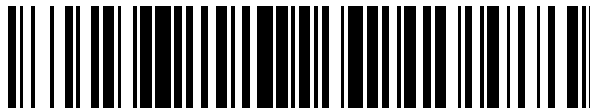


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 477 543**

51 Int. Cl.:

C07D 207/20 (2006.01)

C07D 209/00 (2006.01)

C07K 2/00 (2006.01)

C12P 13/04 (2006.01)

C12P 21/00 (2006.01)

C07K 1/107 (2006.01)

C07K 1/13 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.10.2009 E 09744017 (6)**

97 Fecha y número de publicación de la concesión europea: **16.04.2014 EP 2356092**

54 Título: **Pirrolina-carboxi-lisina generada biosintéticamente y modificaciones de proteínas específicas de sitio mediante derivatización química de restos de pirrolina-carboxi-lisina y pirrolisina**

30 Prioridad:

24.10.2008 US 108434 P

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

17.07.2014

73 Titular/es:

**IRM LLC (100.0%)
131 Front Street P.O. Box HM 2899
Hamilton HM LX, BM**

72 Inventor/es:

**GEIERSTANGER, BERNHARD;
OU, WEIJIA;
CELLITTI, SUSAN E.;
UNO, TETSUO;
CROSSGROVE, TIFFANY;
CHIU, HSIEN-PO;
GRUNEWALD, JAN y
HAO, XUESHI**

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 2 477 543 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

Pirrolina-carboxi-lisina generada biosintéticamente y modificaciones de proteínas específicas de sitio mediante derivatización química de restos de pirrolina-carboxi-lisina y pirrolisina

Campo de la invención

La invención se refiere a la introducción selectiva de aminoácidos genéticamente codificados en proteínas. La invención se refiere también a la derivatización química de dichos aminoácidos.

Antecedente de la invención

Las metilamina metiltransferasas de las archaea metanogénicas de la familia Methanosarcina contienen pirrolisina (PYL) de forma natural. La pirrolisina es un análogo de lisina incorporado traduccionalmente de forma simultánea en codones UAG en el marco en el ARNm respectivo, y se considera el 22º aminoácido natural.

El documento US 2006/166319 A1 describe un método para preparar una célula que, cuando se expone al aminoácido no formal pirrolisina o a uno de sus derivados y se transforma o transfecta con un polinucleótido que comprende un codón UAG o TAG en marco, incorpora el resto de pirrolisina o su derivado en la proteína o polipéptido codificado por el nucleótido.

El documento US 2006/0183888 A1 describe una L-pirrolisina sintética que tiene una fórmula tal como se muestra en el documento US 2006/0183888 A1, y derivados del mismo, métodos químicos para preparar estos aminoácidos y derivados, y los métodos de derivatizar estos aminoácidos tras la incorporación en un péptido o proteína.

Neumann y col. (2008), Nature Chem. Biol., 4, 232-234 describe la incorporación específica de sitio de la N^s-acetillisina en proteínas recombinantes producidas en Escherichia coli mediante la evolución de una pareja de N^s-acetillisil-ARNt sintetasa/ARNt_{CUA}.

Descripción de las figuras

Figura 1. Estructuras of pirrolisina (PYL) y el análogo de pirrolisina pirrolina-carboxi-lisina (PCL: PCL-A o PCL-B).

Figura 2. Comparación de posibles rutas biosintéticas para la pirrolisina (A) y PCL (B). **Figura 3.** Aminoacilación de pylT ARNt con pirrolisina (A) y PCL (B).

Figura 4A. Plásmidos que transportan pylT, pylS, pylB, pylC y pylD para la incorporación de PCL o pirrolisina derivados biosintéticamente en células de mamíferos, y que transportan una construcción génica mutante TAG de sitio único para la proteína modelo hRBP4.

Figura 4B. Construcciones génicas mutantes TAG de sitio único para hRBP4, mEPO, hEPO, y mlgGI. Los sitios únicos están indicados por flechas.

Figura 5. Un plásmido que transportan pylT, pylS, pylB, pylC y pylD para la incorporación de PCL o pirrolisina derivados biosintéticamente en células de Escherichia coli.

Figura 6A. Expresión de hRBP4 en células HEK293F. Construcciones mutantes TAG de hRBP4 (nº 1-9). SDS-PAGE seguido por una transferencia Western con un anticuerpo dirigido contra His. la flecha a 26 kDa indica hRBP4 de longitud completa.

Figura 6B y 6C. SDS-PAGE (A) espectro de masas (B) de hRBP4 purificado

Phe62PCL producido en células HEK293F en presencia de D-ornitina.

Figura 7. El análisis de la espectrometría de masas de una digestión trípica de hRBP4 Phe122PCL indica la incorporación de PCL al sitio de la diana Phe122TAG. Espectro MS/MS asignado de YWGVASF*LQK (F* = PCL) (SEC ID N°:17)

Figura 8. El análisis espectrométrico del espectro de masas (TIC y EIC de 2+ iones de YWGVASF*LQK) (SEC ID N°:17) de una digestión trípica de hRBP4 Phe122PCL indica la incorporación de PCL en el sitio de la diana Phe122TAG.

Figura 9. análisis por espectrometría de masas de una digestión trípica de hRBP4 Phe122PCL. El espectro de masas muestra 3+ y 2+ precursores de YWGVASF*LQK (SEC ID N°:17) indica la incorporación de PCL en el sitio de la diana Phe122TAG.

Figura 10. Detección de PCL, biosintetizado a partir de D-ornitina, en un lisado de células HEK293F.

Figura 11. Incorporación de N-s-ciclopentiloxicarbonil-L-lisina (CYC) en varias proteínas hRBP4 TAG mutantes en células HEK293F. La Figura 11A muestra la incorporación de CYC en diferentes sitios en hRBP4 tal como se ha detectado mediante SDS-PAGE y transferencia Western con anticuerpos dirigidos contra la etiqueta His y anticuerpos dirigidos contra RBP4. La Figura 11B muestra los resultados de SDS-PAGE de hRBP4 Phe62CYC purificada (mutante

nº 2). La Figura 11C muestra un espectro de masas de hRBP4 Phe62CYC.

Figura 12. Verificación por espectrometría de masas de la incorporación de CYC en el sitio TAG de la Phe62CYC mutante de hRBP4.

Figura 13. Incorporación de PCL como función de diversos precursores e incorporación directa de diversos análogos de pirrolisina (incluyendo CYC) en la proteína mutante hRBP4 TAG utilizando células HEK293F (Figura 13A y Figura 13B), e incorporación de PCL utilizando diferentes combinaciones de los genes biosintéticos pylB, pylC y pylD (Figura 13C). La Figura 13A es una transferencia Western de muestras no purificadas con un anticuerpo dirigido contra la etiqueta His y la Figura 13B es un gel SDS-PAGE de la proteína purificada Ni-NTA.

Figura 14. Precursores potenciales de la biosíntesis de PCL (A) y diversos análogos de pirrolisina (B).

Figura 15. Evaluación de diversos precursores y combinaciones génicas para la incorporación de PCL en FASTE en células HK100.

Figura 16. Esquema biosintético potencial para la formación de PCL (PCL-A).

Figura 17A. Incorporación específica del sitio de PCL generado biosintéticamente en los sitios codificados para TAG único (cuatro sitios) en el dominio Fc de la IgG 1 de ratón.

Figura 17B. Incorporación específica del sitio de PCL generado biosintéticamente en los sitios codificados para TAG único (once sitios) en la eritropoyetina (EPO) tal como se detectó mediante SDS-PAGE.

Figura 18. Incorporación específica del sitio de PCL generado biosintéticamente en los sitios codificados para TAG único (dos sitios) en el dominio de la tioesterasa de la ácido graso sintetasa humana (FAS-TE). La Figura 18 muestra la SDS-PAGE (A) y el espectro de masas para la incorporación de PCL (B).

Figura 19. Incorporación específica del sitio de PCL generado biosintéticamente en los sitios codificados para TAG único (un sitio) en FKBP-12. La Figura 19 muestra la SDS-PAGE (A) y el espectro de masas para la incorporación de PCL (B) y un cristal de FKBP12-I90PCL (C).

