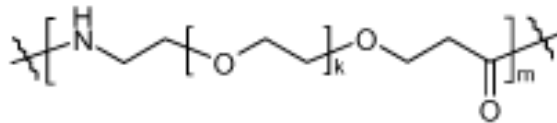
20

en donde b y e se seleccionan cada uno independientemente del grupo que consiste en 0, 1 y 2, y c y f se seleccionan cada uno independientemente del grupo que consiste en 0, 1 y 2 siempre y cuando b es 1 o 2 cuando c es 0, o b es 0 cuando c es 1 o 2, y e es 1 o 2 cuando f es 0, o e es 0 cuando f es 1 o 2, y

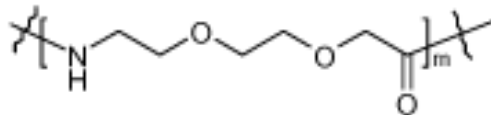
25

D se selecciona del grupo que consiste en

30

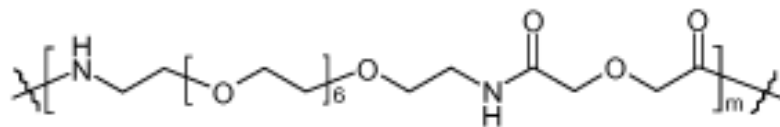


35



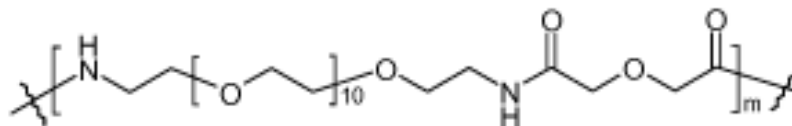
40

45

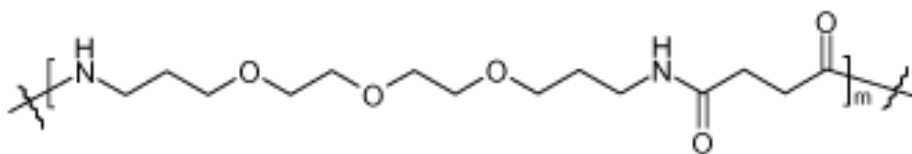


50

55

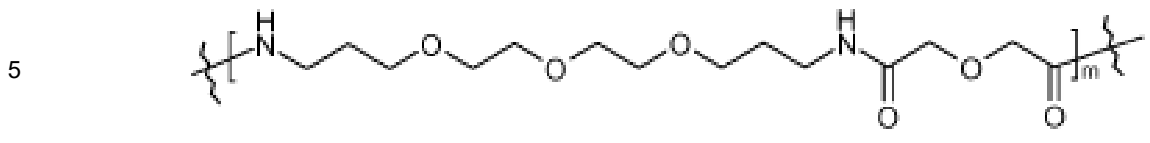


60



65

y



10 en donde x se selecciona del grupo que consiste en 0, 1, 2, 3, 4, 5, 11 y 27 y m se selecciona del grupo que consiste en 0, 1, 2, 3, 4, 5, y 6.

15 37. Un método de acuerdo con la modalidad 36, en donde el ácido carboxílico A-OH, A-C-D-OH, A-B-C-OH o A-B-C-D-OH se activa al agente de acilación activado.

18 38. Un método de acuerdo con la modalidad 35, en donde dicha etapa de acilación tiene lugar en una mezcla de solventes acuosos.

20 39. Un método de acuerdo con la modalidad 38, en donde dicha mezcla de solventes acuosos comprende un solvente polar orgánico.

23 40. Un método de acuerdo con la modalidad 39, en donde dicho solvente polar orgánico se selecciona del grupo que consiste en 1-metil-pirrolidin-2-ona, acetonitrilo, tetrahidrofurano, dimetilsulfóxido, *N,N*-dimetilformamida y *N*-metil-formamida.

25 41. Un método de acuerdo con la modalidad 40, en donde la cantidad de 1-metil-pirrolidin-2-ona, acetonitrilo, tetrahidrofurano, dimetilsulfóxido, *N,N*-dimetilformamida o *N*-metil-formamida es menor que 80 % (v/v).

28 42. Un método de acuerdo con la modalidad 40, en donde la cantidad de 1-metil-pirrolidin-2-ona, acetonitrilo, tetrahidrofurano, dimetilsulfóxido, *N,N*-dimetilformamida o *N*-metil-formamida es menor que 50 % (v/v).

30 43. Un método de acuerdo con la modalidad 40, en donde la cantidad de 1-metil-pirrolidin-2-ona, acetonitrilo, tetrahidrofurano, dimetilsulfóxido, *N,N*-dimetilformamida o *N*-metil-formamida está en el intervalo de aproximadamente 40-80 % (v/v).

33 44. Un método de acuerdo con la modalidad 35, en donde el agente de acilación se agrega a la mezcla de reacción como un sólido.

36 45. Un método de acuerdo con la modalidad 35, en donde dicha etapa de acilación tiene lugar en una mezcla de solventes acuosos, en donde el pH en la mezcla de reacción está entre 9 y 13, preferentemente entre 10 y 12 o incluso con mayor preferencia entre 10,5 y 11,5.

39 46. Un método de acuerdo con la modalidad 35, en donde el agente de acilación, que se activa opcionalmente, comprende uno o más grupos de protección.

42 47. Un método de acuerdo con la modalidad 46, que comprende además la etapa de eliminar uno o más grupos de protección de dicho agente de acilación activado.

45 48. Un método de acuerdo con cualquiera de las modalidades anteriores, que comprende además la etapa de eliminar una extensión propéptido N-terminal de la molécula precursora resultante.

48 49. Un método de acuerdo con cualquiera de las modalidades 1-47, que comprende además la etapa de eliminar una extensión propéptido C-terminal de la molécula precursora resultante.

51 50. Un método de acuerdo con cualquiera de las modalidades anteriores, en donde la célula huésped es una célula huésped de mamífero.

54 51. Un método de acuerdo con cualquiera de las modalidades 1-49, en donde la célula huésped es una célula huésped bacteriana.

57 52. Un método de acuerdo con cualquiera de las modalidades 1-49, en donde la célula huésped es una célula huésped de levadura.

60 53. Un método de acuerdo con la modalidad 52, en donde la célula huésped de levadura es *Saccharomyces cerevisiae*.

65

54. Un método de acuerdo con cualquiera de las modalidades anteriores, en donde el análogo del GLP-1 es un análogo o derivado del GLP-1(7-37) humano, GLP-2 humano, exendina-3 o exendina-4.

5 55. Un método de acuerdo con cualquiera de las modalidades anteriores, en donde el análogo del GLP-1 es un análogo o derivado del GLP-1(7-37) humano.

10 56. Un método de acuerdo con cualquiera de las modalidades anteriores, en donde dicho análogo del GLP-1 se selecciona del grupo que consiste en Aib⁸-GLP-1(7-36)-amida, Aib⁸-GLP-1(7-37), Aib^{8,35} GLP-1(7-37), [Aib⁸,Arg³⁴]GLP-1-(7-37), [DesaminoHis⁷, Glu²²,Arg²⁶,Arg³⁴,Lys³⁷]GLP-1-(7-37), [DesaminoHis⁷, Glu²²,Arg²⁶,Arg³⁴,Lys³⁷]GLP-1-(7-37)amida,

[DesaminoHis⁷,Arg²⁶,Arg³⁴,Lys³⁷]GLP-1-(7-37)amida,

15 [DesaminoHis⁷,Arg²⁶,Arg³⁴,Lys³⁷]GLP-1-(7-37)amida,

[DesaminoHis⁷, Glu²²,Arg²⁶,Arg³⁴,Lys³⁷]GLP-1-(7-37),

[DesaminoHis⁷,Arg²⁶,Arg³⁴,Lys³⁷]GLP-1-(7-37),

20 [DesaminoHis⁷, Glu²²,Arg²⁶, Glu³⁰,Arg³⁴,Lys³⁷]GLP-1-(7-37),

[DesaminoHis⁷, Glu²²,Arg²⁶,Arg³⁴,Lys³⁷]GLP-1(7-37)amida,

25 [DesaminoHis⁷,Arg³⁴]GLP-1-(7-37), [Aib⁸, Glu²²,Arg²⁶,Arg³⁴,Lys³⁷]GLP-1-(7-37)amida,

[DesaminoHis⁷, Glu²² Arg²⁶, Arg³⁴, Phe(m-CF₃)²⁸]GLP-1-(7-37)amida, [DesaminoHis⁷, Glu²²,Arg²⁶,Arg³⁴]GLP-1-(7-37)-Lys,

30 [DesaminoHis⁷, Glu²²,Arg²⁶,Arg³⁴]GLP-1-(7-37)-Lys,

[dDesaminoHis⁷,Arg²⁶,Arg³⁴,]GLP-1-(7-37)-Lys,

[DesaminoHis⁷, Glu²²,Arg²⁶,Arg³⁴,Lys³⁷]GLP-1-(7-37)amida,

35 y análogos de estos.

57. Un método de acuerdo con cualquiera de las modalidades anteriores, en donde el análogo del GLP-1 es la única sustancia biológicamente activa en la composición farmacéutica.

40 58. Un método de acuerdo con cualquiera de las modalidades anteriores, en donde el análogo del GLP-1 es un análogo de exendina-4.

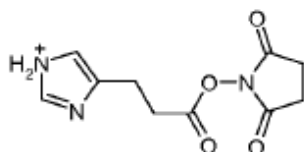
45 59. El método de acuerdo con cualquiera de las modalidades anteriores que comprende una etapa de proteger selectivamente el grupo épsilon amino de al menos un residuo de lisina.

60. El método de acuerdo con la modalidad 59 en donde la etapa se agrega después de la etapa (ii) que separa la molécula precursora expresada del caldo de cultivo.

50 61. El método de acuerdo con la modalidad 59 en donde la etapa se agrega después de la etapa (iv) como se define en la modalidad 1 que aísla el análogo o derivado del GLP-1 resultante por medios adecuados.

62. El método de acuerdo con la modalidad 19, en donde el aminoácido N-terminal es desamino-histidina introducido por el agente de acilación activado 4-{3-[(2,5-dioxo-1-pirrolidinil)oxi]-3-oxopropil}-1Himidazol-1-io

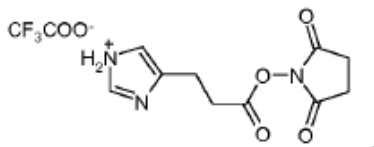
55



60

65 63. El método de acuerdo con la modalidad 19, en donde el aminoácido N-terminal es desamino-histidina introducido por el uso del agente acilado activado trifluoroacetato de 4-{3-[(2,5-dioxo-1-pirrolidinil)oxi]-3-oxopropil}-1Himidazol-1-io

5



10 64. El método de acuerdo con cualquiera de las modalidades anteriores, en donde el acoplamiento de la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos tiene lugar en un disolvente polar orgánico seleccionado del grupo que consiste en 1-metil-pirrolidin-2-ona, acetonitrilo, tetrahidrofurano, dimetilsulfóxido *N,N*-dimetilformamida y *N*-metil-formamida.

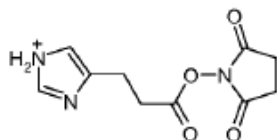
15 65. El método de acuerdo con cualquiera de las modalidades 1-63, en donde el acoplamiento de la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos tiene lugar en una mezcla de solventes acuosos.

20 66. El método de acuerdo con la modalidad 65, en donde la mezcla de solventes acuosos comprende un solvente polar orgánico seleccionado del grupo que consiste en 1-metil-pirrolidin-2-ona, acetonitrilo, tetrahidrofurano, dimetilsulfóxido *N,N*-dimetilformamida y *N*-metil-formamida.

25 67. El método de acuerdo con cualquiera de las modalidades 1-63, en donde el acoplamiento de la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos tiene lugar en una mezcla de solventes acuosos, en donde el pH en la mezcla de reacción está entre 7 y 11, preferentemente entre 8 y 10.

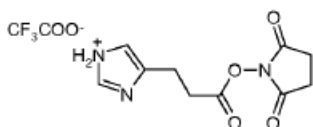
30 68. Un método para introducir des-aminohistidina en un agente insulínico con el uso del agente de acilación activado 4-{3-[(2,5-dioxo-1-pirrolidinil)oxi]-3-oxopropil}-1Himidazol-1-ilo:

35



40 69. Un método para introducir des-aminohistidina en un agente insulínico con el uso del agente acilado activado trifluoroacetato de 4-{3-[(2,5-dioxo-1-pirrolidinil)oxi]-3-oxopropil}-1Himidazol-1-ilo

45



50

La presente invención se describirá ahora en más detalles en los siguientes ejemplos y dibujos no limitantes.

EJEMPLOS

55

Abreviaturas usadas:

Adoc: adamantiloxicarbonilo

60 Aib: ácido α-aminoisobutírico

Boc terc butiloxicarbonilo

CH₃CN acetonitrilo

65

DCM: diclorometano

	DIC:	Diisopropilcarbodiimida
	DIPEA:	diisopropiletilamina
5	DMF:	<i>N,N</i> dimetilformamida
	EtOAc:	Etilacetato
10	Et ₂ O:	éter dietílico
	Fmoc:	9 H-fluoren-9-ilmetoxicarbonilo
	HATU:	hexafluorofosfato de 2-(7-azabenzotriazol-1-il)-1,1,3,3 tetrametiluronio
15	HBTU:	hexafluorofosfato de 2-(1H-benzotriazol-1-il)-1,1,3,3-tetrametiluronio
	H ₂ O:	agua
20	HOAt:	1-hidroxi-7-azabenzotriazol
	HOBt:	1-hidroxibenzotriazol
	MeCN:	acetonitrilo
25	Mtt:	4-Metiltrilito
	PM:	Peso molecular
30	NaOH:	Hidróxido de sodio
	NHS:	N-Hidroxisuccinimida
	NMP:	1-metil-pirrolidin-2-ona
35	OtBu:	éster terc butilo
	PyBop	hexafluorofosfato de (benzotriazol-1-iloxi)tripirrolidinofosfonio
40	PibroP	hexafluorofosfato de bromo-tris-pirrolidinofosfonio
	r.t:	temperatura ambiente
	OSu:	éster de N-hidroxisuccinimida
45	tBu:	terc butilo
	TBME:	terc-butilmetiléter
50	TEA:	triethylamina
	TFA:	ácido trifluoroacético
	TFFH	Hexafluorofosfato de tetrametilfluoroformamidinio
55	TIS:	triisopropilsilano
	Trt:	Tritilo, trifenilmetilo
60	TSTU:	tetrafluoroborato de O-(N-succinimidil)-1,1,3,3-tetrametiluronio

Métodos generales

Método general para la preparación de péptidos (método A)

65 El péptido puede sintetizarse de acuerdo con la estrategia de Fmoc en un sintetizador de péptidos 433 Applied Biosystems en escala de 0,25 mmol o 1,0 mmol mediante el uso de los protocolos FastMoc UV suministrados por el

fabricante que emplean acoplamientos mediados por HBTU hexafluorofosfato de (2-(1H-benzotriazol-1-il)-1,1,3,3-tetrametiluronio) o HATU (hexafluorofosfato de O- (7-azabenzotriazol-1-il)-1,1,3,3-tetrametiluronio) en NMP (N-metilpirrolidona), y monitoreo UV de la desprotección del grupo de protección Fmoc. Las resinas basadas en Wang o tritilo pueden usarse como soporte sólido y los derivados de aminoácidos protegidos usados son Fmoc-aminoácidos estándar (suministrados de por ejemplo Anaspec, o Novabiochem) suministrado en cartuchos previamente pesados adecuados para el sintetizador ABI433A. El aminoácido N terminal se protege con Boc en el grupo alfa amino (por ejemplo, se usó Boc-His(Boc)OH para los péptidos con His en el N-terminal). La síntesis de los péptidos puede mejorarse en algunos casos mediante el uso de dipéptidos protegidos en el enlace amida del dipéptido amídico con un grupo que puede escindirse en condiciones ácidas, tales como, pero sin limitarse a, 2-Fmoc-oxi-4-metoxibencilo o 2,4,6-trimetoxibencilo. En los casos en que está presente una serina o una treonina en el péptido, pueden usarse dipéptidos de pseudoprolina (ver, por ejemplo, el catálogo de Novabiochem 2002/2003 o una versión más reciente, o WR Sampson (1999), J. Pep. Sci. 5, 403).

Método general para la preparación de péptidos (método B)

Un método alternativo de síntesis de péptidos es mediante la química de Fmoc en un sintetizador de péptidos Liberty basado en microondas (CEM Corp., Carolina del Norte). Las resinas de Wang o basadas en tritil pueden usarse como el soporte sólido. La química de acoplamiento se realiza con DIC/HOAt en NMP mediante el uso de soluciones de aminoácidos de 0,4 M en NMP y un exceso molar de 8-10 veces. Las condiciones de acoplamiento son 5 minutos a 70 °C. La desprotección se logra con piperidina al 5 % en NMP hasta 70 °C.

Las moléculas precursoras del GLP-1 pueden por ejemplo purificarse mediante una variedad de procedimientos cromatográficos, por ejemplo cromatografía de intercambio iónico, cromatografía de interacción hidrofóbica, cromatografía de filtración en gel, cromatografía de afinidad, HPLC de fase reversa o similares, en dependencia del tipo de polipéptido en cuestión.

Método general para la preparación de agentes de acilación A-OH A-B-C-D-OH, A-C-D-OH, A-B-C-OH y extensiones de aminoácidos N-terminal Xaa₇-Xaa₈-Xaa₉-Xaa₁₀, Xaa₇-Xaa₈ y Xaa₇

Los ácidos carboxílicos necesarios pueden prepararse en una resina de 2-clorotritilcloruro mediante el uso de química de Fmoc estándar. El primer ácido amino carboxílico protegido con Fmoc puede unirse a la resina de 2-clorotritilcloruro mediante el hinchamiento de la resina en un solvente adecuado como NMP, DCM, DMF o similar, preferentemente DCM. El ácido amino carboxílico Fmoc protegido adecuadamente se agrega junto con una base adecuada tal como DIPEA o TEA, y la resina se agita durante un periodo de tiempo adecuado tal como 30 min a r.t. La resina se lava con un solvente adecuado tal como NMP, DMF o DCM.

La desprotección de Fmoc se logra mediante el uso de piperidina en NMP, preferentemente piperidina al 20 % en NMP a r.t. durante 1 a 30 min, típicamente 10 min, antes de lavarse bien la resina con NMP, DMF o DCM. La etapa se repite hasta que se obtiene la desprotección completa, típicamente 3 veces o más.

El acoplamiento del siguiente ácido carboxílico amino protegido con Fmoc se logra mediante el uso de condiciones de acoplamiento estándar; la resina se hincha en un solvente como NMP, DMF, DCM o una mezcla de estos. Otra solución de ácido carboxílico amino protegido con Fmoc en un solvente como NMP, DMF, DCM o una mezcla de estos y HOBt o HOAt se agrega DIC o un equivalente de estos. Esta mezcla se agrega a la resina y la mezcla se agrega DIPEA o TEA y la mezcla se agita a r.t. durante 1 a 16 horas típicamente 3 horas. Alternativamente, la mezcla se agita durante 1 a 60 min antes de agregar DIPEA o TEA a la mezcla. Si el acoplamiento no se completa de acuerdo con la prueba de TNBS, la prueba de nihdrina o la prueba de cloranil la etapa se repite hasta la prueba negativa.

Alternativamente, la activación del ácido carboxílico amino protegido con Fmoc puede lograrse mediante el uso de los siguientes reactivos de acoplamiento; PyBOP, PyBroP, HBTU, HATU o TFFH.