Figura 20. Incorporación específica del sitio de PCL generado biosintéticamente en los sitios codificados para TAG único (veinte sitios) en el factor de crecimiento de fibroblastos 21 (FGF21). SDS-PAGE muestra la incorporación de PCL en sitios múltiples en FGF21.

Figura 21. La Figura 21A muestra el análisis de SDS-PAGE de la incorporación de PYL o PCL en mTNF- α con glutamina Gln21 (CAA) mutada en un codón de detección TAG en presencia y ausencia de pylB. La Figura 21B SDS-PAGE evalúa la pureza de la proteína. La Figura 21C muestra los espectros de masas intactos de mEGF Tyr10TAG expresados en Escherichia coli que sugieren una mezcla de proteínas con PYL y PCL incorporadas (Figura 21C, parte inferior) y PCL predominantemente (Figura 21C, parte superior).

Figura 22. Esquemas de reacción posibles para la derivatización química de PCL con 2-amino-benzaldehído.

Figura 23. Conjugados de proteínas y cambio de masas postulados tras la derivatización de PCL con 2-amino-benzaldehído, 2-aminoacetofenona y 2-amino-5-nitro-benzofenona.

Figura 24. Análisis por espectrometría de masas de hRBP4 Phe122PCL derivatizada con 2-amino-benzaldehído (2-ABA).

Figura 25. El análisis por espectrometría de masas de una digestión trípica de la proteína hRBP4 Phe122PCL derivatizada con 2-ABA verifica la derivatización del resto de PCL incorporado en el sitio TAG péptido YWGVASF*LQK (SEC ID N°:17)

Figura 26. Evaluación de la dependencia del pH de la derivatización de hRBP4 Phe122PCL con 2-ABA.

Figura 27. Evaluación de la eficacia de reacción como función de la relación de concentración de reactivo a proteína, y de la reactividad con 2-ABA, 2-ANBP y 2-AAP.

Figura 28. Derivatización para relaciones molares más grandes que 4700 (A: 4700 veces en exceso de 2-ABA sobre la proteína) y para hRBP4 incorporada a OMePhe (B: 15400 veces en exceso).

Figura 29. Derivatización de FAS-TE Tyr2454PCL con 2-amino-acetofenona (2-AAP). Espectros de masas de muestras sin reaccionar (A y C) y muestras derivatizadas con 2-AAP a pH 5,0 (B) y pH 7,4 (D).

Figura 30. Esquema de reacción general para la modificación específica de sitio de proteínas mediante la derivatización química de pirrolisina y/o PCL con 2-amino-benzaldehído o análogos de 2-amino-benzaldehído.

Figura 31. Ilustración de una realización de un polímero PEG funcionalizado, 2-amino-acetofenonas-PEG₈ (2-AAP-PEG8; TU3205-044), acopladas a las proteínas mediante un resto de PCL incorporada en una proteína.

Figura 32. Derivatización de hRBP4 Phe122PCL con 2-AAP-PEG8. Espectro de masas hRBP4 con PCL incorporada en la posición 122 tras la derivatización con 2-AAP-PEG8 a pH 7,5 (A) y pH 5,0 (B) en comparación con la de la proteína hRBP4 Phe122PCL sin reaccionar (C) y con hRBP4 natural con (D y E) y sin 2-AAP-PEG8 añadida.

Figura 33. Derivatización de FAS-TE Tyr2454PCL con 2-AAP-PEG8. Espectro de masas de la proteína sin reaccionar (A) y de la proteína FAS-TE Tyr2454PCL reaccionada con 2-AAP-PEG8 (TU3205-044) (B). Las Figuras 33C y 33D muestran la derivatización de FAS-TE Tyr2454PCL con 2,4kDa 2-AAP-PEG (TU3205-048) (Figura 33C derivatizada a temperatura ambiente y Figura 33D a 4°C).

Figura 34. PEGilación de FAS-TE Tyr2454PCL con 0,5kDa 2-AAP-PEG (2-AAP-PEG8), 2,4 kDa 2-AAP-PEG y 23 kDa 2-AAP-PEG en las relaciones molares que se muestran.

Figura 35. Derivatización de FGF21 Lys81PCL con 2-AAP-PEG8. Espectro de masas de FGF21 Lys84PCL sin reaccionar (A) y de FGF21 Lys84PCL tras la derivatización con 2-AAP-PEG8 (B).

Figura 36. PEGilación de proteínas FGF21. Los resultados de SDS-PAGE obtenidos tras la derivatización de siete de los mutantes FGF21 PCL con un PEG-FGF21 que muestra 2-AAP-PEG de 23 kDa, FGF21-PCL de longitud completa (FL) y FGF21-PCL (TR) antes (A) y después la purificación parcial (B).

Figura 37. PEGilación de proteínas EPO. SDS-PAGE tras la derivatización de mutantes EPO PCL de ratón con 2-AAP-PEG de 23 kDa.

Figura 38. Derivatización de PCL con aminoazúcares. Esquema de reacción generalizado donde D-manosamina se acopla a una proteína (ilustrada como R₁) que tiene PCL incorporada en la anterior.

Figura 39. Derivatización de hRBP4 Phe122PCL con D-manosamina. Espectro de masas de hRBP4 con PCL incorporada en la posición 122 tras reacción con manosamina.

Figura 40. Derivatización de FAS-TE Leu2222PCL con D-manosamina. Espectro de masas de la ácido graso sintetasa humana sin reaccionar (FAS-TE) con PCL incorporada en la posición 2222 (FAS-TE Leu2222PCL/Leu2223Ile) (A) y de la proteína reaccionada con manosamina (B).

Figura 41. Ilustración de una realización para la unión específica de sitio de un oligosacáridos a una proteína haciendo reaccionar un resto 2-ABA unido al oligosacárido con PCL incorporada en la proteína.

Figura 42. Ilustración de determinadas realizaciones del conjugado de proteína-proteína (heterodímeros, heterotrímeros, homotrímeros) formados mediante reticulación de proteínas que tienen PCL incorporada en el anterior.

Figura 43. Formación del homodímero con un reticulador bifuncional específico de PCL tal como se ilustra para la proteína FGF21 Lys84PCL. Un ejemplo no limitante de un reticulador bifuncional utilizado para formar un homodímero se muestra en A, y el espectro de masas de la mezcla de reacción de FGF-21 Lys84PCL reticulada utilizando este enlazador bifuncional se muestra en B.

Figura 44. Formación del homodímero de la proteína mutante FGF-21 PCL con un reticulador bifuncional. La Figura 44A muestra el espectro de masas de FGF-21 reticulada, Donde las condiciones de reacción están alteradas a partir de las utilizadas en la Figura 43.

Figura 45. Una realización de un reticulador utilizado para formar trímeros.

Figura 46. Ilustración de diversas realizaciones de marcas específicas de sitio y el marcado utilizando los métodos proporcionados en el presente documento.

Figura 47. La Figura 47A muestra el análisis por espectrometría de masas de mEGF-Y10PCL conjugado con biotina. La Figura 47B muestra la transferencia Western del conjugado de mEGF-Y10PCL-ABA-biotina utilizando una peroxidasa de rábano picante conjugada con un anticuerpo de cabra dirigido contra biotina (HRP). La Figura 47C muestra el análisis por espectrometría de masas ESI de mEGF-Y10PCL conjugado con fluoresceína. La Figura 47D muestra el análisis por espectrometría de masas ESI de mEGF-Y10PCL conjugado con un disacárido.

Figura 48. La Figura 48A muestra el análisis por espectrometría de masas ESI de mEGF-Q21PCL conjugado con un mononitrofenil hapteno. La Figura 48B muestra el análisis por espectrometría de masas ESI de mEGF-Y10PCL

conjugado con un mononitrofenil hapteno. La Figura 48C muestra el análisis por espectrometría de masas ESI de mEGF-Q21PCL conjugado con un dinitrofenil hapteno. La Figura 48D muestra el análisis por espectrometría de masas ESI de mEGF-Q10PCL conjugado con un dinitrofenil hapteno.

5 **Figura 49.** La Figura 49A muestra el análisis por espectrometría de masas ESI de mEGF-Y 10PCL conjugado con un agonista de TLR7, La Figura 49B muestra el análisis por espectrometría de masas ESI de mEGF-Y10PCL conjugado con un fosfolípido.

10 **Figura 50.** La Figura 50A y la Figura 50B muestran el análisis por espectrometría de masas MALDI-TOF de la conjugación de mTNF-Q21PCL con PX2-PADRE a dos diferentes valores de pH (Figura 50A: pH 5.0; Figura 50B: pH 7,5). La Figura 50C muestra el análisis por espectrometría de masas ESI para la conjugación de mTNF-Q21PCL con BHA-exPADRE.

15 **Figura 51.** La Figura 51 es un espectro de masas ESI que muestra el acoplamiento de BHA-exPADRE con mEGF-Y10PCL.

20 **Figura 52.** La Figura 52A es un ensayo de cambio de geles del acoplamiento de BHA-BG1 (7,4 kDa) y BHA-BG2 (7,4 kDa) con mTNF-Q21PCL (19,3 kDa). La Figura 52B es un ensayo de cambio de geles del acoplamiento de BHA-BG2 (7,4 kDa) con mEGF-Y10PCL (7,2 kDa).

Figura 53. La Figura 53 ilustra una realización de dicha unión orientada específica de sitio.

25 **Figura 54.** La Figura 54A muestra el análisis por espectrometría de masas ESI de hFGF21-K150PCL acoplada con 2-ABA y a continuación reducida con 20 mM de NaCNBH₃ durante 1 hora. La Figura 54B muestra el análisis por espectrometría de masas ESI del conjugado de hFGF21-K150PCL 2-ABA reducido tras haberse dializado en tampón fosfato 10 mM (pH 7,5) e incubado a 50°C durante 1 día.

30 **Figura 55.** La Figura 55 demuestra la estabilidad de la unión de PCL para la FGF21 pegilada con y sin reducción usando NaCNBH₃. Se muestran geles SDS-PAGE de muestras reducidas y muestras no reducidas en la Figura 55A. Además, La Figura 55B muestra un gel SDS-PAGE de las muestras no reducidas incubadas durante 60 horas a 4°C, temperatura ambiente 37°C y 50°C, y 95°C.

Figura 56. Análisis de RMN de la reacción de PCL-A con 2-ABA

35 **Figura 57.** La Figura 57A es una estructura propuesta del producto resultante de la reacción entre PCL-A y 2-ABA. La Figura 57B muestra las estructuras de equilibrio propuesta del producto resultante de la reacción entre PCL-A y 2-ABA. La Figura 57C es una estructura propuesta del producto reducido.

40 **Figura 58.** Análisis de RMN de la reacción de PCL-B con 2-ABA

Figura 59. Derivatización de Pirrolisina (Pyl) y PCL incorporadas en mEGF.

Sumario de la invención

45 Se proporcionan en el presente documento proteínas y/o polipéptidos que tienen una o más PCL incorporadas en el anterior, donde la PCL se genera biosintéticamente y se incorpora a las proteínas y/o polipéptidos. Se proporcionan también en el presente documento proteínas y/o polipéptidos que tienen una o más pirrolisinas (PYL) incorporadas en el anterior, donde la PYL se genera biosintéticamente y se incorpora a las proteínas y/o polipéptidos. Se proporcionan también en el presente documento proteínas y/o polipéptidos que tienen una o más pirrolisinas PCL incorporadas en el anterior, donde la PCL y la PYL se generan biosintéticamente y se incorporan en las proteínas y/o polipéptidos.

50 Se proporcionan también en el presente documento proteínas y/o polipéptidos que tienen uno o más restos PCL, donde la PCL se genera biosintéticamente y se incorpora a las proteínas y/o polipéptidos, y los uno o más restos PCL se derivatizan por tanto acoplando a las proteínas y polipéptidos un grupo seleccionado a partir de una marca, un colorante, un polímero, un polímero soluble en agua, un polialquilenglicol, un poli(etilenglicol), un derivado de poli(etilenglicol), un azúcar, un lípido, un fotorreticulador, un compuesto citotóxico, un fármaco, una marca de afinidad, una marca de fotoafinidad, un compuesto reactivo; una resina, un péptido, una segunda proteína o polipéptido o análogo de polipéptido, un anticuerpo o fragmento de anticuerpo, un quelante metálico, un cofactor, un ácido graso, un hidrato de carbono, un polinucleótido, un ADN, un ARN, una sonda de la PCR, un polinucleótido antisentido, un ribooligonucleótido, un desoxirribooligonucleótido, ADN modificado con fosforotioato, ADN y ARN modificados, ácido nucleico peptídico, un sacárido, un disacárido, un oligosacárido, un polisacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial, una nanopartícula, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un grupo funcional novedoso, un grupo que interactúa de forma covalente o no covalente con otras moléculas, un resto fotoestimulado, un resto excitable mediante radiación actínica, un ligando un resto fotoisomerizable, biotina, un análogo de biotina, un resto que incorpora un átomo pesado, un grupo escindible químicamente, un grupo fotoescindible, una cadena lateral alargada, un azúcar unido a carbono, un agente redox

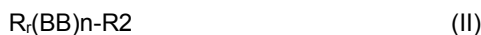
activo, un aminotioácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo cromóforo, un grupo quimioluminiscente, un resto fluorescente, un grupo denso en electrones, un grupo magnético, un grupo intercalante, un grupo quelante, un cromóforo, un agente de transferencia de energía, un agente biológicamente activo, una marca detectable, una molécula pequeña, un ácido nucleico inhibidor, un ARNip, un radionucleótido, un agente de captura de neutrones, un derivado de biotina, punto(s) cuántico(s), un nanotransmisor, un radiotransmisor, una abzima, una enzima, un activador complejo activado, un virus, una toxina, un adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, un epítipo de linfocito T, un fosfolípido, una molécula de tipo LPS, hemocianina de lapa californiana (KLH), un hapteno inmunógeno, un aglicano, un alérgeno, una angioestatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimotopo, un receptor, una micela inversa, detergentes, potenciadores de la respuesta inmune, colorantes fluorescentes, reactivos FRET, sondas de radioimagen, otra sonda de espectroscopía, profármacos, toxinas para inmunoterapia, un soporte sólido, un $-C_{n_2}C_{n_2}-(OCH_2CH_2O)_p-OX^2$, y un $-O-(CH_2CH_2O)_pCH_2CH_2-X^2$ (donde p es 1 a 10.000 y X^2 es H, un alquilo C_{1-8} , un grupo protector o un grupo funcional terminal).

En determinadas realizaciones, dichas proteínas y/o polipéptidos que tienen uno o más restos PYL incorporados en el anterior, donde la PYL se genera biosintéticamente y se incorpora a las proteínas y/o polipéptidos, la pirrolisina se derivatiza acoplando por tanto a las proteínas y/o polipéptidos uno de los grupos anteriormente mencionados proporcionados anteriormente para las proteínas y/o polipéptidos que tienen uno o más restos PCL1 incorporados en la anterior.

En determinadas realizaciones, dichas proteínas y/o polipéptidos que tienen uno o más restos PCL y uno o más restos PYL incorporados en el anterior, donde la PCL y la PYL se generan biosintéticamente y se incorporan a las proteínas y/o polipéptidos, la PCL y la pirrolisina se derivatizan acoplando por tanto a las proteínas y/o polipéptidos uno de los grupos anteriormente mencionados proporcionados anteriormente para las proteínas y/o polipéptidos que tienen uno o más restos PCL1 incorporados en la anterior.

En determinadas realizaciones, dicha biosíntesis anteriormente mencionada se produce en células eucariotas, células de mamíferos, células de levaduras o células de insectos. En determinadas realizaciones, las células son células de *Escherichia coli*, mientras que en otras realizaciones las células de levaduras son células de *Saccharomyces cerevisiae* o *Pichia pastoralis*. En determinadas realizaciones, las células son células CHO, células HeLa, células HEK293F o células sf9.