La incorporación adicional de ácidos carboxílicos Fmoc-amino protegidos adecuadamente puede introducirse mediante el uso del procedimiento proporcionado anteriormente. La última etapa de acoplamiento que comprende la incorporación de un derivado de ácido carboxílico se introduce mediante el uso de las condiciones de acoplamiento descritas anteriormente.

La escisión de la resina puede lograrse mediante el tratamiento de la resina con 20 % de trifluoroetanol en DCM, DCM-TIPS-TFA (95,5:2,5:2) o hexafluoroisopropanol de 5 min a 3 horas para dar el derivado totalmente protegido.

La escisión de la resina también puede lograrse mediante el uso de TFA-TIPS-Agua (95:2,5:2,5) de 5 min a 3 horas para proporcionar los derivados totalmente desprotegidos.

Alternativamente los derivados pueden prepararse en solución mediante el uso de procedimientos estándares descritos en la literatura.

Los bloques de construcción protegidos apropiadamente están disponibles comercialmente o pueden prepararse mediante el uso de procedimientos descritos en la literatura.

Xaa₇ está disponible comercialmente o puede prepararse mediante procedimientos descritos en la literatura.

Método general para la activación de agentes de acilación A-OH A-B-C-D-OH, A-C-D-OH, A-B-C-OH y extensiones de aminoácidos N-terminal Xaa₇-Xaa₈-Xaa₉-Xaa₁₀, Xaa₇-Xaa₈ y Xaa₇

La activación de los aminoácidos requeridos puede lograrse mediante el uso de procedimientos estándar descritos en la literatura. Si se requiere un éster activo tal como un *N*-hidroxisuccinimida éster, el ácido carboxílico se disuelve en un solvente apropiado tal como THF, acetato de etilo, NMP o DMF. Un reactivo tal como hexafluorofosfato de *O*-(*N*-succinimidilo)-*N,N,N,N'*-tetrametiluronio se agrega a la solución junto con una base tal como DIPEA o TEA. La mezcla se agita a r.t. hasta que la reacción se completa típicamente de 3 a 16 horas. Alternativamente el ácido carboxílico puede disolverse junto con *N*-hidroxisuccinimida y tratarse con DIC. La mezcla del producto puede usarse directamente sin purificación adicional o puede someterse a un tratamiento acuoso.

Purificación

El péptido bruto se purificó mediante HPLC preparativa en una columna empacada con sílice C18. Después del secado el péptido bruto se disolvió en ácido acético al 50 % en H₂O, se diluyó con H₂O y se inyectó en la columna que después se eluyó con un gradiente de CH₃CN de en agua y TFA al 0,1 % 10 ml/min durante 50 minutos a 40 °C. Las fracciones que contenían el péptido se recogieron y se liofilizaron después de la dilución con agua.

Análisis

LCMS

El LCMS se realizó en una configuración que consiste en un espectrómetro de masa de cuadrupolo simple Sciex API 100. El control del instrumento y la adquisición de datos se hicieron mediante el programa informático de control Sciex Sample que se ejecuta en una computadora con Windows 2000. La bomba de HPLC se conecta a dos recipientes de eluyente que contienen:

A: Ácido trifluoroacético al 0,05 % en agua

B: Ácido trifluoroacético al 0,05 % en acetonitrilo

El análisis se realizó a temperatura ambiente mediante la inyección de un volumen apropiado de la muestra (preferentemente 20 µl) en la columna que se eluyó con un gradiente de acetonitrilo.

Las condiciones del HPLC, los parámetros del detector y los parámetros del espectrómetro de masas usado se muestran en la tabla siguiente:

Columna: Waters Xterra MS C-18 X 3 mm id 5 µm

Gradiente: 5 % - 90 % de acetonitrilo lineal durante 7,5 min a 1,5 ml/min

Detección: 210 nm (salida análoga del DAD)

ELS (salida análoga del ELS) 40 °C

Modo de ionización de MS: API-ES

HPLC

Método 01_B4_2: El análisis de RP se realizó mediante el uso de un sistema Waters 600S equipado con un detector de matriz de diodos Waters 996. Las detecciones de UV a 214 nm y 254 nm se registraron mediante el uso de un Symmetry 300 C18, 5 µm, 3,9 mm x 150 mm, 42 °C. Se eluyó con un gradiente lineal de 5-95 % de acetonitrilo, 90-0 % de agua, y 5 % de ácido trifluoroacético (1,0 %) en agua durante 15 minutos a una velocidad de flujo de 1,0 ml/min.

Método 02_B4_4: El análisis de RP se realizó mediante el uso de un sistema Alliance Waters 2695 equipado con un detector de banda dual Waters 2487. Las detecciones de UV a 214 nm y 254 nm se registraron mediante el uso de un Symmetry300 C18, 5 µm, 3,9 mm x 150 mm, 42 °C. Se eluyó con un gradiente lineal de 5-95 % de acetonitrilo, 90-0 % de agua, y 5 % de ácido trifluoroacético (1,0 %) en agua durante 15 minutos a una velocidad de flujo de 1,0 ml/min.

Ejemplo 1

Construcción de sistemas de expresión de levaduras y producción de moléculas precursoras del análogo del GLP-1 (R34)GLP-1(8-37)

5 Los plásmidos de expresión son del tipo C-POT, similar a los descritos en el documento EP 171,142. Estos son vectores de expresión basados en 2 μ caracterizados por contener el gen de triosa fosfato isomerasa (POT) de *Schizosaccharomyces pombe* para el propósito de selección y estabilización del plásmido en *S. cerevisiae*. Los plásmidos también contienen el promotor y terminador de triosa fosfato isomerasa de *S. cerevisiae* (Figura 1). Estas secuencias son similares a las secuencias correspondientes en el plásmido PkFN1003 (descrito en el documento WO 90100075) como todas las demás secuencias excepto las siguientes: 1) La secuencia del fragmento EcoRI-XbaI que codifica la proteína de fusión del líder y el producto de insulina. 2) Se ha introducido una mutación silente que resulta en la eliminación de un sitio NcoI en la región 2 μ en el vector de expresión. Para facilitar la clonación de diferentes proteínas de fusión, la secuencia de ADN que codifica el líder pre-pro MF α 1 se ha cambiado para incorporar un sitio de NcoI (ver Figura 2) y se denomina líder pre-pro MF α 1*. Por lo tanto, el fragmento NcoI-XbaI se reemplaza simplemente por un fragmento NcoI-XbaI que codifica la molécula precursora del GLP-1 de interés. Tales fragmentos NcoI-XbaI pueden sintetizarse mediante el uso de oligonucleótidos sintéticos y PCR de acuerdo con técnicas estándares. Además del líder alfa, pueden usarse otros líderes.

20 Los fragmentos de ADN sintéticos que contienen secuencias que codifican (R34)GLP-1(11-37), (E22,R26,R34)GLP-1(8-37)K38 y (E22,R26,R34,K37)GLP-1(8-37) se obtuvieron de Genearth AG, BioPark, Josef-Engert-Str. 11, D-93053 Regensburg, Alemania. El ADN sintético que codifica (E22,R26,R34)GLP-1(8-37)K38 y (E22,R26,R34,K37)GLP-1(8-37) se proporcionó con secuencias de ADN 3' y 5' que codifican la extensión N- y C-terminal para facilitar la expresión en levadura (Figura 2) y se denomina Ext1-(E22,R26,R34)GLP-1(8-37)K38-Ext2 y Ext1-(E22,R26,R34,K37)GLP-1(8-37)-Ext2. El ADN sintético se digirió con NcoI y XbaI y se ligó al fragmento NcoI – XbaI del vector de expresión tipo cPOT modificado (Figura 1). Esto resultó en tres plásmidos de expresión Psa273, Psa277 y Psa278 que codifican (R34)GLP-1(11-37) (SEQ ID NO:6), Ext1-(E22,R26,R34)GLP-1(8-37)K38-Ext2 (SEQ ID NO:7) y Ext1-(E22,R26,R34,K37)GLP-1(8-37)-Ext2 (SEQ ID NO:8), respectivamente.

30 Los plásmidos de expresión se propagaron en *E. coli*, se cultivaron en presencia de ampicilina y se aislaron mediante el uso de técnicas estándar (Sambrook y otros, 1989). El ADN plasmídico se verificó para el inserto mediante nucleasas de restricción apropiadas (por ejemplo, *EcoRI*, *NcoI*, *XbaI*) y se mostró mediante el análisis de secuencias para contener la secuencia adecuada de las moléculas de los precursores del análogo del GLP-1 (R34)GLP-1(11-37), Ext1-(E22,R26,R34)GLP-1(8-37)K38-Ext2 y Ext1-(E22,R26,R34,K37)GLP-1(8-37)-Ext2. (Figura 2).

35 Los plásmidos se transformaron en *S. cerevisiae* cepa ME1719 (MATa/ α leu2/leu2 pep4-3/pep4-3 Δ tpi::LEU2/ Δ tpi::LEU2 Δ ura3/ Δ ura3 Δ yyp1::URA3 / Δ yyp1::ura3 Cir+). Esta cepa se describe en el documento WO 98/01535. Los transformantes de levadura que contienen el plásmido se seleccionaron por la utilización de glucosa como fuente de carbono en YPD (extracto de levadura al 1 %, peptona al 2 %, glucosa al 2 %) placas de agar (2 %). Esto resultó en cepas de tres levaduras: ySA251, ySA259 y ySA260 que expresan (R34)GLP-1(11-37) (SEQ ID NO:9), Ext1-(E22,R26,R34)GLP-1(8-37)K38-Ext2 (SEQ ID NO:10) y Ext1-(E22,R26,R34,K37)GLP-1(8-37)-Ext2 (SEQ ID NO:11), respectivamente.

45 Las cepas de levadura ME1719 ySA251, ySA259 y ySA260 se inocularon en 5 ml de medio de cultivo, por ejemplo un medio que consiste en 5 g/L (NH₄)₂DE SO₄, 184 mg/L (NH₄)₂HPO₄, 2,88 g/L KH₂PO₄, 1,42 g/L MgSO₄·7H₂O, 1,28 g/L, K₂DE SO₄, ácido succínico a 10,00 g/L, casaminoácidos a 10,00 g/L, FeSO₄·7H₂O de 0,0112 g/L, 0,0086 g/L de MnSO₄·H₂O, 0,0014 g/L de CuSO₄·5 H₂O, 0,00185 g/L de ZnSO₄·7H₂O, 0,0129 g/L de CaCl₂·2H₂O, ácido cítrico a 0,071 g/l, m-inositol a 28,0 mg/l, cloruro de colina a 14,0 mg/l, tiamina a 2,8 mg/l, niacinamida a 2,8 mg/l, ácido Ca-pantoténico a 2,1 mg/l, biotina a 0,14 mg/l, ácido fólico a 0,14 mg/l, glucosa a 40 g/l. El cultivo se llevó a cabo en 30 °C durante 3 días. Después de la centrifugación, se eliminó el sobrenadante para el análisis cuantitativo de HPLC mediante el cual se midió la concentración del análogo del GLP-1 secretado. La identidad de las moléculas precursoras del GLP-1 se confirmó por análisis LC/MS. Las extensiones N- y C-terminal empleadas para facilitar la expresión de (E22,R26,R34)GLP-1(8-37)K38 y (E22,R26,R34,K37)GLP-1(8-37) se eliminaron por Proteasa 1 de lisina específica de *Lyromobactor Lyticus* de acuerdo con procedimientos bien establecidos.

Ejemplo 2

60 Construcción del sistema de expresión de levaduras y producción de la molécula precursora del análogo del GLP-1 (R34)GLP-1(9-37)

65 Un fragmento de ADN sintético que codifica (R34)GLP-1(9-37) se construyó por PCR estándar mediante el uso de los cebadores 5'-AGGGGTATCCATGGCTAAGAGAGAAGGTACCTTCACCTCTGAC-3' y 5'-AATCTTAGTTTCTAGAGCCTGCG-3' y un plásmido que contiene la secuencia de ADN que codifica (R34)GLP-1(7-37). Los cebadores se diseñaron con un sitio 5' NcoI y un sitio 3' XbaI. Por lo tanto, el fragmento de PCR podría digerirse por NcoI y XbaI después de la purificación y ligarse al fragmento del vector NcoI – XbaI del vector de

expresión de tipo cPOT modificado descrito en el Ejemplo 1. El plásmido resultante que codifica (R34)GLP-1(9-37) se denominó Psa82.

5 El plásmido de expresión se propagó en *E. coli*, se cultivó en presencia de ampicilina y se aisló mediante el uso de técnicas estándar (Sambrook y otros, 1989). El ADN plasmídico se verificó para el inserto por nucleasas de restricción apropiadas (por ejemplo, *EcoRI*, *NcoI*, *XbaI*) y se mostró mediante el análisis de secuencias para contener la secuencia adecuada de la molécula precursora del análogo del GLP-1 (R34)GLP-1(9-37).

10 El plásmido se transformó en la cepa ME1719 de *S. cerevisiae* y un transformante de levadura que contiene el plásmido se seleccionó como se describe en el Ejemplo 1. La cepa de levadura resultante que expresa (R34)GLP-1(9-37) se denominó ySA96.

15 La cepa de levadura se cultivó en medio de cultivo y (R34)GLP-1(9-37) se recuperó del medio como se describe en el Ejemplo 1.

Ejemplo 3

Construcción de sistemas de expresión en levaduras para precursores del GLP-1

20 El líder y los sitios de escisión Kex2p se optimizaron para aumentar el rendimiento de expresión del péptido secretado y procesado. Los fragmentos de ADN sintético que codifican varios líderes y constructos (R34)GLP-1(11-37), o (R34)GLP-1(9-37) se amplificaron por PCR estándar. Los cebadores directos incluyeron un sitio *NcoI* y la secuencia que codifica la secuencia líder optimizada. Los cebadores inversos incluyeron un sitio *XbaI*. Por lo tanto, permitiendo la clonación de los fragmentos de PCR restringidos por *NcoI-XbaI* en el vector de expresión tipo cPOT restringidos por *NcoI-XbaI*. Consulte la figura 4 para ver las secuencias, la numeración del plásmidos de los constructos de expresión que codifican (R34)GLP-1 (11-37) y los nombres de las cepas.

25 Los plásmidos de expresión se propagaron en *E. coli* en presencia de ampicilina y se aislaron mediante el uso de técnicas estándar (Sambrook y otros, 1989). Los constructos de plásmidos se verificaron por digestión de endonucleasa de restricción mediante el uso de enzimas de restricción apropiadas. Las secuencias codificantes se verificaron por secuencia.

30 Los plásmidos se transformaron en la cepa ME1719 de *S. cerevisiae*. Las células de levadura que contienen el plásmido se seleccionaron como se describe en el Ejemplo 1. El cultivo de las cepas de levadura fue esencialmente como se describe en el Ejemplo 1

35 La optimización de la secuencia en la cercanía del sitio de escisión Kex2p, P1 a P6, condujo a un aumento modesto en el rendimiento (figura 4: B). Se exploró la incorporación de una Prolina en la secuencia (P3 a P6) como un medio potencial para aumentar la exposición del sitio de escisión Kex2p a Kex2p. Aunque, esto se ha vuelto así para no ser el caso para (R34)GLP-1(11-37) puede observarse un efecto beneficioso para (R34)GLP-1(11-37) (figura 4:H, I). La variación de los residuos P7 a P11 tuvo un efecto pronunciado sobre el rendimiento de la expresión, se observaron aumentos de 2-2,5 veces cuando se introdujeron motivos que contienen aminoácidos cargados (Figura 4: M-Z, K5-7, L2-4). Se ha diseñado una serie de constructos para el procesamiento con DAP-1 (dipeptidil aminopeptidasa) en el proceso corriente abajo (figura 4: G, J, K, L). Los constructos J y L contienen una Q en la posición antes de Thr-11, para proteger contra la escisión promiscua por parte de DAP-1 del péptido GLP-1. Esta Gln puede eliminarse posteriormente mediante el uso de las enzimas Q-ciclasa y pGAPasa (Qiagen).

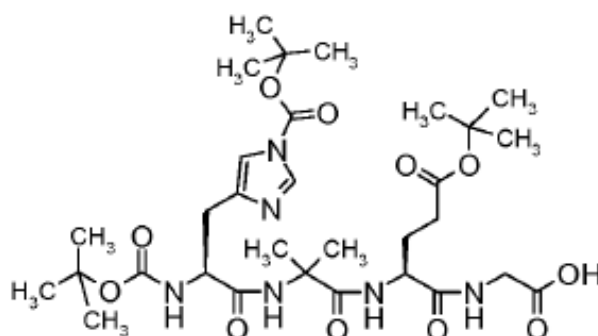
Tabla 1.

SECUENCIA	SECUENCIA REF.	RENDIMIENTO	PLÁSMIDO	NOMBRE DE LA CEPA
•RARYKR	(A)	25	pISNN546	SCI50
•RDLGKR	(B)	125	pISNN547	SCI51
•RDLAKR	(C)	88	pISNN548	SCI52
•RARAKR	(D)	58	pISNN549	SCI53
•RALDKR	(E)	58	pISNN550	SCI54
•RALAKR	(F)	43	pISNN551	SCI55
•PRDLGKR	(H)	68	pISNN585	SCI57
•RPLGKR	(I)	58	pISNN586	SCI58
•RDLGKREA	(G)	200	pISNN589	SCI56
•RDLGKREAQ	(J)	188	pISNN599	SCI59

	•RDLGKREAEA	(K)	240	pISNN590	SCI60
5	•RDLGKREALEKR	(K5)	n.d.	pISNN654	
	•RDLGRREALEKR	(K6)	n.d.	pISNN655	
	•RDLGEALEKR	(K7)	100	pISNN652	
10	•RDLGKREAEAQ	(L)	253	pISNN587	SCI44
	•RDLGKREAEQKR	(L2)	25	pISNN650	
	•RDLGRREAEQKR	(L3)	n.d.	pISNN656	
	•RDLGEAEQKR	(L4)	13	pISNN651	
15	•ERLERDLGKR	(M)	225	pISNN610	SCI63
	•KERLERDLGKR	(N)	100	pISNN611	SCI64
	•ERLEKR	(O)	138	pISNN617	SCI65
	•KERLEKR	(P)	125	pISNN612	SCI66
20	•PERLERDLGKR	(Q)	73	pISNN613	SCI67
	•PERLEKR	(R)	125	pISNN614	SCI68
	•EAEARDLGKR	(S)	175	pISNN615	SCI69
25	•PEAEARDLGKR	(T)	100	pISNN616	SCI70
	•EEAEKR	(U)	125	pISNN623	SCI71
	•EEAERDLGKR	(V)	218	pISNN624	SCI72
30	•RDLGEEAEKR	(X)	218	pISNN625	SCI73
	•EEAELAKR	(Y)	165	pISNN626	SCI74
	•EEAELGKR	(Z)	105	pISNN627	SCI75
35	•KR		100	pISNN545	SCI43

Ejemplo 4

Preparación de la extensión N-terminal Boc-His(Boc)-Aib-Glu(O-tBu)-Gly-OH



Se prehinchó la resina de 2-clorotrilcloruro (1 % de DVB, 1,4 mmol/g, 10 g, 14,0 mmol) en DCM antes de agregar una solución de Fmoc-Gly-OH (8,3 g, 28 mmol) y DIPEA (9,03 g, 70 mmol) en 100 ml de DCM, y se agitó durante 30 min a r.t.. La resina se lavó con NMP (2 x 100 ml) y 2 x 100 ml de DCM-MeOH-DIPEA (80:15:5, 10 min) después NMP (3 x 100 ml), antes del tratamiento con piperidina al 20 % en NMP (3 x 100 ml) durante 10 min cada una. La resina se lavó con NMP (6 x 100 ml), DCM (2 x 100 ml) y MeOH (2 x 100 ml) y se secó durante la noche *al vacío*. La resina se trató con 100 ml de piperidina al 20 % en NMP durante 10 min. La etapa se repitió dos veces antes de lavarse con NMP (5 x 100 ml).