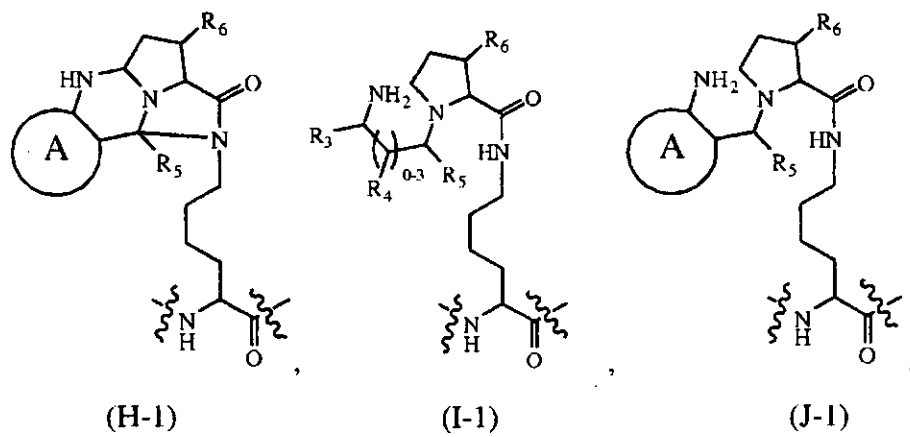
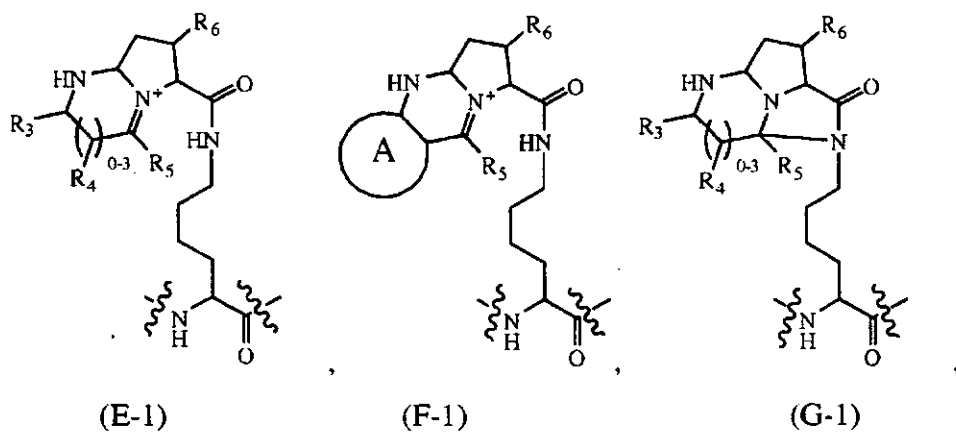
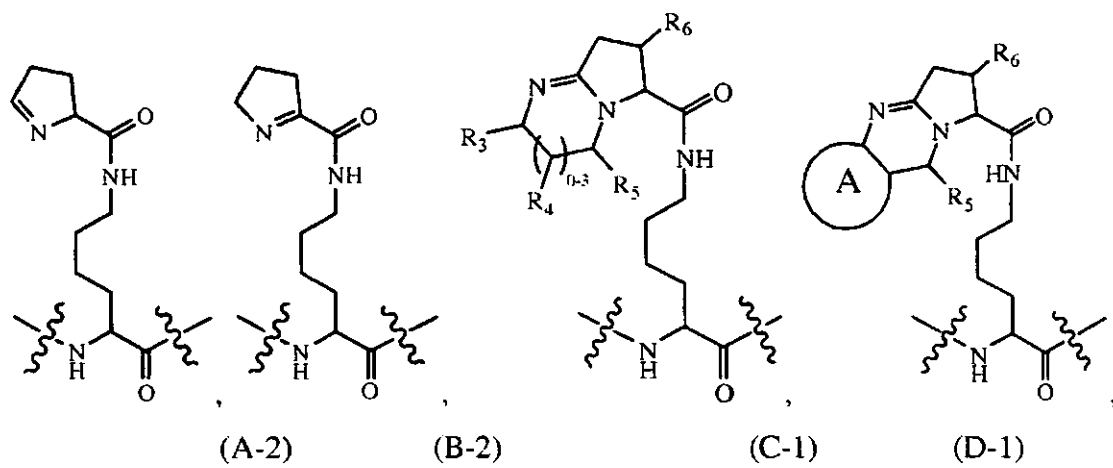
Un aspecto proporcionado en el presente documento son los compuestos que tienen la estructura de la Fórmula (II):

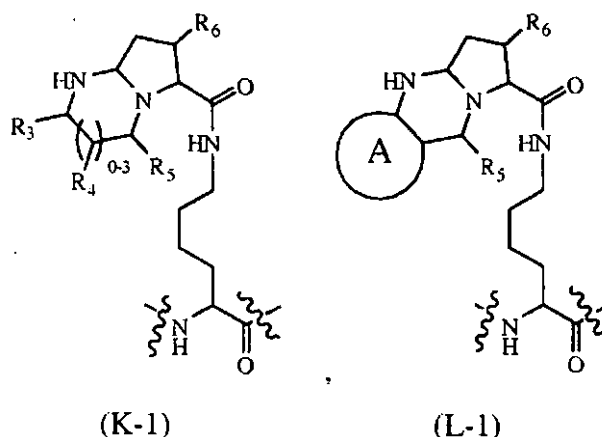


donde:

R_1 es H o un grupo de modificación del amino terminal;
 R_2 es OH o un grupo de modificación del carboxi terminal;
n es un número entero de 1 a 5000;

cada BB se selecciona de forma independiente entre un resto de aminoácido, un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (A-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (B-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (C-1) un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (D-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (E-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (F-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (G-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (H-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (I-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (J-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (K-1) y un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (L-1);



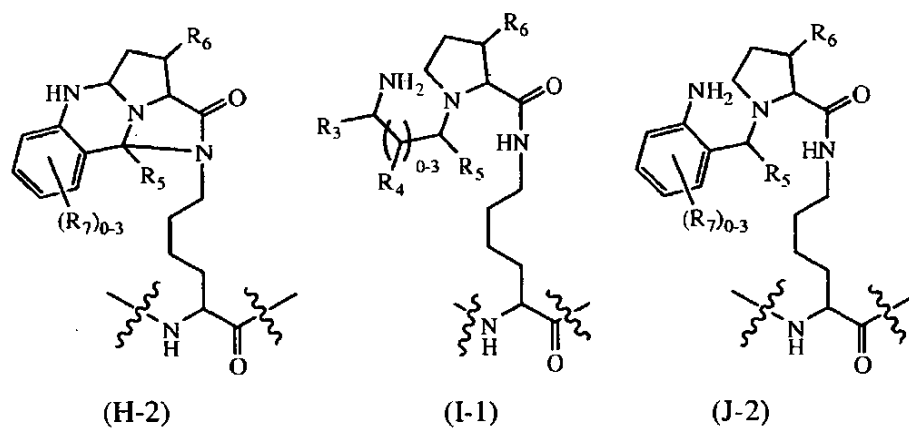
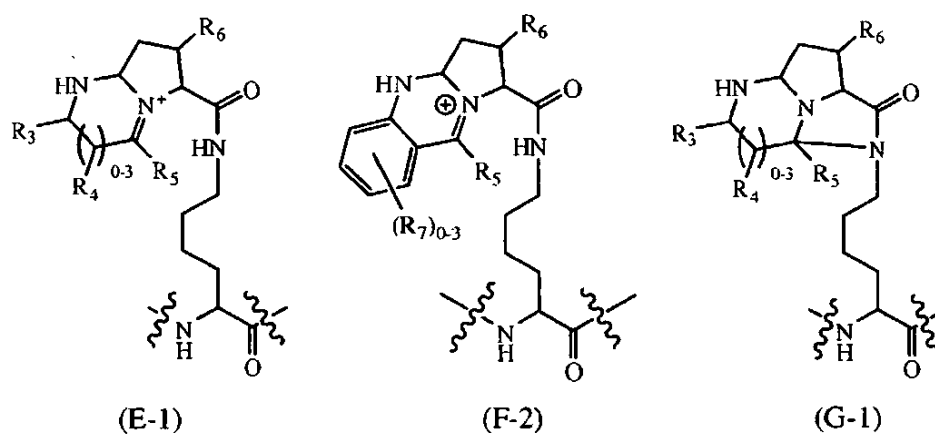
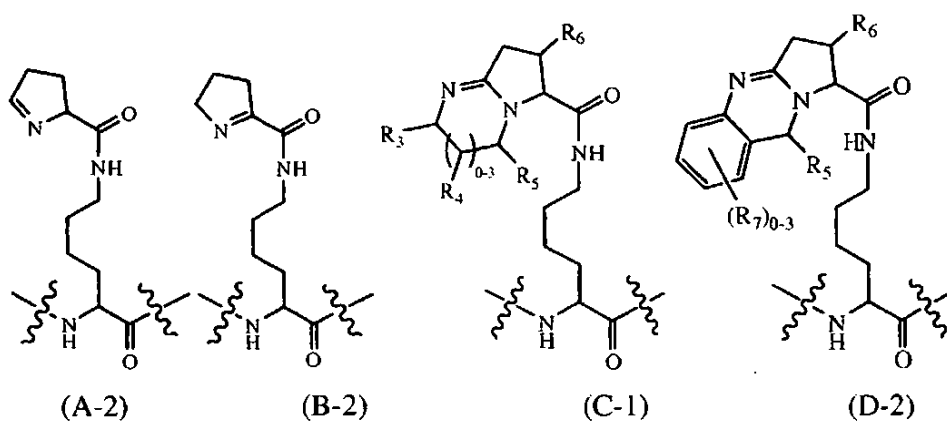