Fmoc-Glu(O-tBu)-OH (23,8 g, 56 mmol) se disolvió en una mezcla de 100 ml de NMP y 20 ml de DCM. Se agregó HOAt (7,62 g, 56 mmol) seguido por adición en forma de gotas de DIC (7,07 g, 56 mmol). La mezcla se agitó durante

15 min antes de agregar DIPEA (9,03 g, 70 mmol), y la mezcla se transfirió a la resina. La resina se agitó durante 16 h antes de lavarse con NMP (4 x 100 ml). La resina se trató con piperidina al 20 % en NMP (100 ml) durante 10 min. La etapa se repitió dos veces antes de lavarse con NMP (5 x 100 ml).

5 Fmoc-Aib-OH (18,22 g, 56 mmol) se disolvió en una mezcla de 100 ml de NMP y 20 ml de DCM. Se agregó HOAt (7,62 g, 56 mmol) seguido por adición en forma de gotas de DIC (7,07 g, 56 mmol). La mezcla se agitó durante 20 min antes de agregarla a la resina. La mezcla se agitó durante 30 min antes de agregar DIPEA. La mezcla se agitó durante 16 h antes de lavar la resina con NMP (4 x 100 ml). La resina se trató con piperidina al 20 % en NMP (100 ml) durante 10 min. La etapa se repitió dos veces antes de lavar con NMP (5 x 100 ml).

10 Se disolvió Boc-His(Boc)-OH (19,9 g, 56 mmol) en una mezcla de 100 ml de NMP y 20 ml de DCM. Se agregó HOAt (7,62 g, 56 mmol) seguido por adición en forma de gotas de DIC (7,07 g, 56 mmol). La mezcla se agitó durante 20 min antes de agregarla a la resina. La mezcla se agitó durante 30 min antes de agregar DIPEA. La mezcla se agitó durante 16 h antes de lavar la resina con NMP (3 x 100 ml), DCM (4 x 100 ml).

15 Se agregó trifluoroetanol-DCM (1:4, 125 ml) y la resina se agitó durante 1 hora. Se recogió el filtrado y se agregó una nueva parte de trifluoroetanol-DCM (1:4, 125 ml) y se agitó durante 15 min. Se recogió el filtrado y se retiró el solvente *al vacío* para dar un aceite claro. Se agregó éter de dietilo frío (300 ml) en donde se formó un precipitado blanco. Se agregó éter de petrol claro (100 ml) y el precipitado se filtró, se lavó con éter de dietilo y se secó *al vacío*.

20 Rendimiento 7,3 g.

HPLC (Método 01_B4): Rt = 7,8 min

25 LCMS: m/z = 944 (M+H)⁺

PM calculado = 943,2

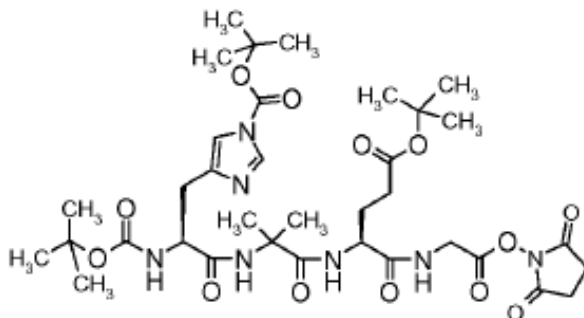
Ejemplo 5

30 Preparación de la extensión N-terminal Boc-His(Boc)-Aib-Glu(O-tBu)-Gly-OSuc

35

40

45



50 Se agregó una solución de Boc-His(Boc)-Aib-Glu(O-tBu)-Gly-OH (1,0 g, 1,47 mmol) en THF seco (60 ml), DIPEA (0,47 g, 3,66 mmol) y hexafluorofosfato de *O*-(*N*-succinimidilo)-*N,N,N,N*-tetrametiluronio (1,10 g, 3,66 mmol) y la mezcla de reacción se agitó durante 24 h a temperatura ambiente. Se agregó diclorometano (250 ml) y la solución se lavó con H₂O (2 x 100 ml). La fase orgánica se secó sobre MgSO₄, se filtró y se concentró *al vacío* para proporcionar 1,1 g de Boc-His(Boc)-Aib-Glu(O-tBu)-Gly-OSu.

55 HPLC (Método 01_B4): Rt = 8,3 min

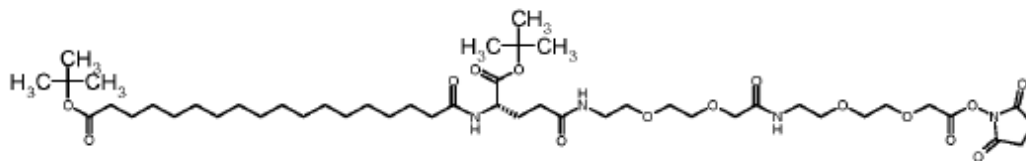
LCMS: m/z = 780,9 (M+H)⁺

60 PM calculado = 779,9

Ejemplo 6

65 Preparación del agente de acilación éster ter-butilo del ácido 17-((*S*)-1-terc-Butoxicarbonil-3-{2-[2-((2-[2-(2,5-dioxopirrolidin-1-iloxicarbonilmetoxi)etoxi]etilcarbamoil)metoxi)etoxi]etilcarbamoil]propilcarbamoil]heptadecanoico

5



10 El compuesto del título se preparó mediante el método general para la preparación y activación de agentes de acilación A-OH A-B-C-D-OH, A-C-D-OH, A-B-C-OH como se describió anteriormente y similar al ejemplo 4.

HPLC (Método 01_B4):Rt = 16,2 min

15 LCMS:m/z = 944 (M+H)⁺

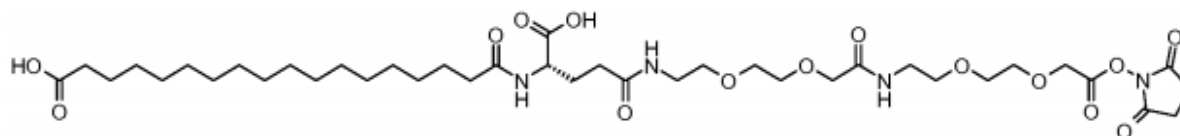
PM calculado = 943,2

Ejemplo 7

20

Preparación del agente de acilación ácido 17-((S)-1-carboxi-3-{2-[2-({2-[2-(2,5-dioxo-pirrolidin-1-ilo)carbonilmetoxi]-etoxi]-etilcarbamoil}-metoxi)-etoxi]-etilcarbamoil}-propilcarbamoil)-heptadecanoico

25



30

El éster de terc-butilo del ácido 17-((S)-1-terc-butoxicarbonil-3-{2-[2-({2-[2-(2,5-dioxopirrolidin-1-ilo)carbonilmetoxi]etoxi]etilcarbamoil} metoxi)etoxi]-etilcarbamoil}-propilcarbamoil) heptadecanoico se disolvió en TFA y la solución resultante se agitó durante 2 h y se evaporó *al vacío* hasta la sequedad. El aceite resultante se coevaporó con tolueno tres veces hasta la sequedad lo que dio lugar a un residuo aceitoso

35

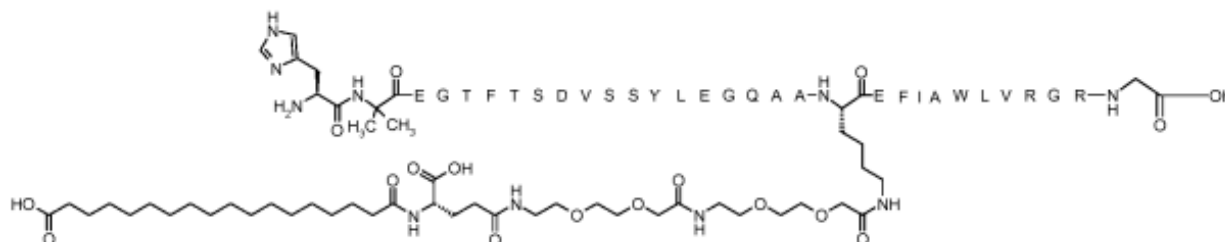
LCMS: m/z = 831 (M+H)⁺

Ejemplo 8

40

Preparación del derivado del GLP-1 [Arg34]GLP-1-(11-37)-N-épsilon26-[2-(2-[2-(2-[2-((S)-4-carboxi-4-(17-carboxiheptadecanoilamino) butirilamino]etoxi]etoxi)acetilamino]etoxi]etoxi)acetil] [Aib8,Arg34,]GLP-1-(7-37).

45



50

55 [Arg34]GLP-1-(11-37) (50 mg, 0,017 mmol) se disolvió en 2 ml de agua y se agregó DIPEA (89 µl, 0,51 mmol). La mezcla se agitó durante 10 min antes de que el pH se midiera a 10,6. Luego una solución del éster terc-butilo del ácido 17-((S)-1-terc-butoxicarbonil-3-{2-[2-({2-[2-(2,5-dioxopirrolidin-1-ilo)carbonilmetoxi]-etoxi]-etilcarbamoil} metoxi)etoxi]etilcarbamoil}-propilcarbamoil)- heptadecanoico (20,5 mg, 0,022 mmol) en MeCN (1.5 ml) se agregó en forma de gotas durante 10 min. La mezcla de reacción se agitó durante 1 hora antes de eliminar MeCN *al vacío* y la solución se liofilizó.

60

El residuo resultante se disolvió en NMP (2 ml) en donde se agregó DIPEA (22 µl, 0,17 mmol) y Boc-His(Boc)-Aib-Glu(OtBu)-Gly-OSu (26 mg, 0,033 mmol). La mezcla de reacción se agitó durante 4 horas antes de agregar una parte adicional de His(Boc)-Aib-Glu(OtBu)-Gly-OSu (13 mg 0,017 mmol) y la mezcla de reacción se agitó durante la noche. Se agregó éter de dietilo frío (20 ml) en donde se formó un precipitado blanco. El precipitado se aisló mediante centrifugación y el sobrenadante se decantó. El precipitado se lavó con 10 ml de éter de dietilo y se secó en el aire.

65

Desprotección

El intermedio bruto se disolvió en TFA-triisopropilsilano-H₂O (95:2,5:2,5, 10 ml) y se agitó durante 3 h, en donde la solución se concentró *al vacio* a la aplicación de 1 ml. Se agregó éter de dietilo frío (20 ml) en donde se formó un precipitado. El precipitado se aisló mediante centrifugación y el sobrenadante se decantó. El precipitado se lavó con éter dietílico y se secó. El compuesto bruto se disolvió en ácido acético-H₂O (3:7) (10 ml) y se purificó en HPLC.

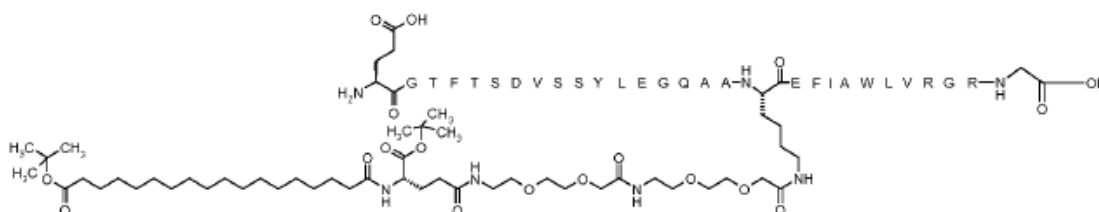
HPLC (Método 02_B4): Rt = 9,6 min

LCMS: m/z = 1372,5 (M+3H)³⁺

PM calculado = 4113,7

Ejemplo 9

Preparación del derivado del GLP-1 péptido N-épsilon26-[2-(2-[2-(2-[2-(2-[4-(17-*terc*-butoxicarbonilheptadecanoil-amino)-4(S)-*terc*-butoxicarbonilbutirilamino]etoxi)etoxi]acetilamino)etoxi]etoxi]acetil]-[Arg34]GLP-1-(9-37).



A una solución del éster de *terc*-butilo del ácido 17-((S)-1-*terc*-butoxicarbonil-3-[2-(2-[[2-(2-carboxiloxietoxi)etilcarbamoil]metoxi]etoxi)etilcarbamoil]propilcarbamoil] heptadecanoico) (94,0 mg, 0,112mmol) en NMP (2,0 ml) se agregó hexafluorofosfato de O-(1-succinimidil)-N,N,N',N'-tetrametiluronio (TSTU, 41,7 mg, 0,139mmol) y TEA (96,0 µl, 0,69 mmol). Después de agitar durante 30 min. a temperatura ambiente la mezcla se agregó en forma de gotas a una solución del péptido [Arg34]GLP-1-(9-37) (220 mg, 0,069 mmol) en agua (9 ml) y TEA (96,0 µl, 0,69 mmol). Luego de agitarse durante 30 min adicionales a temperatura ambiente, la mezcla se diluyó con agua (30 ml) y se liofilizó durante 16 h dando una mezcla oleosa/amorfa que se disolvió en una mezcla de ácido acético-acetonitrilo-agua (30:20:50, 180 ml), se filtró y purificó por HPLC para dar péptido N-épsilon26-[2-(2-[2-(2-[2-(2-[4-(17-*terc*-butoxicarbonilheptadecanoil-amino)-4(S)-*terc*-butoxicarbonilbutirilamino]etoxi)etoxi]acetilamino)etoxi]etoxi]acetil]-[Arg34]GLP-1-(9-37).

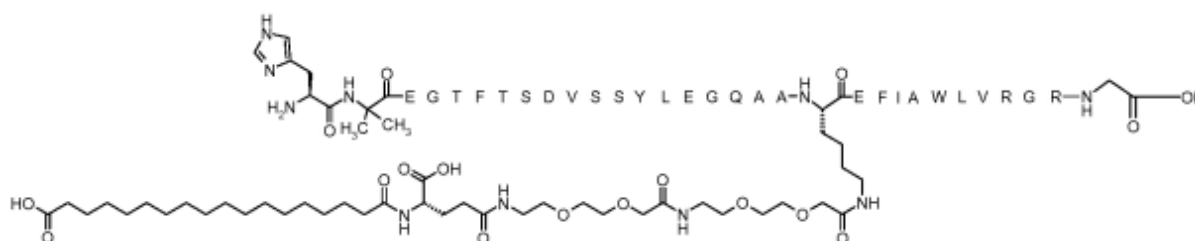
HPLC (Método 02_B4_4): Rt = 12,1 min

LCMS: m/z = 1335,5 (M+3H)³⁺

PM calculado = 4003,6

Ejemplo 10

Preparación del análogo del GLP-1 péptido N-épsilon26-[2-(2-[2-(2-[2-(2-[(S)-4-Carboxi-4-(17-carboxi-heptadecanoilamino)-butirilamino]etoxi)etoxi]acetilamino]-etoxi)etoxi]-acetil] [Aib8,Arg34]GLP-1-(7-37).



A una solución de Boc-His(Boc)-Aib-OH (23 mg, 0,052 mmol) en NMP (2,0 ml) se agregaron hexafluorofosfato de 2-(1H-9-azabenzotriazol-1-il)-1,1,3,3-tetrametiluronio (HATU, 20 mg, 0,052 mmol) y TEA (29 µl, 0,208 mmol). Luego de agitarse durante 45 min. se agregó la mezcla a una solución de N-épsilon26-[2-(2-[2-(2-[2-(2-[4-(17-*terc*-butoxicarbonilheptadecanoilamino)-4(S)-*terc*-butoxicarbonilbutirilamino]etoxi)etoxi]acetilamino)-etoxi]etoxi]acetil]-péptido [Arg34]GLP-1-(9-37) (83 mg, 0,021mmol) y TEA 3,6 ul, 0,026 mmol). Después de agitarse durante 14 días la

mezcla de reacción se diluyó con agua (100 ml) y se purificó mediante HPLC. Las fracciones purificadas se agruparon y se retiró el solvente orgánico *al vacío* antes de liofilizar el residuo durante 16 h.

Al residuo se agregó una mezcla de trifluoroacético-triisopropilsilano-agua (94:3:3,10 ml). Luego de agitar durante 2 horas, la mezcla de reacción se evaporó *al vacío*, el residuo se disolvió en una mezcla de amoniaco-agua (99:1,100 ml) y se purificó mediante HPLC preparativa para dar el compuesto del título.

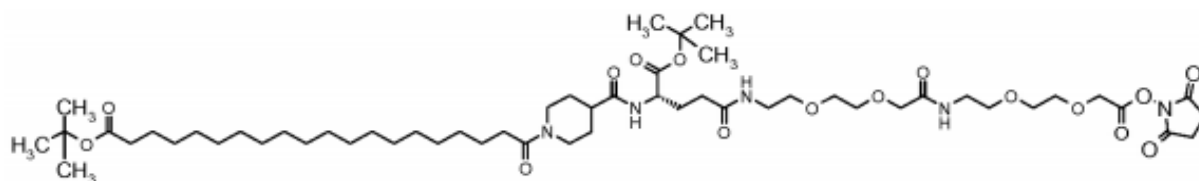
HPLC (Método 02_B4_4): Rt = 9,4 min

LCMS: m/z = 1372 (M+3H)³⁺

PM calculado = 4113,7

Ejemplo 11

Preparación del agente de acilación éster terc-butilo del ácido 20-[4-((S)-1-tert-Butoxicarbonil-3-{2-[2-({2-[2-(2,5-dioxo-pirrolidin-1-iloxycarbonilmetoxi)-etoxi]etilcarbamoil})metoxi]etoxi]etilcarbamoil)-propilcarbamoil]-piperidin-1-il]-20-oxo-icosanoico



El compuesto del título se preparó mediante el método general para la preparación y activación de agentes de acilación A-OH A-B-C-D-OH, A-C-D-OH, A-B-C-OH como se describió anteriormente y similar al ejemplo 4.

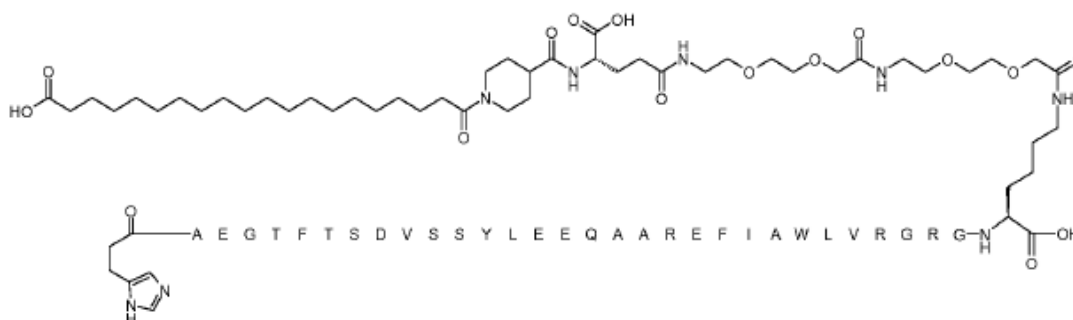
LCMS: m/z = 1111 (M+H)⁺

LCMS: m/z = 1082,6 (M+H)⁺

PM calculado = 1082,4

Ejemplo 12

Preparación del análogo del GLP-1 péptido [Des amino-His7,Glu22,Arg26,34,Lys37(Lys(2-{2-[2-({2-[2-((S)-4-Carboxi-4-[[1-(19-carboxinonadecanoil)piperidina-4-carbonil]amino]-butirilamino)etoxi]-etoxi]-acetilamino)etoxi]etoxi]-acetil)-GLP-1(7-37)



A una solución del péptido [Glu22,Arg26,34,Lys37]GLP-1-(8-37) (46 mg, 0,013 mmol) en 2 ml de agua se agregó DIPEA (50,4 mg, 0,39 mmol) y la mezcla se agitó durante 10 min antes de agregar en forma de gotas una solución del éster terc-butilo del ácido 20-[4-((S)-1-tert-butoxicarbonil-3-{2-[2-({2-[2-(2,5dioxo-pirrolidin-1-iloxycarbonilmetoxi)-etoxi]-etilcarbamoil})metoxi]-etoxi]-etilcarbamoil)-propilcarbamoil]-piperidin-1-il]-20-oxo-icosanoico (32,4 mg, 0,03 mmol) en un 1 ml de MeCN en el transcurso de 15 min. La mezcla de reacción se agitó por 20 min antes de agregar agua (20 ml), y la solución se liofilizó.

El residuo resultante se disolvió en NMP (1 ml) y la mezcla del ácido 4-[2-(2,5-dioxo-pirrolidin-1-iloxycarbonil)-etil]-imidazol-1-carboxílico adamantan-1-il éster (éster de Adoc-des amino His-OSu) se agitó durante 3 horas antes de

agregar en forma de gotas a dietil éter frío (20 ml). El precipitado blanco se aisló mediante centrifugación y se lavó dos veces con éter de dietilo (20 ml) y se secó.

5 El residuo resultante se trató con TFA-TIPS-agua (95:2,5:2,5, 1 ml) durante 3 horas antes de agregar en forma de gotas a dietil éter frío (10 ml). El precipitado blanco se aisló mediante centrifugación y se lavó dos veces con éter de dietilo (10 ml). El residuo se purificó mediante HPLC preparativa para dar el compuesto del título

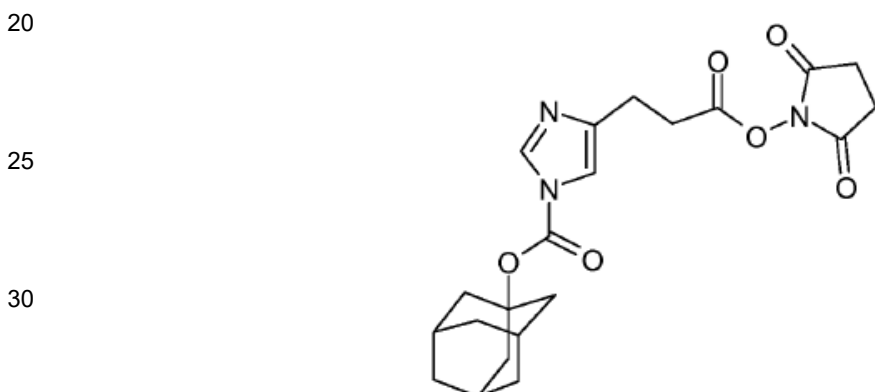
HPLC (Método 01_B4_2): Rt = 11.1 min

10 LCMS: m/z = 1484 (M+3H)³⁺

PM calculado = 4452,1

Ejemplo 13

15 Preparación del ácido 4-[2-(2,5-dioxo-pirrolidin-1-iloxicarbonil)etil]imidazol-1-carboxílico adamantan-1-il éster



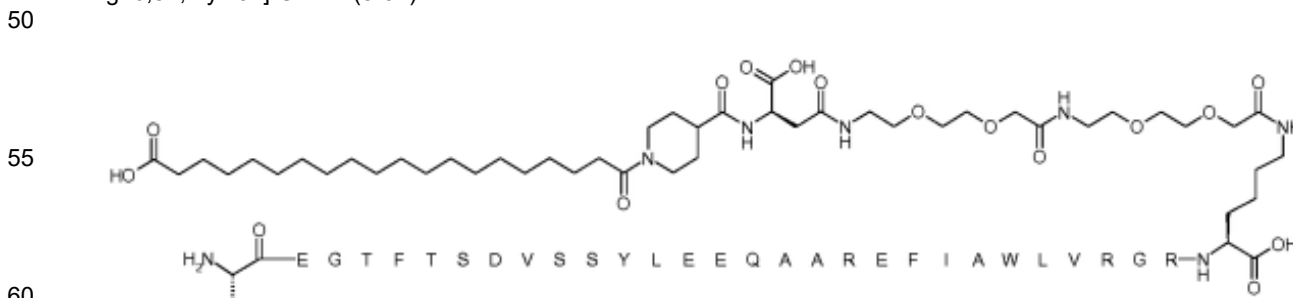
35 Se disolvió ácido 4-(2-carboxietil)imidazol-1-carboxílico adamantan-1-il éster (1 g, 3,1 mmol) en THF (5 ml) y DMF (10 ml). La solución se enfrió hasta 0 °C y se agregaron TSTU y DIEA. La solución se agitó a 0 °C durante 1 h y a temperatura ambiente durante 16 h. La muestra se concentró al vacío. EtOAc (100 ml) se agregó, y la solución se lavó con 0,2 N de HCl (2 x 50 ml), se secó sobre MgSO₄, y se concentró al vacío para producir un residuo cristalino pegajoso (1,18 g, 90 % de rendimiento). El producto bruto se usó sin purificación adicional.

40 LCMS:m/z = 416,2 (M+H)

¹H-NMR (DMSO, 300 Mhz) (señales seleccionadas) δ 8,22 (s, 1H), 7,37 (s, 1H), 3,02 (t, 2H), 2,82-2,88 (m, 2H), 2,81 (s, 4H), 2,19 (s-br, 9H), 1,66 (s, 6H).

45 Ejemplo 14

50 Preparación del análogo del GLP-1 péptido N-épsilon37{2-[2-(2-[2-((R)-3-carboxi-3-{{1-(19-carboxinadecanoil)piperidina-4-carbonil]amino}propionilamino)etoxi]etoxi]acetilamino)etoxi]etoxi]acetil [Glu22, Arg26,34, Lys 37] GLP-1 (8-37)



65 El análogo del GLP-1 se preparó mediante el uso de un sintetizador de péptido CEM Liberty que comienza con una resina de Wang Fmoc-Lys(Mtt). Después que se eliminó el grupo Fmoc de Ala8, el N terminal se protegió al realizar un doble acoplamiento con di-*tert*-butil dicarbonato. El grupo Mtt se eliminó mediante el lavado de la resina con DCM, el tratamiento de la resina con hexafluoroisopropanol durante 15 min a r.t y el lavado de la resina con DCM. Las modificaciones de N-épsilon 37 se prepararon en el sintetizador de péptido CEM Libert mediante el uso de

reactivos protegidos con Fmoc y ácido eicosanodioico mono-terc-butil éster. El péptido bruto se escindió de la resina TFA-TIPS-Agua (95:2,5:2,5) y se precipitó en éter frío, y se aisló mediante centrifugación. El péptido bruto se purificó mediante HPLC preparativa (4 cm de diámetro x 200 mm, C18, 60 ml/min, 33-53 % de acetonitrilo).

5 LCMS: m/z = 1065,8 (M+4H)⁴⁺

Ejemplo 15

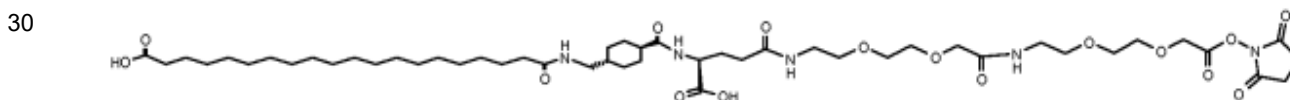
10 Preparación del análogo del GLP-1 péptido N-épsilon 37{2-[2-(2-[2-((R)-3-carboxi-3-[[1-(19-carboxinonadecanoil)piperidina-4-carbonil]amino}propionilamino)etoxi]etoxi)acetilamino)etoxi]etoxi]acetil [Des amino-His(Adoc) 7, Glu22, Arg26,34, Lys 37]GLP-1(7-37)

15 El análogo del GLP-1 péptido N-épsilon37{2-[2-(2-[2-((R)-3-carboxi-3-[[1-(19-carboxinonadecanoil)piperidina-4-carbonil]amino}propionilamino)etoxi]etoxi)acetilamino)etoxi]etoxi]acetil[Glu22,Arg26,34, Lys 37]GLP-1 -(8-37) (2,1 mg, 0,001 mmol) se disolvió en NMP (105 µL). Una solución de ácido 4-[2-(2,5-dioxo-pirrolidin-1-iloxicarbonil)etil]imidazol-1-carboxílico adamantan-1-il éster se preparó al disolver 25,3 mg en 1296 µl de NMP, y 25 µl de esta solución se agregó a la solución del péptido seguido por DIEA (0,9 µl). La conversión al producto se siguió por análisis de LCMS. Después de 2 h a r.t., se observó que la relación de la señal de ELS del producto al péptido inicial fue 976:28 (97 % de conversión).

20 LCMS: m/z = 1141,1 (M+4H)⁴⁺

Ejemplo 16

25 Preparación del agente de acilación ácido 19-[[trans-4-((S)-1-Carboxi-3-{2-[2-({2-[2-(2,5-dioxo-pirrolidin-1-iloxicarbonil)metoxi]-etoxi]-etilcarbamoil}-metoxi)-etoxi]-etilcarbamoil}-propilcarbamoil)-ciclohexilmetil]-carbamoil]-nonadecanoico



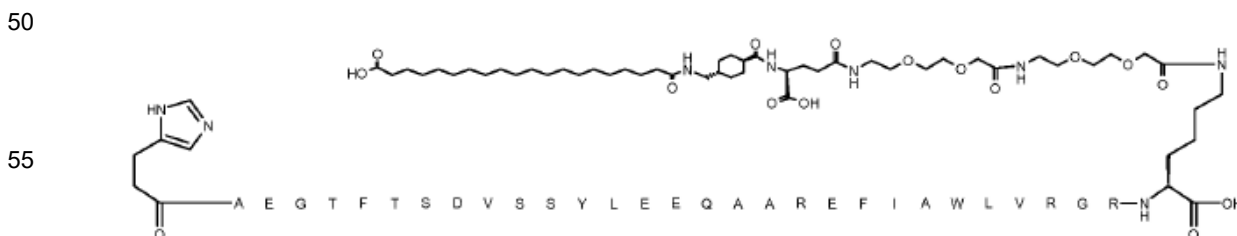
35 Ácido 19-[[trans-4-((S)-1-terc-butoxicarbonil-3-{2-[2-({2-[2-(2,5-dioxo-pirrolidin-1-iloxi carbonil)metoxi]-etoxi]-etilcarbamoil}-metoxi)-etoxi]-etilcarbamoil}-propilcarbamoil)-ciclohexilmetil]-carbamoil]-nonadecanoico terc-butil éster (1,0 g, 0,9 mmol) se agitó en TFA (10 ml) durante 1h 45 min antes de agregar 100 ml de éter de dietilo frío. El precipitado blanco se aisló mediante centrifugación y se lavó 3 x con éter de dietilo para dar el compuesto del título.

40 El precipitado se secó *al vacío* durante 16 horas.

LCMS: m/z = 999 (M+1)⁺

Ejemplo 17

45 Preparación del análogo del GLP-1 péptido N-épsilon37-[2-(2-[2-[2-(2-[2-((S)-4-carboxi-4-({trans-4-[[19-carboxi-nonadecanoilamino]metil]ciclohexanocarbonil]amino)butirilamino]etoxi]etoxi)acetilamino]etoxi]etoxi]acetil[[Desamino His7,Glu22,Arg26,Arg34,Lys37]GLP-1-(7-37)



60 [Glu22, Arg26, Arg34, Lys37]GLP-1-(8-37) preparado por medio de una técnica recombinante (300 mg, 0,09 mmol) disuelta en 6 ml de agua se agregó DIPEA (373 µl, 2,19 mmol) y se agitó durante 10 min en donde una solución de ácido 19-[[trans-4-((S)-1-carboxi-3-{2-[2-({2-[2-(2,5-dioxo-pirrolidin-1-iloxi carbonil)metoxi]-etoxi]-etilcarbamoil}-metoxi)-etoxi]-etilcarbamoil}-propilcarbamoil)-ciclohexilmetil]-carbamoil]-nonadecanoico (175,7 mg, 0,18 mmol) en NMP (1,7 ml) se agregó en forma de gotas. La mezcla se agitó durante 90 min antes de agregar Na₂HPO₄·7H₂O (470,6 mg, 1,76 mmol) y el pH se ajustó a 8,4 mediante el uso de 1 N de HCl.

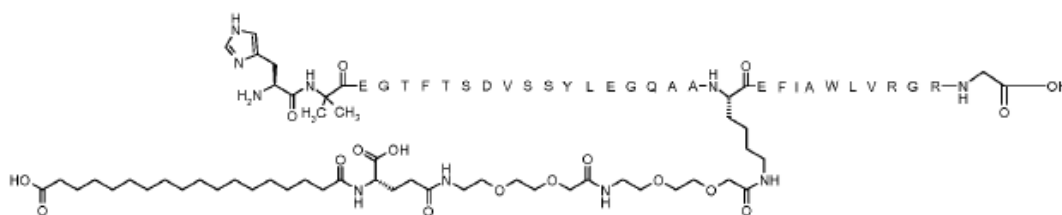
65

En otro matraz, ácido 3-(1H-Imidazol-4-il)-propiónico (93,2 mg, 0,53 mmol), tetrafluoroborato O-(N-succinimidilo)-N,N,N',N'-tetrametiluronio de (159,0 mg, 0,53 mmol) y DIPEA (180 µl, 1,06 mmol) se mezcló en NMP (0,75 ml) y la solución se agitó durante 1 hora antes de agregar en forma de gotas la mezcla de reacción que contiene el péptido. La mezcla se agitó durante 14 horas, en donde se agregó piperidina (350 µl, 3,51 mmol). La mezcla se agitó durante otras dos horas, antes de diluirla con agua hasta un volumen total de 20 ml, y purificarse mediante HPLC preparativa, usando un gradiente de 33 a 53% MeCN en agua.

LCMS: m/z = 1475,8 (M+3H)³⁺

Ejemplo 18

Preparación alternativa del análogo del GLP-1 N-épsilon26-[2-(2-[2-(2-[2-(2-[(S)-4-Carboxi-4-(17-carboxiheptadecanoilamino)butirilamino]etoxi)etoxi)acetilamino]etoxi)etoxi)acetil][Aib8,Arg34]GLP-1-(7-37)



[Arg34]GLP-1-(11-37) preparado mediante técnicas recombinantes (100 mg, 0,033 mmol) disuelto en 2 ml de agua se agregó a DIPEA (113 µl, 0,66 mmol) y se agitó durante 10 min en donde se agregó el ácido 17-((S)-1-carboxi-3-{2-[2-({2-[2-(2,5-dioxo-pirrolidin-1-iloxicarbonil-metoxi)-etoxi]-etilcarbamoil)-metoxi]-etoxi]-etilcarbamoil}-propil-carbamoil)-heptadecanoico (55 mg, 0,066 mmol) (Preparado de acuerdo con los ejemplos 2, 3 y 13) en pequeñas porciones durante 15 min. La mezcla se agitó durante 60 min antes de agregar Na₂HPO₄·7H₂O (353,9 mg, 1,32 mmol) y el pH se ajustó a 8,3 mediante el uso de 1 N de HCl. Fmoc-His-Aib-Glu-Gly-OSu (246,2 mg, 0,33 mmol) se agregó a la mezcla de reacción que contiene el péptido y la mezcla se agitó durante 14 horas después de lo cual se agregó piperidina (400 µl, 4,05 mmol). La mezcla se agitó durante otros 30 min, antes de diluirla con agua-MeCN (9:1) hasta un volumen total de 40 ml, y purificarse mediante HPLC preparativa, mediante el uso de un gradiente de 30 a 50 % MeCN en agua.

LCMS: m/z = 1372,5 (M+3H)³⁺

Ejemplo 19

Preparación alternativa del análogo del GLP-1 péptido N-épsilon26-[2-(2-[2-(2-[2-[(S)-4-Carboxi-4-(17-carboxiheptadecanoilamino)butirilamino]etoxi)etoxi)acetilamino]etoxi)etoxi)acetil][Aib8,Arg34]GLP-1-(7-37)

Acilación: [Arg34]GLP-1-(11-37) recombinante (99 mg; 33 µmol) se suspendió en H₂O (4 ml) en la cámara de reacción de 70 ml de un titulador plus Metrohm 848 Titrino. El pH se ajustó a pH=11,3 controlado por el programa de autotitulación SET (por adición de 3,8 ml, 0,1 M de NaOH (ac)). A pH = 11,3 el péptido se disolvió completamente. El ácido 17-((S)-1-carboxi-3-{2-[2-({2-[2-(2,5-dioxo-pirrolidin-1-iloxi carbonilmetoxi)etoxi]-etilcarbamoil)-metoxi]-etoxi]etilcarbamoil}-propilcarbamoil)-heptadecanoico de cadena lateral activada (55 mg, 66 µmol, 2 eq) se disolvió en NMP (250 µl) y se transfirió a una jeringa Hamilton de 250 µl. Mediante la acción de una bomba automatizada con jeringa se agregaron 250 µl (2,0 eq) de esta solución durante 10 minutos. Después de 60 minutos adicionales se extrajo una muestra (10 µl), se diluyó en 90 µl de MeCN-H₂O (1:1). UPLC mostró un pico principal de producto acilado N26 (aprox. 95 %) + traza de un péptido sin reaccionar + traza de cadena lateral hidrolizada.

Ligación: Éster de NHS tetrapéptido Fmoc-His-Aib-Glu-Gly-OSu (123 mg, 165 µmol, 5 eq) se disolvió en NMP (250 µl). El pH de la mezcla de reacción se ajustó a 6,9 con ácido acético diluido. Luego el pH se ajustó a 7,0 mediante la función pH-stat en el titulador. Se agregaron 250 µl de éster NHS durante 20 minutos. El pH se estabilizó en pH=7,0 ± 0,1.

La muestra se extrajo después de 30 min adicionales y UPLC mostró una relación 1:1 del intermedio acilado y el producto esperado (con protección de Fmoc en la posición N-terminal).

Se agregó otro 5 eq del éster de NHS tetrapéptido como anteriormente. La solución se agitó durante 16 h. La relación fue ahora 2:8 del intermedio acilado y el producto.

Se agregó otro 2,5 eq (total de 12,5 eq) del éster de NHS tetrapéptido como anteriormente. La solución se agitó durante 16 h. La relación fue ahora aprox. 1:9 del intermedio acilado y el producto.

ES 2 725 492 T3

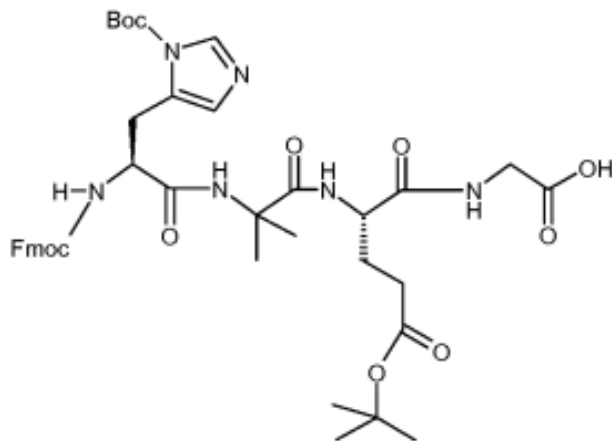
Se agregó piperidina (1,6 ml) a la mezcla de reacción y la agitación continuó durante 1h.