donde:

- 5 R_3 , R_5 y cada R_4 se seleccionan independientemente entre H, -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidroxi-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹; R_6 es H o alquilo C₁;
- 10 A es un cicloalquilo C₃₋₈, heterocicloalquilo C₃₋₈, un arilo monocíclico de 5-6 miembros, un heteroarilo monocíclico de 5-6 miembros, un anillo bicíclico condensado de 9-10 miembros o un anillo tricíclico condensado de 13-14 miembros, en donde A está opcionalmente sustituido con 1 a 5 sustituyentes seleccionados independientemente entre -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidroxi-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹;
- 15 L se selecciona entre un enlace, alquileo C₁₋₈, alquileo C₁₋₈ halo-sustituido, alquileo C₁₋₈ hidroxi-sustituido, alquilenilo C₂₋₈, alquilenilo C₂₋₈ halo-sustituido, alquilenilo C₂₋₈ hidroxi-sustituido, un polialquilenglicol, un polietilenglicol), -O(CR¹¹R¹²)_k-, -S(CR¹¹R¹²)_k-, -S(O)_k(CR¹¹R¹²)_k-, -O(CR¹¹R¹¹)_k-NR¹¹C(O)-, -O(CR¹¹R¹²)_kC(O)NR¹¹-, -C(O)-, -C(O)(CR¹¹R¹²)_k-, -C(S)-, -C(S)(CR¹¹R¹²)_k-, -C(O)NR¹¹-, -NR¹¹C(O)-, -NR¹¹(CR¹¹R¹²)_k-, -CONR¹¹(CR¹¹R¹²)_k-, -N(R¹¹)CO(CR¹¹R¹²)_k-, -C(O)NR¹¹(CR¹¹R¹²)_k-, NR¹¹C(O)(CR¹¹R¹²)_k-, en donde cada R¹¹ y R¹² son independientemente H, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, o alquilo C₁₋₈ hidroxi-sustituido, y k es un número entero de 1 a 12, y
- 20 X¹ se ha seleccionado de una marca, un colorante, un polímero, un polímero soluble en agua, un polialquilenglicol, un poli(etilenglicol), un derivado de poli(etilenglicol), un azúcar, un fotorreticulador, un compuesto citotóxico, un fármaco, una marca de afinidad, una marca de fotoafinidad, un compuesto reactivo; una resina, un péptido, una segunda proteína o polipéptido o análogo de polipéptido, un anticuerpo o fragmento de anticuerpo, un quelante metálico, un cofactor, un ácido graso, un hidrato de carbono, un polinucleótido, un ADN, un ARN, una sonda de la PCR, un polinucleótido antisentido, un ribooligonucleótido, un desoxirribooligonucleótido, ADN modificado con fosforotioato, ADN y ARN modificados, ácido nucleico peptídico, un sacárido, un disacárido, un oligosacárido, un polisacárido, un dendrímero soluble en agua, una ciclodextrina, una biomaterial, una nanopartícula, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un grupo funcional novedoso, un grupo que interactúa de forma covalente o no covalente con otras moléculas, un resto fotoestimulado, un resto excitable mediante radiación actínica, un ligando un resto fotoisomerizable, biotina, un análogo de biotina, un resto que incorpora un átomo pesado, un grupo escindible químicamente, un grupo fotoescindible, una cadena lateral alargada, un azúcar unido a carbono, un agente redox activo, un aminotioácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo cromóforo, un grupo quimioluminiscente, un resto fluorescente, un grupo denso en electrones, un grupo magnético, un grupo intercalante, un grupo quelante, un cromóforo, un agente de transferencia de energía, un agente biológicamente activo, una marca detectable, una molécula pequeña, un ácido nucleico inhibidor, un ARNip, un radionucleótido, un agente de captura de neutrones, un derivado de biotina, punto(s) cuántico(s), un nanotransmisor, un radiotransmisor, una abzima, una enzima, un activador complejo activado, un virus, una toxina, un adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, un epítipo de linfocito T, un fosfolípido, una molécula de tipo LPS, hemocianina de lapa californiana (KLH), un hapteno inmunógeno, un aglicano, un alérgeno, una angioestatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimotopo, un receptor, una micela inversa, detergentes, potenciadores de la respuesta inmune, colorantes fluorescentes, reactivos FRET, sondas de radiografía, otra sonda de espectroscopía, profármacos, toxinas para inmunoterapia, un soporte sólido, -CH₂CH₂-(OCH₂CH₂O)_p-OX², -O-(CH₂CH₂O)_pCH₂CH₂-X², y cualquier combinación de los mismos, donde p es de 1 a 10.000, y X² es H, un alquilo C₁₋₈, un grupo protector o un grupo funcional terminal, y donde al menos un BB es un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (C-1) o la Fórmula (D-1) o la Fórmula (E-1) o la Fórmula (F-1) o la Fórmula (G-1) o la Fórmula (H-1) o la Fórmula (I-1) o la Fórmula (J-1) o la Fórmula (K-1) o la Fórmula (L-1).
- 50

En determinadas realizaciones de dichos compuestos, el anillo A se selecciona entre fenilo, naftaleno y piridina.