UPLC mostró desprotección total. El producto bruto se purificó en HPLC de fase inversa para proporcionar 62 mg (45 %) del producto deseado (98 % puro).

LCMS: m/z = 1372,5 (M+3H)³⁺

Ejemplo 20

Preparación de Fmoc-His(Boc)-Aib-Glu(O-tBu)-Gly-OH



	Nombre	Dens	PM:	Mol	n	W	V	
		[g/ml]	[g/mol]	Relación	[mmol]	[g]	ml	
	Producto: C41H52N6O11		804,905	1	98	78,9		
1ro AA	resina de 2-clorotritilcloruro 1 % DVB			1	98	140		
	Fmoc-Gly-OH		297,3	2	196	58,3		
	DIPEA	0,76	129,3	2	196		66,7	pH 9
	Sust. Resina Fmoc-Gly: 0,7 mmol/g							
2do AA	Fmoc-Glu(OtBu)-OH		425,5	2	196	86,9		
	TCTU		355,5	2	196	69,7		
	DIPEA	0,76	129,3	2	196		66,7	
	Fmoc-Glu(OtBu)-OH		425,5	0,5	49	21,7		Reacoplamiento
	TCTU		355,5	0,5	49	17,4		
	DIPEA	0,76	129,3	0,5	49		17,6	
3er AA	Fmoc-Aib-OH		325,4	2	196	63,8		
	TCTU		355,5	2	196	69,7		
	DIPEA	0,76	129,3	2	196		66,7	
4to AA	Fmoc-His(boc)-OH		477,5	2	196	93,6		
	TCTU		355,5	2	196	69,7		
	DIPEA	0,76	129,3	2	196		66,7	

Carga de glicina

5 La resina se preinchó en DCM (1 litro, 20 min) antes de agregar una solución de Fmoc-Gly-OH y DIPEA en 500 ml de DCM, y se agitó durante 60 min. La resina se lavó con DMF (2 x 500 ml) y 2 x 500 ml de DCM-MeOH-DIPEA (80:15:5, 10 min) y después DMF (3 x 500 ml), DCM (500 ml) y se dejó durante la noche a temperatura ambiente.

Acoplamiento 1

10 Se usó piperidina al 20 % en DMF (500 ml) durante 50 min para la desprotección de Fmoc, luego la resina se lavó con DMF (8 x 500 ml).

15 Fmoc-Glu(OtBu)-OH y TCTU se disolvieron en 700 ml de DMF agregado al reactor y se agregó DIPEA. La mezcla se agitó durante 6 h. De acuerdo con la HPLC, los grupos amina libres de prueba estaban todavía presentes; el acoplamiento se repitió mediante el uso de $\frac{1}{4}$ de reactivos. Cuando se terminó la reacción de acoplamiento (sin grupos amino presentes), la resina se lavó con DMF (4 x 500 ml)

Acoplamiento 2

20 Fmoc-Aib-OH y TCTU se disolvieron en 700 ml de solución de DMF, se colocó en el reactor y se agregó DIPEA. La mezcla se agitó durante 50 min. Cuando terminó la reacción de acoplamiento (sin grupos amino presentes), la resina se lavó con DMF (4 x 500 ml).

Acoplamiento 3

25 Se usó piperidina al 20 % en DMF (500 ml) durante 30 min para la desprotección de Fmoc y luego la resina se lavó con DMF (8 x 500 ml).

30 Fmoc-His(Boc)-OH y TCTU se disolvieron en 700 ml de DMF, se agregó la solución en el reactor y se agregó DIPEA. La mezcla se agitó durante 50 min. Cuando terminó la reacción de acoplamiento (sin grupos amino presentes), la resina se lavó con DMF (3 x 500 ml) y DCM (2 x 500 ml).

1^{ra} Escisión

35 La 1^{ra} escisión se realizó mediante el uso de 1/7 de resina. Se agregó trifluoroetanol-DCM (1:4, 500 ml) y la resina se agitó durante 70 min. Se recogió el filtrado y el solvente se eliminó al vacío para proporcionar aceite. El aceite se disolvió en acetato de etilo y se dejó en refrigerador para formar cristales. Los cristales se filtraron, se lavaron mediante el uso de éter de dietilo y se secaron al aire, para proporcionar 6,1 g del producto (rendimiento de 45 %), pureza mediante HPLC de 99,0 %.

40 2^{da} Escisión

45 La 2^{da} escisión se realizó mediante el uso de 3/7 de resina. Se agregó trifluoroetanol-DCM (1:4, 1000 ml) y la resina se agitó durante 90 min. Se recogió el filtrado y el solvente se eliminó al vacío para proporcionar aceite. El aceite se disolvió en acetato de etilo y se dejó en refrigerador durante toda la noche para formar cristales. Los cristales se filtraron, se lavaron con éter de dietilo y se secaron al aire para proporcionar 33,1 g de producto (rendimiento de 80 %), pureza mediante HPLC de 98,7 %.

Rendimiento: 6,1 g (45 %)

50 ESI+MS m/z: 805,1 (M+H)⁺, 1610,7 (2M+H)⁺.

HPLC R_t (Luna 4,6 x 250, 5 ul, 100A; acetonitrilo / tampón* 30:70 a 60:40, 30 min, 1 ml/min., 220 nm): 23,31 min.; pureza de 99,0%

55 y 33,1 g (80 %)

ESI+MS m/z: 805,1 (M+H)⁺, 1610,8 (2M+H)⁺.

60 HPLC R_t (Luna 4,6 x 250, 5 ul, 100A, acetonitrilo / tampón* 30:70 a 60:40, 30 min, 1 ml/min., 220 nm): 23,16 min.; pureza de 98,7%

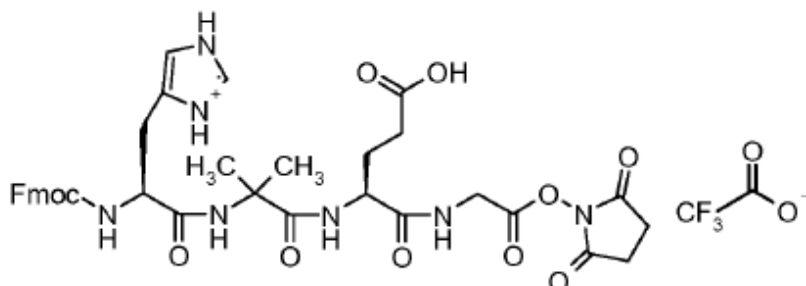
*tampón: 2,71 g de KH₂PO₄, 7,13 g de NaH₂PO₄·2H₂O disuelto en H₂O (2L);

Ejemplo 21

65

5

10



Preparación de Fmoc-His-Aib-Glu-Gly-OSuc. Sal de TFA

15 Fmoc-His(Boc)-Aib-Glu(O-tBu)-Gly-OH (1,25 g, 1,56 mmol) se disolvió en THF (10 ml) y se agregó DIPEA (0,64 g, 3,73 mmol) y TSTU (0,56 g, 1,87 mmol, 1,2 eq). La mezcla se agitó a temperatura ambiente durante la noche.

LCMS: conversión de la aplicación de 95 %. La mezcla se filtró y el solvente se eliminó *al vacío*. EtOAc (100 ml) se agregó y la fase orgánica se lavó dos veces con 0,1 N de HCl frío y se secó sobre Na₂SO₄, se filtró y se

20

evaporó para dar Fmoc-His(Boc)-Aib-Glu(OtBu)-Gly-OSu bruto (1,1 g, 79 %).

El producto bruto se disolvió en DCM (2 ml) y se agregó TFA (1 ml). Se agitó durante 1h. LCMS: 100 % de desprotección. Después de la evaporación del solvente orgánico se agregó dietiléter frío al residuo y se recogió el precipitado resultante mediante filtración y se secó en un horno de vacío para dar Fmoc-His-Aib-Glu-Gly-OSu (430 mg)

25

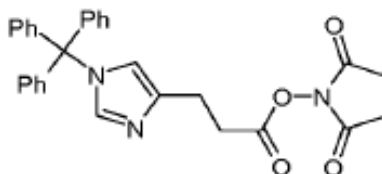
LCMS: m/z = 746,3 (M+H)⁺

30 Ejemplo 22

Preparación de Ácido 3-(1-tritil-1H-imidazol-4-il)-propiónico 2,5-dioxo-pirrolidin-1-il éster

35

40



45 Ácido 3-(*N*-tritacilimidazol-4-il)propiónico (Jung y otros, Bioorg. Med. Chem. Lett., 6, 2317-2336, 1998) (50 g, 0,13 mol) se disolvió en DCM (500 ml), *N*-hidroxisuccinimida (24 g, 0,209 mol) y se agregó *N,N'*-diciclohexilcarbodiimida (35 g 0,170 mol), mientras que se mantiene la temperatura interna por debajo de 30 °C con un baño de hielo. La mezcla resultante se agitó a temperatura ambiente durante la noche. La mezcla se filtró y la solución resultante se concentró *al vacío*, se volvió a disolver en THF (350 ml) tras el calentamiento y se mezcló con 2-propanol (350 ml) hasta una temperatura resultante de 32 °C, luego se enfrió hasta 5 °C. El producto cristalino se filtró y se lavó con 2-propanol (150 ml) y se secó al vacío para proporcionar el compuesto del título (47,3 g, rendimiento de 75 %).

50

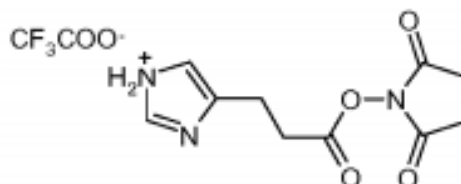
¹H NMR (*D*₆-DMSO, 400Mhz): δ 7,15-6,99 (m, 9H), 6,82 (s, 1H), 6,81-6,62 (m, 6H), 6,38 (s, 1H), 2,63 (t, 3H), 2,4 (m, 6H).

55 Ejemplo 23

Preparación de trifluoroacetato de 4-[2-(2,5-dioxo-pirrolidin-1-iloxicarbonil)-etil]-1*H*-imidazol-1-*io*

60

65





10 Arg34-GLP-1[9-37] · 4TFA (103 mg; 0,028 mmol) se disolvió en 2 ml de agua y se transfirió a una cámara de
 reacción de titulación con 2 x 1 ml adicionales de agua. La cámara de reacción se equipó con un imán de agitación,
 electrodo de pH y un tubo de titulante. El pH de la solución fue 1,9. El titulador se programó a un método 'STAT' a
 pH fijo = 11,3. El titulante, el hidróxido de sodio (3,27 ml, 0,1 M ac), se agregó a la titulación para alcanzar pH = 11,3.
 15 (Boc)₂O (186 mg, 0,85 mmol, 3 eq) se disolvió en NMP (2,500 ml) y se llenó una jeringa de Hamilton de 250 µl con
 esta solución (es decir, 250 µl; 18,6 mg de (Boc)₂O, 3 eq).

Se agregó agua (1,0 ml) a la cámara de reacción para obtener un volumen total de 8,3 ml y la solución de (Boc)₂O
 NMP se agregó después durante 20 min mediante la ayuda de una bomba de jeringa a temperatura ambiente. El
 titulador mantuvo el pH = 11,3 mediante la adición automática del titulante. Después de agitar durante 10 minutos
 20 adicionales, LCMS mostró 88 % del producto Lys26(NHBoc) junto con un producto diboc (9 %) y material de partida
 (3 %). El pH de la solución se ajustó a 6,5 con un volumen pequeño de HOAc diluido (ac) y el titulador se ajustó en
 pH = 7,5 para la reacción adicional.

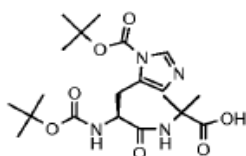
Una solución de Boc-His(Boc)-Aib-OSu (305 mg, 0,568 mmol, 20 eq) se disolvió en NMP (500 µl) lo que proporcionó
 25 un volumen total de 715 µl. De esta solución 2 x 250 µl (2 x 7 eq) se agregó a la mezcla de reacción durante 2 x 30
 minutos, manteniendo el pH = 7,5. Se observó precipitación inmediata de Boc-His(Boc)-Aib-OSu, pero al agitar todo
 se disolvió. Después del final de la adición, la mezcla se agitó durante otras 2 h a pH = 7,5 y a temperatura ambiente
 para permitir la disolución y la reacción del reactivo.

30 La mezcla se liofilizó y se trató con TFA que contenía 3 % de TIPS y 3 % de agua durante 2 h. Los solventes se
 evaporaron y el residuo se purificó mediante HPLC para dar el producto deseado.

LCMS m/z: 849,88 (M+4H)⁴⁺, 1132,86 (M+3H)³⁺. Calculado 3396,698 (M+H)⁺.
 Ejemplo 26

35 Preparación de Boc-His(Boc)-Aib-OH

40



45

50 Carga de Aib: La resina de 2-clortritilcloruro al 1 % DVB (Carga 1,1 mmol/g) (10,0 g, 11 mmol) se prehinchó en DCM
 antes de agregar una solución de Fmoc-Aib-OH (7,16 g, 22 mmol) y DIPEA (9,4 ml, 55 mmol) en DCM (100 ml). La
 mezcla se agitó durante 60 min. La resina se lavó con NMP (2 x 100 ml) y DCM-MeOH-DIPEA (80:15:5) (2 x 100 ml,
 10 minutos cada uno), y NMP (3 x 100 ml). La resina lavada se trató con piperidina al 20 % en NMP (3 x 100 ml, 10
 min cada una). La resina se lavó con NMP (6 x 100 ml).

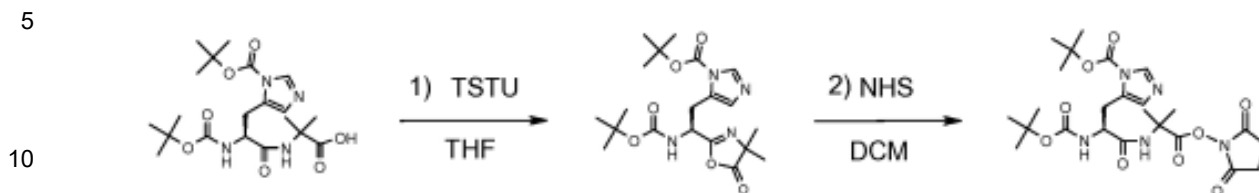
55 Acoplamiento: Se disolvió Boc-His(Boc)-OH (15,64 g, 44 mmol) en NMP (100 ml) y se agregó DCM (20 ml) y HOBt
 (5,95 g, 44 mmol) seguido por la adición lenta de DIC (6,81 ml, 44 mmol). La mezcla se agitó durante 15 min antes
 de agregar DIPEA (9,40 ml, 55 mmol). La mezcla activada se agregó a la resina y la resina se agitó durante 16 h
 antes de lavarse con NMP (4 x 100 ml) y DCM (10 x 100 ml).

60 Escisión: Se agregó trifluoroetanol-DCM (1:4, 100 ml) y la resina se agitó durante 1 hora. Se recogió el filtrado y se
 agregó y agitó una nueva parte de trifluoroetanol-DCM (1:4, 100 ml) durante 15 min y se recogió el filtrado. Los
 filtrados combinados se evaporaron *al vacío* y se agregó petroléter (50 ml). El precipitado resultante se lavó con
 petroléter, se recogió y se secó en horno de vacío durante 16 h a 40 °C para dar Boc-His(Boc)-Aib-OH (1,00 g,
 65 rendimiento de 21 %, pureza de 95 %).

LCMS m/z: 441,18 (M+H)⁺. Calculado 440,227 (M+H)⁺.

Ejemplo 27

Preparación de Boc-His(Boc)-Aib-OSu

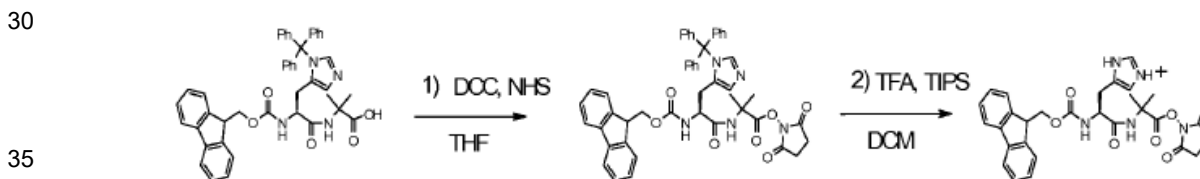


15 Se disolvió Boc-His(Boc)-Aib-OH (0,70 g, 1,59 mmol) en THF (5,0 ml) y se agregó TSTU (1,44 g, 4,77 mmol) y DIPEA (1,36 ml, 7,95 mmol). La mezcla de reacción se agitó a rt. durante 2 h. LCMS mostró una conversión casi completa en el producto de oxazolona. La mezcla de reacción se filtró y el solvente se evaporó. El residuo se volvió a disolver en DCM (10 ml) y se agregó N-hidroxisuccinimida (292 mg, 2,56 mmol, 1,6 eq). La mezcla se agitó a temperatura ambiente durante la noche. UPLC muestra un pico principal del éster de Su deseado (85 %), traza de material inicial sin reaccionar y 5 % de la oxazolona anillo cerrado. El solvente se evaporó a temperatura ambiente a presión reducida y se agregó H₂O (25 ml) para disolver N-hidroxisuccinimida sin reaccionar. El H₂O se decantó y el residuo se disolvió en EtOAc (100 ml) se secó sobre MgSO₄, se filtró y se evaporó para dar una espuma blanca que contiene 70 % del producto deseado.

20 LCMS m/z: 538,18 (M+H)⁺. Calculado 538,251 (M+H)⁺

Ejemplo 28

Preparación de Fmoc-His-Aib-OSu



40 Fmoc-His(Trt)-Aib-OH (5 g, 7,1 mmol) se disolvió en THF (50 ml). Se agregó DCC (1,61 g, 7,8 mmol) y NHS (0,94 g, 8,2 mmol) y la solución resultante se agitó durante 16 h a temperatura ambiente. DCU se filtró y la solución se evaporó *al vacío*. El aceite resultante se disolvió en DCM (34 ml) y se agregó ácido trifluoroacético (34 ml) y TIPS (6 ml). La solución se agitó durante 2 h a temperatura ambiente, se evaporó *al vacío* a 25 ml, se agregó en forma de gotas a TBME (200 ml), se agitó durante 60 minutos, y se recogió el precipitado blanco resultante mediante filtración, y se lavó con TBME adicional. El precipitado recogido se secó al vacío durante 16 h.

45 LCMS: m/z = 560 (M+H)⁺

Ejemplo 29

Preparación del derivado del GLP-1 péptido [Aib8,Arg34]GLP-1-(7-37)



El péptido [Arg34] GLP-1(9-37) • 4TFA (103 mg; 0,028 mmol) se disolvió en 0,1 M de TEA (ac., 7,5 ml) se agregó TEA adicional (140 µl) para un pH resultante de 11,28. FmocOSu (90 mg) se disolvió en NMP (1600 µl) y se agregó 334 µl (0,055 mmol) de esta solución en forma de gotas a la solución de péptido acuosa. El pH se ajustó a 7 mediante la adición de 1M de H₂SO₄ (ac). La sal de trifluoroacetato FmocHisAibOSu (90 mg, 0,134 mmol) se disolvió en NMP (800 µl) y se agregó en forma de gotas a la solución mientras se mantuvo el pH entre 7,0 y 7,5 mediante la adición continua de 1M de NaOH(ac). Se agregó piperidina (2 ml) y la solución resultante se agitó durante 60 min a

ES 2 725 492 T3

temperatura ambiente. Se agregó TBME (5 ml) y la mezcla se agitó durante 30 min. La fase acuosa se sembró y se liofilizó.