En determinadas realizaciones de dichos compuestos, cada BB se selecciona de forma independiente entre un resto de aminoácido, un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (A-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (B-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (C-1) un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (D-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (E-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (F-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (G-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (H-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (I-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (J-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (K-1) y un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (L-2);



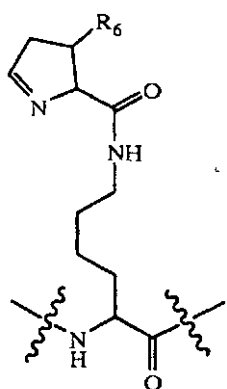
poli(etilenglicol), alquileno C₁₋₈ halo sustituido, o un alquileno C₁₋₈ hidroxi sustituido

Otro aspecto proporcionado en el presente documento es un método para derivatizar una proteína, donde la proteína tiene la estructura de acuerdo con la Fórmula (I), comprendiendo el método poner en contacto la proteína con un reactivo de Fórmula (III) o Fórmula (IV); donde la Fórmula (I) corresponde a:

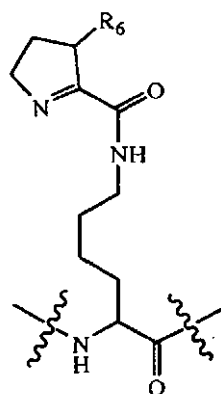


donde:

R₁, es H o un grupo de modificación del amino terminal; R₂ es OH o un grupo de modificación del carboxi terminal; n es un número entero de 1 a 5000; cada AA se selecciona de forma independiente entre un resto de aminoácido, un resto del aminoácido pirrolisina, un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (A-1) y un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (B-1);



(A-1)

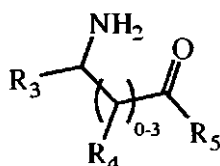


(B-1)

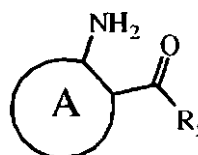
R₆ es H o alquilo C₁, y

al menos un AA es un resto del aminoácido pirrolisina o un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (A-1) o la Fórmula (B-1);

y en el que la Fórmula (II) y la Fórmula (III) corresponden a:



(III)



(IV)

donde:

R₃, R₅ y cada R₄ se seleccionan independientemente entre H, -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidroxi-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹;

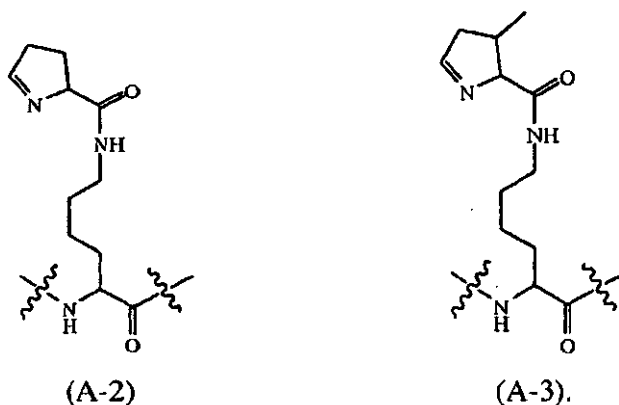
A es un cicloalquilo C₃₋₈, heterocicloalquilo C₃₋₈, un arilo monocíclico de 5-6 miembros, un heteroarilo monocíclico de 5-6 miembros, un anillo bicíclico condensado de 9-10 miembros o un anillo tricíclico condensado de 13-14 miembros, en donde A está opcionalmente sustituido con 1 a 5 sustituyentes seleccionados independientemente entre -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidroxi-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹;

L se selecciona entre un enlace, alquileno C₁₋₈, alquileno C₁₋₈ halo-sustituido, alquileno C₁₋₈ hidroxi-sustituido, alquilenilo C₂₋₈, alquilenilo C₂₋₈ halo-sustituido, alquilenilo C₂₋₈ hidroxi-sustituido, un polialquilenglicol, un polietilenglicol), -O(CR¹¹R¹²)_k-, -S(CR¹¹R¹²)_k-, -S(O)_k(CR¹¹R¹²)_k-, -O(CR¹¹R¹²)_k-NR¹¹C(O)-, -O(CR¹¹R¹²)_kC(O)NR¹¹-, -C(O)-, -C(O)(CR¹¹R¹²)_k-, -C(S)-, -C(S)(CR¹¹R¹²)_k-, -C(O)NR¹¹-, -NR¹¹C(O)-, -NR¹¹(CR¹¹R¹²)_k-, -CONR¹¹(CR¹¹R¹²)_k-, -N(R¹¹)CO(CR¹¹R¹²)_k-, -C(O)NR¹¹(CR¹¹R¹²)_k-, -NR¹¹C(O)(CR¹¹R¹²)_k-, en donde cada R¹¹ y R¹² son independientemente H, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, o alquilo C₁₋₈

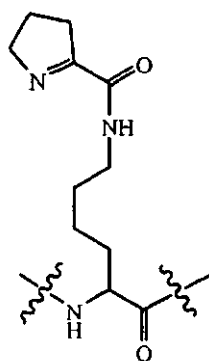
hidroxi-sustituido, y k es un número entero de 1 a 12, y

X¹ se ha seleccionado de una marca, un colorante, un polímero, un polímero soluble en agua, un polialquilenglicol, un poli(etilenglicol), un derivado de poli(etilenglicol), un azúcar, un lípido, un fotorreticulador, un compuesto citotóxico, un fármaco, una marca de afinidad, una marca de fotoafinidad, un compuesto reactivo; una resina, un péptido, una segunda proteína o polipéptido o análogo de polipéptido, un anticuerpo o fragmento de anticuerpo, un quelante metálico, un cofactor, un ácido graso, un hidrato de carbono, un polinucleótido, un ADN, un ARN, una sonda de la PCR, un polinucleótido antisentido, un ribooligonucleótido, un desoxirribooligonucleótido, ADN modificado con fosforotioato, ADN y ARN modificados, ácido nucleico peptídico, un sacárido, un disacárido, un oligosacárido, un polisacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial, una nanopartícula, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un grupo funcional novedoso, un grupo que interactúa de forma covalente o no covalente con otras moléculas, un resto fotoestimulado, un resto excitable mediante radiación actínica, un ligando un resto fotoisomerizable, biotina, un análogo de biotina, un resto que incorpora un átomo pesado, un grupo escindible químicamente, un grupo fotoescindible, una cadena lateral alargada, un azúcar unido a carbono, un agente redox activo, un aminotioácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo cromóforo, un grupo quimioluminiscente, un resto fluorescente, un grupo denso en electrones, un grupo magnético, un grupo intercalante, un grupo quelante, un cromóforo, un agente de transferencia de energía, un agente biológicamente activo, una marca detectable, una molécula pequeña, un ácido nucleico inhibidor, un ARNip, un radionucleótido, un agente de captura de neutrones, un derivado de biotina, punto(s) cuántico(s), un nanotransmisor, un radiotransmisor, una abzima, una enzima, un activador complejo activado, un virus, una toxina, un adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, un epítipo de linfocito T, un fosfolípido, una molécula de tipo LPS, hemocianina de lapa californiana (KLH), un hapteno inmunógeno, un aglicano, un alérgeno, una angioestatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimotopo, un receptor, una micela inversa, detergentes, potenciadores de la respuesta inmune, colorantes fluorescentes, reactivos FRET, sondas de radioimagen, otra sonda de espectroscopía, profármacos, toxinas para inmunoterapia, un soporte sólido, -CH₂CH₂-(OCH₂CH₂O)_p-OX², -O-(CH₂CH₂O)_pCH₂CH₂X², y cualquier combinación de los mismos, donde p es de 1 a 10.000 y X² es H, un alquilo C₁₋₈, un grupo protector o un grupo funcional terminal.

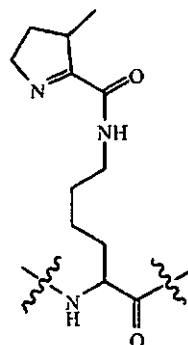
En determinadas realizaciones del método anteriormente mencionado, el resto de aminoácido de la Fórmula (A-1) es un resto de aminoácido que tiene la estructura de la Fórmula (A-2) o la Fórmula (A-3):



y donde el resto de aminoácido de la Fórmula (B-1) es un resto de aminoácido que tiene la estructura de la Fórmula (B-2) o la Fórmula (B-3):



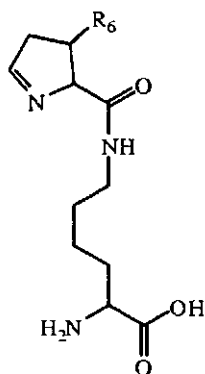
(B-2)



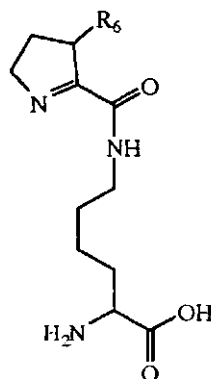
(B-3).