LCMS: m/z = 1134 (M+3H)³⁺

5

Listado de secuencias

<110> Novo Nordisk A/S

10

<120> Producción semirrecombinante de análogos del GLP-1

<130> 7724.204-WO

15

<160> 11

<170> PatentIn versión 3.3

<210> 1

20

<211> 31

<212> PRT

<213> Homo sapiens

25

<400> 1

His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15

30

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly
20 25 30

35

<210> 2

<211> 40

40

<212> PRT

<213> Heloderma suspectum

<400> 2

45

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15

50

Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser
20 25 30

55

Ser Gly Ala Pro Pro Pro Ser Gly
35 40

60

<210> 3

<211> 35

<212> PRT

65

<213> Artificial

<220>
<223> artificial

5 <220>
<221> CARACTERÍSTICA_MISC
<222> (6)..(6)

10 <223> Xaa en la posición 6 es Val o Leu
<220>

15 <221> CARACTERÍSTICA_MISC
<222> (8)..(8)

20 <223> Xaa en la posición 8 es Ser, Lys o Arg
<220>

<221> CARACTERÍSTICA_MISC
<222> (9)..(9)

25 <223> Xaa en la posición 9 es Tyr o Gln
<220>

30 <221> CARACTERÍSTICA_MISC
<222> (10)..(10)

<223> Xaa en la posición 10 es Leu o Met

35 <220>
<221> CARACTERÍSTICA_MISC

40 <222> (12)..(12)

<223> Xaa en la posición 12 es Gly o Glu
<220>

45 <221> CARACTERÍSTICA_MISC
<222> (13)..(13)

50 <223> Xaa en la posición 13 es Gln, Glu, Lys o Arg
<220>

<221> CARACTERÍSTICA_MISC

55 <222> (15)..(15)

<223> Xaa en la posición 15 es Ala o Val

60 <220>
<221> CARACTERÍSTICA_MISC

<222> (16)..(16)

65 <223> Xaa en la posición 16 es Lys, Glu o Arg

<220>
<221> CARACTERÍSTICA_MISC
5 <222> (17)..(17)
<223> Xaa en la posición 17 es Glu o Leu
<220>
10 <221> CARACTERÍSTICA_MISC
<222> (20)..(20)
15 <223> Xaa en la posición 20 es Ala, Glu o Arg
<220>
<221> CARACTERÍSTICA_MISC
20 <222> (23)..(23)
<223> Xaa en la posición 23 es Val, Lys o Arg
<220>
25 <221> CARACTERÍSTICA_MISC
<222> (24)..(24)
30 <223> Xaa en la posición 24 es Lys, Glu, Asn, His o Arg
<220>
35 <221> CARACTERÍSTICA_MISC
<222> (25)..(25)
<223> Xaa en la posición 25 es Gly
40 <220>
<221> CARACTERÍSTICA_MISC
45 <222> (26)..(26)
<223> Xaa en la posición 26 es Arg, Gly o Lys
<220>
50 <221> CARACTERÍSTICA_MISC
<222> (27)..(27)
55 <223> Xaa en la posición 27 es Gly, Ala, Glu, Pro, Lys o está ausente
<220>
<221> CARACTERÍSTICA_MISC
60 <222> (28)..(28)
<223> Xaa en la posición 28 es Lys, Ser, o está ausente
65 <220>

<221> CARACTERÍSTICA_MISC
 <222> (29)..(29)
 5 <223> Xaa en la posición 29 es Ser, Lys, o está ausente
 <220>
 <221> CARACTERÍSTICA_MISC
 10 <222> (30)..(30)
 <223> Xaa en la posición 30 es Gly o está ausente
 15 <220>
 <221> CARACTERÍSTICA_MISC
 <222> (31)..(31)
 20 <223> Xaa en la posición 31 es Ala o está ausente
 <220>
 25 <221> CARACTERÍSTICA_MISC
 <222> (32)..(32)
 <223> Xaa en la posición 32 es Pro o está ausente
 30 <220>
 <221> CARACTERÍSTICA_MISC
 35 <222> (33)..(33)
 <223> Xaa en la posición 33 es Pro o está ausente
 <220>
 40 <221> CARACTERÍSTICA_MISC
 <222> (34)..(34)
 45 <223> Xaa en la posición 34 es Pro o está ausente
 <220>
 <221> CARACTERÍSTICA_MISC
 50 <222> (35)..(35)
 <223> Xaa en la posición 35 es Ser o está ausente
 55 <400> 3

Thr Phe Thr Ser Asp Xaa Ser Xaa Xaa Xaa Glu Xaa Xaa Ala Xaa Xaa
 1 5 10 15

Xaa Phe Ile Xaa Trp Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 20 25 30

65

Xaa Xaa Xaa
35

- 5 <210> 4
<211> 37
<212> PRT
10 <213> Artificial
<220>
15 <223> artificial
<220>
<221> CARACTERÍSTICA_MISC
20 <222> (1)..(1)
<223> Xaa en la posición 1 es Glu o Asp
<220>
25 <221> CARACTERÍSTICA_MISC
<222> (2)..(2)
30 <223> Xaa en la posición 2 es Gly o Ala
<220>
<221> CARACTERÍSTICA_MISC
35 <222> (8)..(8)
<223> Xaa en la posición 8 es Val o Leu
40 <220>
<221> CARACTERÍSTICA_MISC
<222> (10)..(10)
45 <223> Xaa en la posición 10 es Ser, Lys o Arg
<220>
50 <221> CARACTERÍSTICA_MISC
<222> (11)..(11)
<223> Xaa en la posición 11 es Tyr o Gln
55 <220>
<221> CARACTERÍSTICA_MISC
60 <222> (12)..(12)
<223> Xaa en la posición 12 es Leu o Met
<220>
65 <221> CARACTERÍSTICA_MISC

<222> (14)..(14)
<223> Los Xaa en la posición 14 son Gly o Glu
5 <220>
<221> CARACTERÍSTICA_MISC
<222> (15)..(15)
10 <223> Xaa en la posición 15 es Gln, Glu, Lys o Arg
<220>
15 <221> CARACTERÍSTICA_MISC
<222> (17)..(17)
<223> Xaa en la posición 17 es Ala o Val
20 <220>
<221> CARACTERÍSTICA_MISC
25 <222> (18)..(18)
<223> Xaa en la posición 18 es Lys, Glu o Arg
<220>
30 <221> CARACTERÍSTICA_MISC
<222> (19)..(19)
35 <223> Xaa en la posición 19 es Glu o Leu
<220>
<221> CARACTERÍSTICA_MISC
40 <222> (22)..(22)
<223> Xaa en la posición 22 es Ala, Glu o Arg
45 <220>
<221> CARACTERÍSTICA_MISC
<222> (25)..(25)
50 <223> Xaa en la posición 25 es Val, Lys o Arg
<220>
55 <221> CARACTERÍSTICA_MISC
<222> (26)..(26)
<223> Xaa en la posición 26 es Lys, Glu, Asn, His o Arg
60 <220>
<221> CARACTERÍSTICA_MISC
65 <222> (27)..(27)

<223> Xaa en la posición 27 es Gly
 <220>

5 <221> CARACTERÍSTICA_MISC
 <222> (28)..(28)
 <223> Xaa en la posición 28 es Arg, Gly o Lys
 10 <220>
 <221> CARACTERÍSTICA_MISC
 15 <222> (29)..(29)
 <223> Xaa en la posición 29 es Gly, Ala, Glu, Pro, Lys o está ausente
 <220>

20 <221> CARACTERÍSTICA_MISC
 <222> (30)..(30)
 25 <223> Xaa en la posición 30 es Lys, Ser, o está ausente
 <220>
 <221> CARACTERÍSTICA_MISC
 30 <222> (31) .. (31)
 <223> Xaa en la posición 31 es Ser, Lys, o está ausente
 35 <220>
 <221> CARACTERÍSTICA_MISC
 <222> (32)..(32)
 40 <223> Xaa en la posición 32 es Gly o está ausente
 <220>

45 <221> CARACTERÍSTICA_MISC
 <222> (33)..(33)
 <223> Xaa en la posición 33 es Ala o está ausente
 50 <220>
 <221> CARACTERÍSTICA_MISC
 55 <222> (34)..(34)
 <223> Xaa en la posición 34 es Pro o está ausente
 <220>

60 <221> CARACTERÍSTICA_MISC
 <222> (35)..(35)
 65 <223> Xaa en la posición 35 es Pro o está ausente

<220>
 <221> CARACTERÍSTICA_MISC
 5 <222> (36)..(36)
 <223> Xaa en la posición 36 es Pro o está ausente
 <220>
 10 <221> CARACTERÍSTICA_MISC
 <222> (37)..(37)
 15 <223> Xaa en la posición 37 es Ser o está ausente
 <400> 4

20 Xaa Xaa Thr Phe Thr Ser Asp Xaa Ser Xaa Xaa Xaa Glu Xaa Xaa Ala
 1 5 10 15

25 Xaa Xaa Xaa Phe Ile Xaa Trp Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 20 25 30

30 Xaa Xaa Xaa Xaa Xaa
 35

<210> 5
 35 <211> 38
 <212> PRT
 <213> Artificial
 40 <220>
 <223> artificial
 45 <220>
 <221> CARACTERÍSTICA_MISC
 <222> (1)..(1)
 50 <223> Xaa en la posición 1 es Ala, Gly, Val, Leu, Ile o Lys
 <220>
 55 <221> CARACTERÍSTICA_MISC
 <222> (2)..(2)
 <223> Xaa en la posición 2 es Glu o, Asp
 60 <220>
 <221> CARACTERÍSTICA_MISC
 <222> (3)..(3)
 65 <223> Xaa en la posición 3 es Gly o Ala

<220>
 <221> CARACTERÍSTICA_MISC
 5 <222> (9)..(9)
 <223> Xaa puede ser cualquier aminoácido de origen natural
 <220>
 10 <221> CARACTERÍSTICA_MISC
 <222> (11)..(11)
 15 <223> Xaa en la posición 11 es Ser, Lys o Arg
 <220>
 <221> CARACTERÍSTICA_MISC
 20 <222> (12)..(12)
 <223> Xaa en la posición 12 es Tyr o Gln
 <220>
 25 <221> CARACTERÍSTICA_MISC
 <222> (13)..(13)
 30 <223> Xaa en la posición 13 es Leu o Met
 <220>
 <221> CARACTERÍSTICA_MISC
 35 <222> (15)..(15)
 <223> Xaa en la posición 15 es Gly o Glu
 <220>
 40 <221> CARACTERÍSTICA_MISC
 <222> (16)..(16)
 45 <223> Xaa en la posición 16 es Gln, Glu, Lys o Arg
 <220>
 50 <221> CARACTERÍSTICA_MISC
 <222> (18)..(18)
 <223> Xaa en la posición 18 es Ala o Val
 55 <220>
 <221> CARACTERÍSTICA_MISC
 60 <222> (19)..(19)
 <223> Xaa en la posición 19 es Lys, Glu o Arg
 <220>
 65 <221> CARACTERÍSTICA_MISC

<222> (20)..(20)
<223> Xaa en la posición 20 es Glu o Leu
5 <220>
<221> CARACTERÍSTICA_MISC
<222> (23)..(23)
10 <223> Xaa en la posición 23 es Ala, Glu o Arg
<220>
15 <221> CARACTERÍSTICA_MISC
<222> (26)..(26)
<223> Xaa en la posición 26 es Val, Lys o Arg
20 <220>
<221> CARACTERÍSTICA_MISC
25 <222> (27)..(27)
<223> Xaa en la posición 27 es Lys, Glu, Asn, His o Arg
<220>
30 <221> CARACTERÍSTICA_MISC
<222> (28)..(28)
35 <223> Xaa en la posición 28 es Gly
<220>
<221> CARACTERÍSTICA_MISC
40 <222> (29)..(29)
<223> Xaa en la posición 29 es Arg, Gly o Lys
45 <220>
<221> CARACTERÍSTICA_MISC
<222> (30)..(30)
50 <223> Xaa en la posición 30 es Gly, Ala, Glu, Pro, Lys o está ausente
<220>
55 <221> CARACTERÍSTICA_MISC
<222> (31)..(31)
<223> Xaa en la posición 31 es Lys, Ser, o está ausente
60 <220>
<221> CARACTERÍSTICA_MISC
65 <222> (32)..(32)

<223> Xaa en la posición 32 es Ser, Lys, o está ausente
 <220>
 5 <221> CARACTERÍSTICA_MISC
 <222> (33)..(33)
 <223> Xaa en la posición 33 es Gly o está ausente
 10 <220>
 <221> CARACTERÍSTICA_MISC
 15 <222> (34)..(34)
 <223> Xaa en la posición 34 es Ala o está ausente
 <220>
 20 <221> CARACTERÍSTICA_MISC
 <222> (35)..(35)
 25 <223> Xaa en la posición 35 es Pro o está ausente
 <220>
 <221> CARACTERÍSTICA_MISC
 30 <222> (36)..(36)
 <223> Xaa en la posición 36 es Pro o está ausente
 35 <220>
 <221> CARACTERÍSTICA_MISC
 <222> (37)..(37)
 40 <223> Xaa en la posición 37 es Pro o está ausente
 <220>
 45 <221> CARACTERÍSTICA_MISC
 <222> (38)..(38)
 <223> Xaa en la posición 38 es Ser o está ausente
 50 <400> 5

Xaa Xaa Xaa Thr Phe Thr Ser Asp Xaa Ser Xaa Xaa Xaa Glu Xaa Xaa
 55 1 5 10 15

Ala Xaa Xaa Xaa Phe Ile Xaa Trp Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 60 20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa
 65 35

ES 2 725 492 T3

<210> 6
 <211> 81
 5 <212> ADN
 <213> Artificial
 <220>
 10 <223> Artificial
 <400> 6
 15 acttttactt ctgatgttct ttcttatttg gaaggccaag ctgctaaaga atttattgct 60
 tggttgggta gaggtagagg t 81
 20
 <210> 7
 <211> 117
 25 <212> ADN
 <213> Artificial
 <220>
 30 <223> Artificial
 <400> 7
 35 gaagaagctg aaaaggctga aggtactttt acttctgatg tttcttctta tttggaagaa 60
 caagctgcta gagaatttat tgcttgggtg gttagaggta gaggtaaaga agctgaa 117
 40
 <210> 8
 <211> 114
 45 <212> ADN
 <213> Artificial
 <220>
 50 <223> artificial
 <400> 8
 55 gaagaagctg aaaaggctga aggtactttt acttctgatg tttcttctta tttggaagaa 60
 caagctgcta gagaatttat tgcttgggtg gttagaggta gaaaagaagc tgaa 114
 60
 <210> 9
 <211> 27
 65 <212> PRT

ES 2 725 492 T3

<213> Artificial
 <220>
 5 <223> artificial
 <400> 9

10 Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly Gln Ala Ala Lys
 1 5 10 15

15 Glu Phe Ile Ala Trp Leu Val Arg Gly Arg Gly
 20 25

<210> 10
 <211> 39
 20 <212> PRT
 <213> Artificial
 25 <220>
 <223> artificial
 <400> 10

30 Glu Glu Ala Glu Lys Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser
 1 5 10 15

35 Tyr Leu Glu Glu Gln Ala Ala Arg Glu Phe Ile Ala Trp Leu Val Arg
 20 25 30

40 Gly Arg Gly Lys Glu Ala Glu
 35

<210> 11
 45 <211> 38
 <212> PRT
 <213> Artificial
 50 <220>
 <223> artificial
 <400> 11

55 Glu Glu Ala Glu Lys Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser
 1 5 10 15

60 Tyr Leu Glu Glu Gln Ala Ala Arg Glu Phe Ile Ala Trp Leu Val Arg
 20 25 30

65 Gly Arg Lys Glu Ala Glu
 35

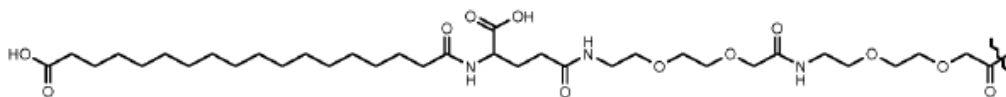
ES 2 725 492 T3

<210> 12
<211> 31
<212> PRT
5 <213> ARTIFICIAL
<220>
10 <223> Análogo de Exendina
<220>
<221> CARACTERÍSTICA_MISC
15 <222> (31)..(31)
<223> Pro o Tyr
20 <400> 12

His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu
1 5 10 15
25
Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Xaa
20 25 30
30 <210> 13
<211> 40
<212> PRT
35 <213> Heloderma sp.
<400> 13
40
His Ser Asp Gly Thr Phe Ile Thr Ser Asp Leu Ser Lys Gln Met Glu
1 5 10 15
45
Glu Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro
20 25 30
50 Ser Ser Gly Ala Pro Pro Pro Ser
35 40

REIVINDICACIONES

1. Un método para producir un derivado de un análogo del GLP-1 que comprende uno o más aminoácidos no proteinogénicos en la parte N-terminal, dicho método que comprende las etapas de:
 - (i) cultivar una célula huésped que comprende una secuencia de nucleótidos que codifica una molécula precursora de dicho análogo del GLP-1 en condiciones adecuadas para la expresión de dicha molécula precursora;
 - (ii) separar la molécula precursora expresada del caldo de cultivo;
 - (iii) acilar el grupo épsilon amino de un residuo de lisina en la molécula precursora expresada con un agente de acilación, que se activa opcionalmente, para obtener un derivado de la molécula precursora, y opcionalmente aislar dicho derivado de la molécula precursora;
 - (iv) acoplar una extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos al derivado de la molécula precursora expresado; y
 - (v) aislar el derivado resultante de un análogo del GLP-1 por medios adecuados;
 en donde
 - (a) dicha etapa de acilación (iii) tiene lugar en una mezcla de solventes acuosos con un pH entre 9 y 13;
 - (b) dicha extensión N-terminal tiene una longitud de 2 o 4 aminoácidos, y
 - (c) dicho derivado resultante de un análogo del GLP-1 es el péptido N-épsilon26-[2-(2-[2-(2-{(S)-4-Carboxi-4-(17-carboxi-heptadecanoilamino)-butirilamino]etoxi)etoxi)acetilamino]-etoxi)etoxi)-acetil][Aib8,Arg34] GLP-1-(7-37)
2. El método de acuerdo con la reivindicación 1, en donde dicha molécula precursora es [Arg34]GLP-1 (11-37).
3. El método de acuerdo con la reivindicación 1, en donde dicha molécula precursora es [Arg34]GLP-1(9-37).
4. El método de acuerdo con cualquiera de las reivindicaciones anteriores, en donde la extensión de aminoácidos N-terminal se protege con uno o más grupos de protección antes de su uso en la etapa (iv) y se desprotege otra vez después del acoplamiento de la extensión N-terminal.
5. El método de acuerdo con cualquiera de las reivindicaciones anteriores, donde el acoplamiento de la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos tiene lugar en un disolvente polar orgánico seleccionado del grupo que consiste en 1-metil-pirrolidin-2-ona, acetonitrilo, tetrahidrofurano, dimetilsulfóxido *N,N*-dimetilformamida y *N*-metil-formamida.
6. El método de acuerdo con cualquiera de las reivindicaciones anteriores, donde el acoplamiento de la extensión de aminoácidos N-terminal que comprende uno o más aminoácidos no proteinogénicos tiene lugar en una mezcla de solventes acuosos.
7. El método de acuerdo con cualquier reivindicación anterior, que comprende además la etapa de eliminar (i) una extensión propéptido N-terminal o (ii) una extensión propéptido C-terminal de la molécula precursora resultante.
8. El método de acuerdo con cualquiera de las reivindicaciones anteriores, en donde dicha etapa de acilación (iii) tiene lugar en una mezcla de solventes acuosos con un pH entre 10 y 12, tal como entre 10,5 y 11,5.
9. El método de acuerdo con cualquiera de las reivindicaciones anteriores, en donde dicha célula huésped se selecciona de una célula huésped de mamífero, una célula huésped de ave, una célula huésped de insectos, una célula huésped vegetal, una célula huésped bacteriana, una célula huésped fúngica y una célula huésped de levadura.
10. El método de acuerdo con cualquier reivindicación anterior, en donde dicho agente de acilación es un análogo de ácido carboxílico de la fórmula general:
 A-C-D-OH
 en donde
 un ácido carboxílico de terminal libre se activa opcionalmente, y
 A-C-D- se selecciona del grupo que consiste en:



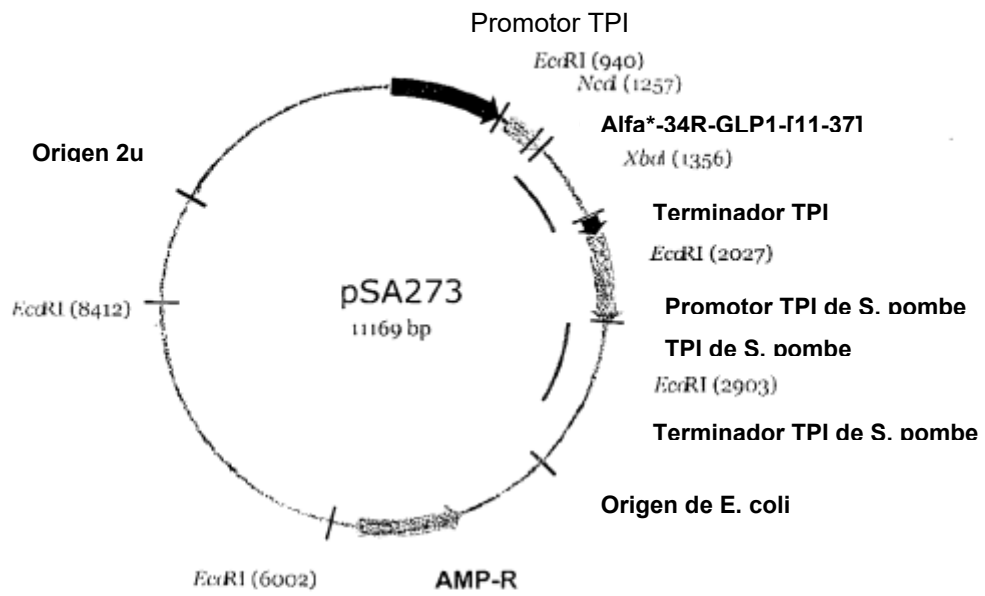


Fig. 1

a)

NcoI
~~~~~

GlyThrSerMet AlaLysArg ThrPheThr SerAspValSer SerTyrLeu GluGlyGln  
 1 GGTACCTCCA TGGCTAAAAG AACTTTTFACT TCTGATGTTT CTCCTTATTT GGAAGGTCAA  
 CCATGGAGGT ACCGATTTTC TTGAAAATGA AGACTACAAA GAAGAATAAA CCTTCCAGTT  
XbaI  
~~~~~

AlaAlaLysGlu PheIleAla TrpLeuVal ArgGlyArgGly ***
 61 GCTGCTAAAG AATTTATTGC TTGGTTGGTT AGAGGTAGAG GTTAGATCTA GAGATCAT
 CGACGATTTTC TTAAATAACG AACCAACCAA TCTCCATCTC CAATCTAGAT CTCTAGTA

b)

NcoI
~~~~~

GlyThrSerMet AlaLysArg GluGluAla GluLysAlaGlu GlyThrPhe ThrSerAsp  
 1 GGTACCTCCA TGGCTAAAAG AGAAGAAGCT GAAAAGGCTG AAGGTACTTT TACTTCTGAT  
 CCATGGAGGT ACCGATTTTC TCTTCTTCGA CTTTTCCGAC TTCCATGAAA ATGAAGACTA

**ValSerSerTyr LeuGluGlu GlnAlaAla ArgGluPheIle AlaTrpLeu ValArgGly**  
 61 GTTCTTCTT ATTTGGAAGA ACAAGCTGCT AGAGAATTTA TTGCTTGGTT GGTAGAGGT  
 CAAAGAAGAA TAAACCTTCT TGTTTCGACGA TCTCTTAAAT AACGAACCAA CCAATCTCCA  
XbaI  
~~~~~

ArgGlyLysGlu AlaGlu***
 121 AGAGGTAAAG AAGCTGAATA GAATCTAGAG ATCAT
 TCTCCATTTT TCGACTTAT CTTAGATCTC TAGTA

c)

NcoI
~~~~~

GlyThrSerMet AlaLysArg GluGluAla GluLysAlaGlu GlyThrPhe ThrSerAsp  
 1 GGTACCTCCA TGGCTAAAAG AGAAGAAGCT GAAAAGGCTG AAGGTACTTT TACTTCTGAT  
 CCATGGAGGT ACCGATTTTC TCTTCTTCGA CTTTTCCGAC TTCCATGAAA ATGAAGACTA

**ValSerSerTyr LeuGluGlu GlnAlaAla ArgGluPheIle AlaTrpLeu ValArgGly**  
 61 GTTCTTCTT ATTTGGAAGA ACAAGCTGCT AGAGAATTTA TTGCTTGGTT GGTAGAGGT  
 CAAAGAAGAA TAAACCTTCT TGTTTCGACGA TCTCTTAAAT AACGAACCAA CCAATCTCCA  
XbaI  
~~~~~

ArgLysGluAla Glu***
 121 AGAAAAGAAG CTGAATAGAA TCTAGAGATC AT
 TCTTTTCTT GACTTATCTT AGATCTCTAG TA

Fig. 2 a)-c)

d)

```

      NcoI
      ~~~~~
GlyValSerMet AlaLysArg GluGlyThr PheThrSerAsp ValSerSer TyrLeuGlu
1  GGGGTATCCA TGGCTAAGAG AGAAGGTACC TTCACCTCTG ACGTCTCGAG TTA CTTGGAA
   CCCCATAGGT ACCGATTCTC TCTTCCATGG AAGTGGAGAC TGCAGAGCTC AATGAACCTT

GlyGlnAlaAla LysGluPhe IleAlaTrp LeuValArgGly ArgGly***
61 GGCCAAGCTG CTAAGGAGTT CATCGCTTGG TTGGTTAGAG GCCGCGGTTA GACGCAGCCC
   CCGGTTTCGAC GATTCTCTCA GTAGCGAACC AACCAATCTC CGGCGCCAAT CTGCGTCGGG
      XbaI
      ~~~~~
121 GCAGGCTCTA GAAACTAA
   CGTCCGAGAT CTTTGATT

```

Fig. 2 d)


```

KR)
      NcoI
      -----
GlyValSerMet AlaLysArg ThrPheThr SerAspValSer SerTyrLeu GluGlyGln
235 GGGGTATCCA TGGCTAAAAG AACTTTTACT TCTGATGTTT CTTCTTATTT GGAAGGTCAA
      CCCCATAGGT ACCGATTTTC TTGAAAATGA AGACTACAAA GAAGAATAAA CCTTCCAGT
                                          XbaI
                                          -----
AlaAlaLysGlu PheIleAla TrpLeuVal ArgGlyArgGly ***
295 GCTGCTAAAG AATTTATTGC TTGGTTGGTT AGAGGTAGAG GTTAGATCTA GA
      CGACGATTC TAAATAACG AACCAACCAA TCTCCATCTC CAATCTAGAT CT
    
```

Fig. 3

A)

NcoI
~~~~~

GlyValSerMet **AlaArgAla ArgTyrLys ArgThrPheThr** SerAspVal SerSerTyr  
 235 GGGGTATCCA TGGCTAGAGC TAGATATAAA AGAACTTTTA CTTCTGATGT TTCTTCTTAT  
 CCCCATAGGT ACCGATCTCG ATCTATATTT TCTTGAAAAT GAAGACTACA AAGAAGAATA XbaI  
 ~~~

LeuGluGlyGln AlaAlaLys GluPheIle AlaTrpLeuVal ArgGlyArg Gly***
 295 TTGGAAGGTC AAGCTGCTAA AGAATTTATT GCTTGGTTGG TTAGAGGTAG AGGTTAGATC
 AACCTTCCAG TTCGACGATT TCTTAAATAA CGAACCAACC AATCTCCATC TCCAATCTAG XbaI
 ~~~

355 TAGAAAC  
 ATCTTTG

B)

NcoI  
~~~~~

GlyValSerMet **AlaArgAsp LeuGlyLys ArgThrPheThr** SerAspVal SerSerTyr
 235 GGGGTATCCA TGGCTAGAGA TTTGGGTAAA AGAACTTTTA CTTCTGATGT TTCTTCTTAT
 CCCCATAGGT ACCGATCTCT AAACCCATTT TCTTGAAAAT GAAGACTACA AAGAAGAATA XbaI
 ~~~

LeuGluGlyGln AlaAlaLys GluPheIle AlaTrpLeuVal ArgGlyArg Gly\*\*\*  
 295 TTGGAAGGTC AAGCTGCTAA AGAATTTATT GCTTGGTTGG TTAGAGGTAG AGGTTAGATC  
 AACCTTCCAG TTCGACGATT TCTTAAATAA CGAACCAACC AATCTCCATC TCCAATCTAG XbaI  
 ~~~

355 TAGAAAC
 ATCTTTG

C)

NcoI
~~~~~

GlyValSerMet **AlaArgAsp LeuAlaLys ArgThrPheThr** SerAspVal SerSerTyr  
 235 GGGGTATCCA TGGCTAGAGA TTTGGCTAAA AGAACTTTTA CTTCTGATGT TTCTTCTTAT  
 CCCCATAGGT ACCGATCTCT AAACCGATTT TCTTGAAAAT GAAGACTACA AAGAAGAATA XbaI  
 ~~~

LeuGluGlyGln AlaAlaLys GluPheIle AlaTrpLeuVal ArgGlyArg Gly***
 295 TTGGAAGGTC AAGCTGCTAA AGAATTTATT GCTTGGTTGG TTAGAGGTAG AGGTTAGATC
 AACCTTCCAG TTCGACGATT TCTTAAATAA CGAACCAACC AATCTCCATC TCCAATCTAG XbaI
 ~~~

355 TAGAAAC  
 ATCTTTG

**Fig. 4 A)–C)**

D)

NcoI  
~~~~~

GlyValSerMet **AlaArgAla ArgAlaLys ArgThrPheThr SerAspVal SerSerTyr**
 235 GGGGTATCCA TGGCTAGAGC TAGAGCTAAA AGAACTTTTA CTTCTGATGT TTCTTCTTAT
 CCCCATAGGT ACCGATCTCG ATCTCGATTT TCTTGAAAAT GAAGACTACA AAGAAGAATA
XbaI
~~

LeuGluGlyGln AlaAlaLys GluPheIle AlaTrpLeuVal ArgGlyArg Gly***
 295 TTGGAAGGTC AAGCTGCTAA AGAATTTATT GCTTGGTTGG TTAGAGGTAG AGGTTAGATC
 AACCTTCCAG TTCGACGATT TCTTAAATAA CGAACCAACC AATCTCCATC TCCAATCTAG
 XbaI
~~~~

355 TAGAAAC  
 ATCTTTG

E)

NcoI  
~~~~~

GlyValSerMet **AlaArgAla LeuAspLys ArgThrPheThr SerAspVal SerSerTyr**
 235 GGGGTATCCA TGGCTAGAGC TTTGGATAAA AGAACTTTTA CTTCTGATGT TTCTTCTTAT
 CCCCATAGGT ACCGATCTCG AAACCTATTT TCTTGAAAAT GAAGACTACA AAGAAGAATA
XbaI
~~

LeuGluGlyGln AlaAlaLys GluPheIle AlaTrpLeuVal ArgGlyArg Gly***
 295 TTGGAAGGTC AAGCTGCTAA AGAATTTATT GCTTGGTTGG TTAGAGGTAG AGGTTAGATC
 AACCTTCCAG TTCGACGATT TCTTAAATAA CGAACCAACC AATCTCCATC TCCAATCTAG
 XbaI
~~~~

355 TAGAAAC  
 ATCTTTG

F)

NcoI  
~~~~~

GlyValSerMet **AlaArgAla LeuAlaLys ArgThrPheThr SerAspVal SerSerTyr**
 235 GGGGTATCCA TGGCTAGAGC TTTGGCTAAA AGAACTTTTA CTTCTGATGT TTCTTCTTAT
 CCCCATAGGT ACCGATCTCG AAACCGATTT TCTTGAAAAT GAAGACTACA AAGAAGAATA
XbaI
~~

LeuGluGlyGln AlaAlaLys GluPheIle AlaTrpLeuVal ArgGlyArg Gly***
 295 TTGGAAGGTC AAGCTGCTAA AGAATTTATT GCTTGGTTGG TTAGAGGTAG AGGTTAGATC
 AACCTTCCAG TTCGACGATT TCTTAAATAA CGAACCAACC AATCTCCATC TCCAATCTAG
 XbaI
~~~~

355 TAGAAAC  
 ATCTTTG

**Fig. 4 D)-F)**

G)

NcoI  
~~~~~

GlyValSerMet **AlaArgAsp LeuGlyLys ArgGluAlaThr** PheThrSer AspValSer
 235 GGGGTATCCA TGGCTAGAGA TTTGGGTAAA AGAGAAGCTA CTTTACTTC TGATGTTTCT
 CCCCATAGGT ACCGATCTCT AAACCCATTT TCTCTTCGAT GAAAATGAAG ACTACAAAGA
 SerTyrLeuGlu GlyGlnAla AlaLysGlu PheIleAlaTrp LeuValArg GlyArgGly
 295 TCTTATTTGG AAGGTCAAGC TGCTAAAGAA TTTATTGCTT GGTTGGTTAG AGGTAGAGGT
 AGAATAAACC TTCCAGTTCG ACGATTTCTT AAATAACGAA CCAACCAATC TCCATCTCCA
 XbaI
~~~~~

\*\*\*

355 TAGAATCTAG AAAC  
 ATCTTAGATC TTTG

H)

NcoI  
~~~~~

GlyValSerMet **AlaProArg AspLeuGly LysArgThrPhe** ThrSerAsp ValSerSer
 235 GGGGTATCCA TGGCTCCAAG AGATTTGGGT AAAAGAACTT TTAFTTCTGA TGTTTCTTCT
 CCCCATAGGT ACCGAGGTTC TCTAAACCCA TTTTCTTGAA AATGAAGACT ACAAGAAGA
 TyrLeuGluGly GlnAlaAla LysGluPhe IleAlaTrpLeu ValArgGly ArgGly***
 295 TATTTGGAAG GTCAAGCTGC TAAAGAATTT ATTGCTTGGT TGGTTAGAGG TAGAGGTTAG
 ATAAACCTTC CAGTTCGACG ATTTCTTAAA TAACGAACCA ACCAATCTCC ATCTCCAATC
 XbaI
~~~~~

355 AATCTAGAAA C  
 TTAGATCTTT G

I)

NcoI  
~~~~~

GlyValSerMet **AlaArgPro LeuGlyLys ArgThrPheThr** SerAspVal SerSerTyr
 235 GGGGTATCCA TGGCTAGACC ATTGGGTAAA AGAACTTTTA CTTCTGATGT TTCTTCTTAT
 CCCCATAGGT ACCGATCTGG TAACCCATTT TCTTGAAAAT GAAGACTACA AAGAAGAATA
 XbaI

LeuGluGlyGln AlaAlaLys GluPheIle AlaTrpLeuVal ArgGlyArg Gly***
 295 TTGGAAGGTC AAGCTGCTAA AGAATTTATT GCTTGGTTGG TTAGAGGTAG AGGTTAGAAT
 AACCTTCCAG TTCGACGATT TCTTAAATAA CGAACCAACC AATCTCCATC TCCAATCTTA
 XbaI
~~~~~

355 CTAGAAAC  
 GATCTTTG

**Fig. 4 G)-I)**

J)

NcoI  
-----

GlyValSerMet Ala**ArgAsp** **LeuGlyLys** **ArgGluAlaGln** ThrPheThr SerAspVal  
 235 GGGCTATCCA TGGCTAGAGA TTTGGGTAAA AGAGAAGCTC AAAC TTTTAC TTCTGATGTT  
 CCCCATAGGT ACCGATCTCT AAACCCATTT TCTCTTCGAG TTTGAAAATG AAGACTACAA  
 SerSerTyrLeu GluGlyGln AlaAlaLys GluPheIleAla TrpLeuVal ArgGlyArg  
 295 TCTTCTTATT TGGAAAGGTCA AGCTGCTAAA GAATTTATTG CTTGGTTGGT TAGAGGTAGA  
 AGAAGAATAA ACCTTCCAGT TCGACGATTT CTTAAATAAC GAACCAACCA ATCTCCATCT  
 XbaI  
-----

Gly\*\*\*  
 355 GGTTAGAATC TAGAAAC  
 CCAATCTTAG ATCTTTG

K)

NcoI  
-----

GlyValSerMet Ala**ArgAsp** **LeuGlyLys** **ArgGluAlaGlu** AlaThrPhe ThrSerAsp  
 235 GGGGTATCCA TGGCTAGAGA TTTGGGTAAA AGAGAAGCTG AAGCTACTTT TACTTCTGAT  
 CCCCATAGGT ACCGATCTCT AAACCCATTT TCTCTTCGAC TTCGATGAAA ATGAAGACTA  
 ValSerSerTyr LeuGluGly GlnAlaAla LysGluPheIle AlaTrpLeu ValArgGly  
 295 GTTCTTCTT ATTTGGAAGG TCAAGCTGCT AAAGAATTTA TTGCTTGGTT GGTTAGAGGT  
 CAAAGAAGAA TAAACCTTCC AGTTCGACGA TTTCTTAAAT AACGAACCAA CCAATCTCCA  
 XbaI  
-----

355 AGAGGTTAGA ATCTAGAAAC  
 TCTCCAATCT TAGATCTTTG

K5)

NcoI  
-----

GlyValSerMet Ala**ArgAsp** **LeuGlyLys** **ArgGluAlaGlu** LeuGluLys ArgThrPhe  
 235 GGGGTATCCA TGGCTAGAGA TTTGGGTAAA AGAGAAGCTG AATTGGAAAA AAGAACTTTT  
 CCCCATAGGT ACCGATCTCT AAACCCATTT TCTCTTCGAC TTAACCTTTT TTCTTGAAAA  
 ThrSerAspVal SerSerTyr LeuGluGly GlnAlaAlaLys GluPheIle AlaTrpLeu  
 295 ACTTCTGATG TTTCTTCTTA TTGGAAGGT CAAGCTGCTA AAGAATTTAT TGCTTGGTTG  
 TGAAGACTAC AAAGAAGAAT AAACCTTCCA GTTCGACGAT TTCTTAAATA ACGAACCAAC  
 XbaI  
-----

ValArgGlyArg Gly\*\*\*  
 355 GTTAGAGGTA GAGGTTAGAA TCTAGAAAC  
 CAATCTCCAT CTCCAATCTT AGATCTTTG

**Fig. 4 J)–K5)**

**K6)**

NcoI  
~~~~~