En determinadas realizaciones del método anteriormente mencionado, el resto de aminoácido de la Fórmula (A-1) es un resto de un aminoácido de Fórmula (V) y el resto de aminoácido de la Fórmula (B-1) es un resto de un aminoácido de Fórmula (VI):

5



(V)



(VI),

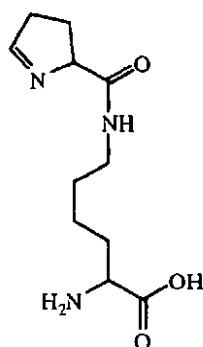
donde R_6 es H o alquilo C_1

10

En determinadas realizaciones, el aminoácido de Fórmula (V) o Fórmula (VI) está generado biosintéticamente en una célula que comprende un gen *pylB*, un gen *pylC* y un gen *pylD*, y la célula está en contacto con un medio de crecimiento que comprende un precursor. En otras realizaciones, el aminoácido de Fórmula (V) o Fórmula (VI) está generado biosintéticamente en una célula que comprende un gen *pylC* y un gen *pylD*, y la célula está en contacto con un medio de crecimiento que comprende un precursor.

15

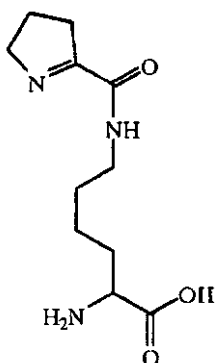
En determinadas realizaciones, el aminoácido de la Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (VII):



(VII),

20

y el aminoácido de la Fórmula (VI) es un aminoácido que tiene la estructura de la Fórmula (VIII)

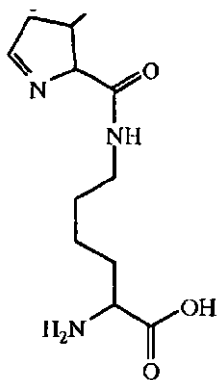


(VIII),

5 y el precursor es D-ornitina, D-arginina, ácido (2S)-2-amino-6-(2,5-diaminopentanamido)hexanoico o ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico.

10 En determinadas realizaciones, el aminoácido de Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (VII) y el precursor es D-ornitina o D-arginina. En determinadas realizaciones, el aminoácido de Fórmula (VI) es un aminoácido que tiene la estructura de la Fórmula (VII) y el precursor es D-ornitina o D-arginina. En determinadas realizaciones, el aminoácido de Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (VII) y el precursor es el ácido (2S)-2-amino-6-(2,5-diaminopentanamido)hexanoico En determinadas realizaciones, el aminoácido de Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (VII) y el precursor es el ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico En determinadas realizaciones, el aminoácido de Fórmula (VI) es un aminoácido que tiene la estructura de la Fórmula (VIII) y el precursor es el ácido (2S)-2-amino-6-(2,5-diaminopentanamido)hexanoico. En determinadas realizaciones, el aminoácido de Fórmula (VI) es un aminoácido que tiene la estructura de la Fórmula (VIII) y el precursor es el ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico.

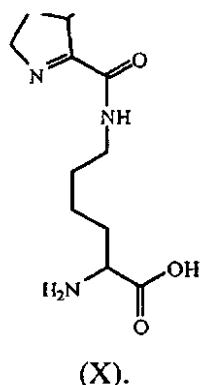
20 En determinadas realizaciones, el aminoácido de Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (IX) y el precursor es D-ornitina, D-arginina o ácido 2,5-diamino-3-metilpentanoico, ácido (2S)-2-amino-6-(2,5-diaminopentanamido)hexanoico o ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico.



(IX).

25 y el aminoácido de Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (X) y el precursor es D-ornitina, D-arginina o ácido 2,5-diamino-3-metilpentanoico, ácido (2R,3S)-2,5-diamino-3-metilpentanoico, ácido (2R,3R)-2,5-diamino-3-metilpentanoico, ~> ácido (2S)-2-amino-6-(2,5-diaminopentanamido)hexanoico o ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico.

30



5 En determinadas realizaciones, el aminoácido de Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (IX) y el precursor es el ácido D-2,5-diamino-3-metilpentanoico. En determinadas realizaciones, el aminoácido de Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (X) y el precursor es el ácido D-2,5-diamino-3-metilpentanoico.

10 En determinadas realizaciones, el aminoácido de Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (IX) y el precursor es el ácido (2R,3S)-2,5-diamino-3-metilpentanoico. En determinadas realizaciones, el aminoácido de Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (X) y el precursor es el ácido (2R,3S)-2,5-diamino-3-metilpentanoico.

15 En determinadas realizaciones, el aminoácido de Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (IX) y el precursor es el ácido (2R,3R)-2,5-diamino-3-metilpentanoico. En determinadas realizaciones, el aminoácido de Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (X) y el precursor es el ácido (2R,3R)-2,5-diamino-3-metilpentanoico. En determinadas realizaciones, el aminoácido de Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (IX) y el precursor es D-ornitina o D-arginina o el ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico. En determinadas realizaciones, el aminoácido de Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (X) y el precursor es D-ornitina o D-arginina o el ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico.

25 En determinadas realizaciones de los métodos anteriormente mencionados, el aminoácido de Fórmula (V), Fórmula (VI), Fórmula (VII), Fórmula (VII), Fórmula (IX) o Fórmula (X) se incorpora a una proteína en la célula mediante un ARNt ortogonal (ARNt-O) y una aminoacil ARNt sintetasa ortogonal (O-RS), donde los aminoacilatos O-RS del ARNt-O con el aminoácido de Fórmula (V) o Fórmula (VI) y el ARNt-O reconocen al menos un codón selector de un ARNm en la célula y el codón selector es un codón ámbar.

30 En determinadas realizaciones de los métodos anteriormente mencionados, la célula comprende además un gen *pylS* y un gen *pylT* y el aminoácido de Fórmula (V), Fórmula (VI), Fórmula (VII), Fórmula (VII), Fórmula (IX) o Fórmula (X) se incorpora a una proteína en la célula mediante una aminoacil ARNt sintetasa y un ARNt que reconoce al menos un codón selector de ARNm en la célula, donde la aminoacil ARNt sintetasa es un producto génico del gen *pylS* y el ARNt es un producto génico del gen *pylT*.

35 En determinadas realizaciones de los métodos anteriormente mencionados, la célula es una célula procariota, mientras que en otras realizaciones la célula es una célula eucariota. En determinadas realizaciones, la célula es una célula de *Escherichia coli*, mientras que en otras realizaciones la célula es una célula de mamífero, una célula de levadura o una célula de insecto. En determinadas realizaciones, la célula de mamífero es una célula CHO, una célula HeLa o una célula HEK293F. En determinadas realizaciones, la célula de insecto es una célula sf9,

40 Otros aspecto proporcionado en el presente documento son proteínas derivatizadas obtenidas usando los métodos anteriormente mencionados, donde dichas proteínas derivatizadas tienen la estructura de acuerdo con la Fórmula (II),