GlyValSerMet **AlaArgAsp LeuGlyArg ArgGluAlaGlu** LeuGluLys ArgThrPhe
 235 GGGGTATCCA TGGCTAGAGA TTTGGGTAGA AGAGAAGCTG AATTGGAAAA AAGAACTTTT
 CCCCATAGGT ACCGATCTCT AAACCCATCT TCTCTTCGAC TTAACCTTTT TTCTTGAAAA
 ThrSerAspVal SerSerTyr LeuGluGly GlnAlaAlaLys GluPheIle AlaTrpLeu
 295 ACTTCTGATG TTTCTTCTTA TTTGGAAGGT CAAGCTGCTA AAGAATTTAT TGCTTGGTTG
 TGAAGACTAC AAAGAAGAAT AAACCTTCCA GTTCGACGAT TTCTTAAATA ACCGAACCAAC
 XbaI
 ~~~~~

ValArgGlyArg Gly\*\*\*  
 355 GTTAGAGGTA GAGGTTAGAA TCTAGAAAC  
 CAATCTCCAT CTCCAATCTT AGATCTTTG

**K7)**

NcoI  
~~~~~

GlyValSerMet **AlaArgAsp LeuGlyGlu AlaGluLeuGlu LysArgThr** PheThrSer
 235 GGGGTATCCA TGGCTAGAGA TTTGGGTGAA GCTGAATTGG AAAAAAGAAC TTTTACTTCT
 CCCCATAGGT ACCGATCTCT AAACCCACTT CGACTTAACC TTTTTTCTTG AAAATGAAGA
 AspValSerSer TyrLeuGlu GlyGlnAla AlaLysGluPhe IleAlaTrp LeuValArg
 295 GATGTTTCTT CTTATTTGGA AGGTCAAGCT GCTAAAGAAT TTATTGCTTG GTTGGTTAGA
 CTACAAAGAA GAATAAACCT TCCAGTTCGA CGATTCTTA AATAACGAAC CAACCAATCT
 XbaI
 ~~~~~

GlyArgGly\*\*\*  
 355 GGTAGAGGTT AGAATCTAGA AAC  
 CCATCTCCAA TCTTAGATCT TTG

**L)**

NcoI  
~~~~~

GlyValSerMet **AlaArgAsp LeuGlyLys ArgGluAlaGlu AlaGlnThr** PheThrSer
 235 GGGGTATCCA TGGCTAGAGA TTTGGGTAAA AGAGAAGCTG AAGCTCAAAC TTTTACTTCT
 CCCCATAGGT ACCGATCTCT AAACCCATTT TCTCTTCGAC TTCGAGTTG AAAATGAAGA
 AspValSerSer TyrLeuGlu GlyGlnAla AlaLysGluPhe IleAlaTrp LeuValArg
 295 GATGTTTCTT CTTATTTGGA AGGTCAAGCT GCTAAAGAAT TTATTGCTTG GTTGGTTAGA
 CTACAAAGAA GAATAAACCT TCCAGTTCGA CGATTCTTA AATAACGAAC CAACCAATCT
 XbaI
 ~~~~~

GlyArgGly\*\*\*  
 355 GGTAGAGGTT AGAATCTAGA AAC  
 CCATCTCCAA TCTTAGATCT TTG

**Fig. 4 K6)–L)**

L2)

NcoI  
-----

GlyValSerMet Ala**ArgAsp** Leu**GlyLys** Arg**GluAlaGlu** Ala**GlnLys** Arg**ThrPhe**  
 235 GGGGTATCCA TGGCTAGAGA TTTGGGTAAA AGAGAAGCTG AAGCTCAAAA AAGAACTTTT  
 CCCCATAGGT ACCGATCTCT AAACCCATTT TCTCTTCGAC TTCGAGTTTT TTCTTGAAAA  
 ThrSerAspVal SerSerTyr LeuGluGly GlnAlaAlaLys GluPheIle AlaTrpLeu  
 295 ACTTCTGATG TTTCTCTTA TTTGGAAGGT CAAGCTGCTA AAGAATTTAT TGCTTGGTTG  
 TGAAGACTAC AAAGAAGAAT AAACCTTCCA GTTCGACGAT TTCTTAAATA ACGAACCAAC  
 XbaI  
 -----

ValArgGlyArg Gly\*\*\*  
 355 GTTAGAGGTA GAGGTTAGAA TCTAGAAAC  
 CAATCTCCAT CTCCAATCTT AGATCTTTG

L3)

NcoI  
-----

GlyValSerMet Ala**ArgAsp** Leu**GlyArg** Arg**GluAlaGlu** Ala**GlnLys** Arg**ThrPhe**  
 235 GGGGTATCCA TGGCTAGAGA TTTGGGTAGA AGAGAAGCTG AAGCTCAAAA AAGAACTTTT  
 CCCCATAGGT ACCGATCTCT AAACCCATCT TCTCTTCGAC TTCGAGTTTT TTCTTGAAAA  
 ThrSerAspVal SerSerTyr LeuGluGly GlnAlaAlaLys GluPheIle AlaTrpLeu  
 295 ACTTCTGATG TTTCTCTTA TTTGGAAGGT CAAGCTGCTA AAGAATTTAT TGCTTGGTTG  
 TGAAGACTAC AAAGAAGAAT AAACCTTCCA GTTCGACGAT TTCTTAAATA ACGAACCAAC  
 XbaI  
 -----

ValArgGlyArg Gly\*\*\*  
 355 GTTAGAGGTA GAGGTTAGAA TCTAGAAAC  
 CAATCTCCAT CTCCAATCTT AGATCTTTG

L4)

NcoI  
-----

GlyValSerMet Ala**ArgAsp** Leu**GlyGlu** Ala**GluAlaGln** Lys**ArgThr** Phe**ThrSer**  
 235 GGGGTATCCA TGGCTAGAGA TTTGGGTGAA GCTGAAGCTC AAAAAAGAAC TTTTACTTCT  
 CCCCATAGGT ACCGATCTCT AAACCCACTT CGACTTCGAG TTTTCTCTTG AAAATGAAGA  
 AspValSerSer TyrLeuGlu GlyGlnAla AlaLysGluPhe IleAlaTrp LeuValArg  
 295 GATGTTTCTT CTTATTTGGA AGGTCAAGCT GCTAAAGAAT TTATTGCTTG GTTGGTTAGA  
 CTACAAAGAA GAATAAACCT TCCAGTTCGA CGATTTCTTA AATAACGAAC CAACCAATCT  
 XbaI  
 -----

GlyArgGly\*\*\*  
 355 GGTAGAGGTT AGAATCTAGA AAC  
 CCATCTCCAA TCTTAGATCT TTG

Fig. 4 L2)–L4)

M)

NcoI  
~~~~~

GlyValSerMet Ala**GluArg** Leu**GluArg** AspLeu**GlyLys** ArgThrPhe ThrSerAsp
 235 GGGGTATCCA TGGCTGAAAG ATTGGAAAGA GATTGGGTA AAAGAACTTT TACTTCTGAT
 CCCCATAGGT ACCGACTTTC TAACCTTTCT CTAAACCCAT TTTCTTGAAA ATGAAGACTA
 ValSerSerTyr LeuGluGly GlnAlaAla LysGluPheIle AlaTrpLeu ValArgGly
 295 GTTTCTTCTT ATTTGGAAGG TCAAGCTGCT AAAGAATTTA TTGCTTGGTT GGTTAGAGGT
 CAAAGAAGAA TAAACCTTCC AGTTCGACGA TTTCTTAAAT AACGAACCAA CCAATCTCCA
 XbaI
~~~~~

ArgGly\*\*\*  
 355 AGAGGTTAGA ATCTAGAAAC  
 TCTCCAATCT TAGATCTTTG

N)

NcoI  
~~~~~

GlyValSerMet Ala**LysGlu** ArgLeu**Glu** ArgAspLeu**Gly** LysArgThr PheThrSer
 235 GGGGTATCCA TGGCTAAAGA AAGATTGGAA AGAGATTTGG GTAAAAGAAC TTTTACTTCT
 CCCCATAGGT ACCGATTCTT TTCTAACCTT TCTCTAAACC CATTCTCTTG AAAATGAAGA
 AspValSerSer TyrLeuGlu GlyGlnAla AlaLysGluPhe IleAlaTrp LeuValArg
 295 GATGTTTCTT CTTATTTGGA AGGTCAAGCT GCTAAAGAAT TTATTGCTTG GTTGGTTAGA
 CTACAAAGAA GAATAAACCT TCCAGTTCGA CGATTCTTA AATAACGAAC CAACCAATCT
 XbaI
~~~~~

GlyArgGly\*\*\*  
 355 GGTAGAGGTT AGAATCTAGA AAC  
 CCATCTCCAA TCTTAGATCT TTG

O)

NcoI  
~~~~~

GlyValSerMet Ala**GluArg** Leu**GluLys** ArgThrPheThr SerAspVal SerSerTyr
 235 GGGGTATCCA TGGCTGAAAG ATTGGAAAAA AGAACTTTTA CTTCTGATGT TTCTTCTTAT
 CCCCATAGGT ACCGACTTTC TAACCTTTTT TCTTGAAAAT GAAGACTACA AAGAAGAATA
 XbaI

LeuGluGlyGln AlaAlaLys GluPheIle AlaTrpLeuVal ArgGlyArg Gly***
 295 TTGGAAGGTC AAGCTGCTAA AGAATTTATT GCTTGGTTGG TTAGAGGTAG AGGTTAGAAT
 AACCTTCCAG TTCGACGATT TCTTAAATAA CGAACCAACC AATCTCCATC TCCAATCTTA
 XbaI
~~~~~

355 CTAGAAAC  
 GATCTTTG

Fig. 4 M)-O)



P)

NcoI  
-----

GlyValSerMet Ala**LysGlu ArgLeuGlu** **LysArg**ThrPhe ThrSerAsp ValSerSer  
 235 GGGGTATCCA TGGCTAAAGA AAGATTGGAA AAAAGAACTT TTA~~CT~~TCTGA TGTTTCTTCT  
 CCCCATAGGT ACCGATTCTT TTCTAACCTT TTTTCTTGAA AATGAAGACT ACAAAGAAGA  
 TyrLeuGluGly GlnAlaAla LysGluPhe IleAlaTrpLeu ValArgGly ArgGly\*\*\*  
 295 TATTTGGAAG GTCAAGCTGC TAAAGAATTT ATTGCTTGGT TGGTTAGAGG TAGAGGTTAG  
 ATAAACCTTC CAGTTCGACG ATTTCTTAAA TAACGAACCA ACCAATCTCC ATCTCCAATC  
 XbaI  
-----

355 AATCTAGAAA C  
 TTAGATCTTT G

Q)

NcoI  
-----

GlyValSerMet Ala**ProGlu ArgLeuGlu** **ArgAspLeuGly** **LysArg**Thr PheThrSer  
 235 GGGGTATCCA TGGCTCCAGA AAGATTGGAA AGAGATTGG GTAAAAGAAC TTTACTTCT  
 CCCCATAGGT ACCGAGGTCT TTCTAACCTT TCTCTAAACC CATTTTCTTG AAAATGAAGA  
 AspValSerSer TyrLeuGlu GlyGlnAla AlaLysGluPhe IleAlaTrp LeuValArg  
 295 GATGTTTCTT CTTATTTGGA AGGTCAAGCT GCTAAAGAAT TTATTGCTTG GTTGGTTAGA  
 CTACAAAGAA GAATAACCT TCCAGTTCGA CGATTCTTA AATAACGAAC CAACCAATCT  
 XbaI  
-----

GlyArgGly\*\*\*  
 355 GGTAGAGGTT AGAATCTAGA AAC  
 CCATCTCCAA TCTTAGATCT TTG

R)

NcoI  
-----

GlyValSerMet Ala**ProGlu ArgLeuGlu** **LysArg**ThrPhe ThrSerAsp ValSerSer  
 235 GGGGTATCCA TGGCTCCAGA AAGATTGGAA AAAAGAACTT TTA~~CT~~TCTGA TGTTTCTTCT  
 CCCCATAGGT ACCGAGGTCT TTCTAACCTT TTTTCTTGAA AATGAAGACT ACAAAGAAGA  
 TyrLeuGluGly GlnAlaAla LysGluPhe IleAlaTrpLeu ValArgGly ArgGly\*\*\*  
 295 TATTTGGAAG GTCAAGCTGC TAAAGAATTT ATTGCTTGGT TGGTTAGAGG TAGAGGTTAG  
 ATAAACCTTC CAGTTCGACG ATTTCTTAAA TAACGAACCA ACCAATCTCC ATCTCCAATC  
 XbaI  
-----

355 AATCTAGAAA C  
 TTAGATCTTT G

Fig. 4 P)–R)

S)

NcoI  
~~~~~

GlyValSerMet Ala**GluAla GluAlaArg** AspLeuGlyLys ArgThrPhe ThrSerAsp
 235 GGGGTATCCA TGGCTGAAGC TGAAGCTAGA GATTTGGGTA AAAGAACTTT TACTTCTGAT
 CCCCATAGGT ACCGACTTCG ACTTCGATCT CTAAACCCAT TTTCTTGAAA ATGAAGACTA
 ValSerSerTyr LeuGluGly GlnAlaAla LysGluPheIle AlaTrpLeu ValArgGly
 295 GTTTCTTCTT ATTTGGAAGG TCAAGCTGCT AAAGAATTTA TTGCTTGGETT GGTTAGAGGT
 CAAAGAAGAA TAAACCTTCC AGTTCGACGA TTTCTTAAAT AACGAACCAA CCAATCTCCA
 XbaI
~~~~~

ArgGly\*\*\*  
 355 AGAGGTTAGA ATCTAGAAAC  
 TCTCCAATCT TAGATCTTTG

T)

NcoI  
~~~~~

GlyValSerMet Ala**ProGlu AlaGluAla** ArgAspLeuGly LysArgThr PheThrSer
 235 GGGGTATCCA TGGCTCCAGA AGCTGAAGCT AGAGATTTGG GTAAAAGAAC TTTTACTTCT
 CCCCATAGGT ACCGAGGTCT TCGACTTCGA TCTCTAAACC CATTCTTCTG AAAATGAAGA
 AspValSerSer TyrLeuGlu GlyGlnAla AlaLysGluPhe IleAlaTrp LeuValArg
 295 GATGTTTCTT CTTATTTGGA AGGTCAAGCT GCTAAAGAAT TTATTGCTTG GTTGGTTAGA
 CTACAAAGAA GAATAAACCT TCCAGTTCGA CGATTCTTA AATAACGAAC CAACCAATCT
 XbaI
~~~~~

GlyArgGly\*\*\*  
 355 GGTAGAGGTT AGAATCTAGA AAC  
 CCATCTCCAA TCTTAGATCT TTG

U)

NcoI  
~~~~~

GlyValSerMet Ala**GluGlu AlaGluLys** ArgThrPheThr SerAspVal SerSerTyr
 235 GGGGTATCCA TGGCTGAAGA AGCTGAAAAA AGAACTTTTA CTTCTGATGT TTCTTCTTAT
 CCCCATAGGT ACCGACTTCT TCGACTTTTT TCTTGAAAAAT GAAGACTACA AAGAAGAATA
 XbaI
~~~~~

LeuGluGlyGln AlaAlaLys GluPheIle AlaTrpLeuVal ArgGlyArg Gly\*\*\*  
 295 TTGGAAGGTC AAGCTGCTAA AGAATTTATT GCTTGGTTGG TTAGAGGTAG AGGTTAGAAT  
 AACCTTCCAG TTCGACGATT TCTTAAATAA CGAACCAACC AATCTCCATC TCCAATCTTA  
 XbaI  
~~~~~

355 CTAGAAAC
 GATCTTTG

Fig. 4 s)-U)

V)

NcoI
~~~~~

GlyValSerMet **AlaGluGlu AlaGluArg AspLeuGlyLys ArgThrPhe ThrSerAsp**  
 235 GGGGTATCCA TGGCTGAAGA AGCTGAAAGA GATTTGGGTA AAAGAACTTT TACTTCTGAT  
 CCCCATAGGT ACCGACTTCT TCGACTTTCT CTAAACCCAT TTTCTTGAAA ATGAAGACTA  
 ValSerSerTyr LeuGluGly GlnAlaAla LysGluPheIle AlaTrpLeu ValArgGly  
 295 GTTTCTTCTT ATTTGGAAGG TCAAGCTGCT AAAGAATTTA TTGCTTGGTT GGTTAGAGGT  
 CAAAGAAGAA TAAACCTTCC AGTTCGACGA TTTCTTAAAT AACGAACCAA CCAATCTCCA  
 XbaI  
~~~~~

ArgGly***
 355 AGAGGTTAGA ATCTAGAAAC
 TCTCCAATCT TAGATCTTTG

X)

NcoI
~~~~~

GlyValSerMet **AlaArgAsp LeuGlyGlu GluAlaGluLys ArgThrPhe ThrSerAsp**  
 235 GGGGTATCCA TGGCTAGAGA TTTGGGTGAA GAAGCTGAAA AAAGAACTTT TACTTCTGAT  
 CCCCATAGGT ACCGATCTCT AAACCCACTT CTTCGACTTT TTTCTTGAAA ATGAAGACTA  
 ValSerSerTyr LeuGluGly GlnAlaAla LysGluPheIle AlaTrpLeu ValArgGly  
 295 GTTTCTTCTT ATTTGGAAGG TCAAGCTGCT AAAGAATTTA TTGCTTGGTT GGTTAGAGGT  
 CAAAGAAGAA TAAACCTTCC AGTTCGACGA TTTCTTAAAT AACGAACCAA CCAATCTCCA  
 XbaI  
~~~~~

ArgGly***
 355 AGAGGTTAGA ATCTAGAAAC
 TCTCCAATCT TAGATCTTTG

Y)

NcoI
~~~~~

GlyValSerMet **AlaGluGlu AlaGluLeu AlaLysArgThr PheThrSer AspValSer**  
 235 GGGGTATCCA TGGCTGAAGA AGCTGAATTG GCTAAAAGAA CTTTTACTTC TGATGTTTCT  
 CCCCATAGGT ACCGACTTCT TCGACTTAAC CGATTTTCTT GAAAATGAAG ACTACAAAGA  
 SerTyrLeuGlu GlyGlnAla AlaLysGlu PheIleAlaTrp LeuValArg GlyArgGly  
 295 TCTTATTTGG AAGGTCAAGC TGCTAAAGAA TTTATTGCTT GGTTGGTTAG AGGTAGAGGT  
 AGAATAAACC TTCCAGTTCG ACGATTTCTT AAATAACGAA CCAACCAATC TCCATCTCCA  
 XbaI  
~~~~~

 355 TAGAATCTAG AAAC
 ATCTTAGATC TTTG

Fig. 4 v)–y)

Z)

NcoI
~~~~~

GlyValSerMet Ala**GluGlu** **AlaGluLeu** **GlyLysArg**Thr PheThrSer AspValSer  
 235 GGGGTATCCA TGGCTGAAGA AGCTGAATTG GGTAAAAGAA CTTTACTTC TGATGTTTCT  
 CCCCATAGGT ACCGACTTCT TCGACTTAAC CCATTTTCTT GAAAATGAAG ACTACAAAGA  
 SerTyrLeuGlu GlyGlnAla AlaLysGlu PheIleAlaTrp LeuValArg GlyArgGly  
 295 TCTTATTTGG AAGGTCAAGC TGCTAAAGAA TTTATTGCTT GGTGGTTFAG AGGTAGAGGT  
 AGAATAAACC TTCCAGTTCG ACGATTTCTT AAATAACGAA CCAACCAATC TCCATCTCCA

XbaI  
~~~~~

355 TAGAATCTAG AAAC
 ATCTTAGATC TTTG

Fig. 4 z)