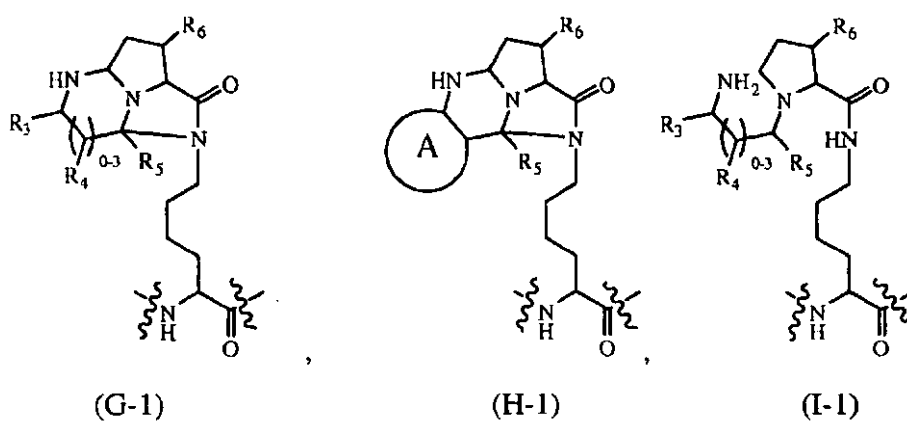
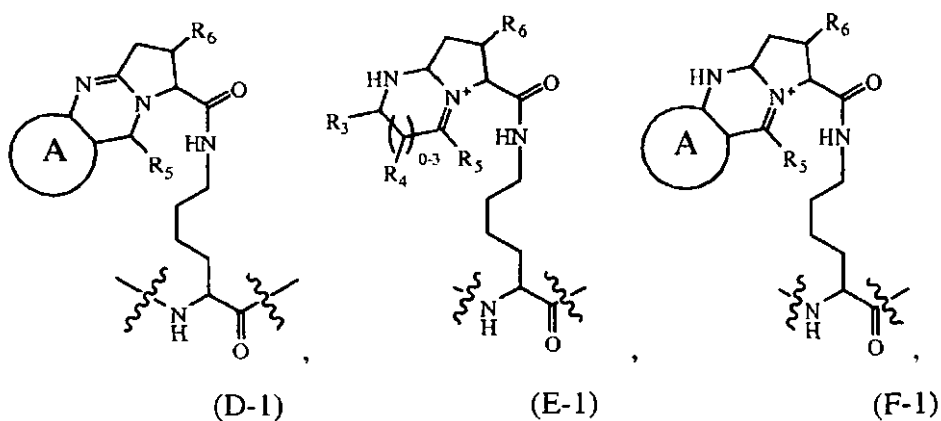
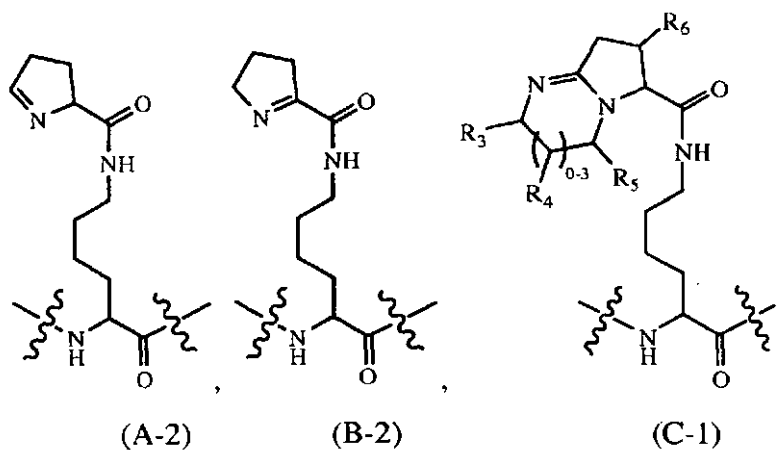


45 donde:

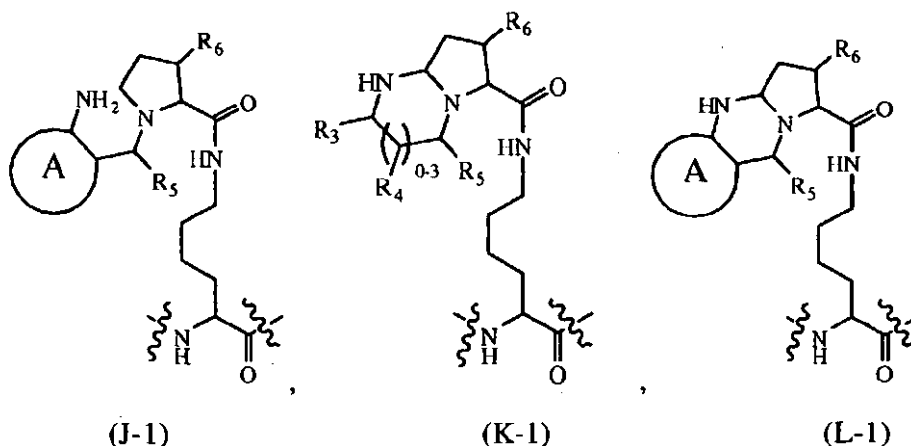
R_1 es H o un grupo de modificación del amino terminal; R_2 es OH o un grupo de modificación del carboxi terminal; n es un número entero de 1 a 5000; cada BB se selecciona de forma independiente entre un resto de aminoácido, un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (A-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (B-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (C-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (D-1), un resto de

5

aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (E-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (F-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (G-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (H-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (I-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (J-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (K-1) y un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (L-1);



10

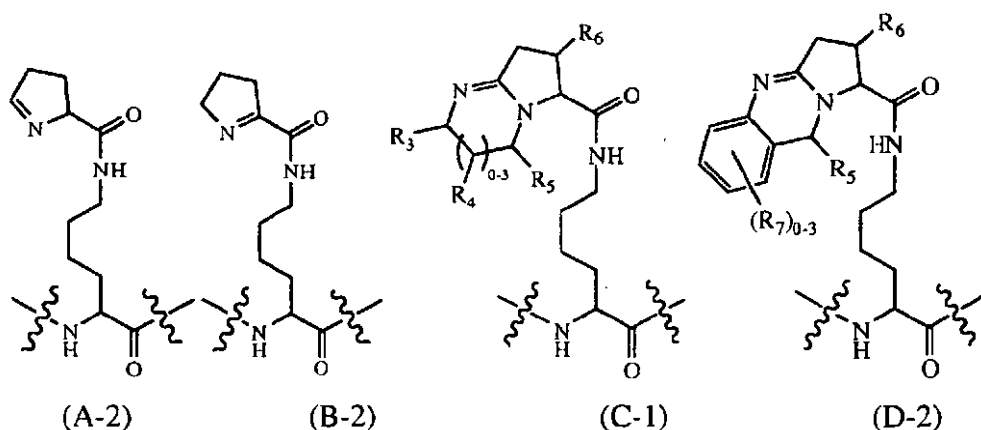


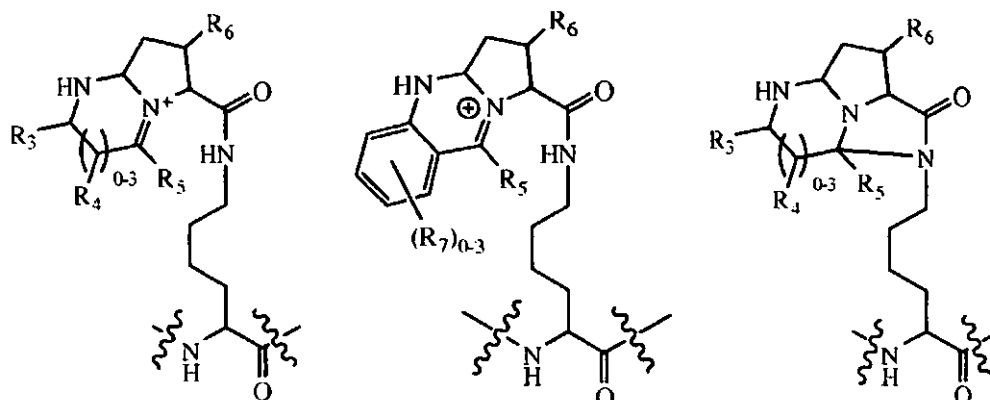
5 donde al menos un BB es un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (C-1) o la Fórmula (D-1) o la Fórmula (E-1) o la Fórmula (F-1) o la Fórmula (G-1) o la Fórmula (H-1) o la Fórmula (I-1) o la Fórmula (J-1) o la Fórmula (K-1) o la Fórmula (L-1).

En determinadas realizaciones de los métodos anteriormente mencionados y dichas proteínas derivatizadas, el anillo A se selecciona entre fenilo, naftalilo y piridilo.

10 En determinadas realizaciones de los métodos anteriormente mencionados, cada BB se selecciona de forma independiente entre un resto de aminoácido, un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (A-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (B-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (C-1) un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (D-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (E-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (F-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (G-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (H-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (I-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (J-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (K-1) y un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (L-2);

20

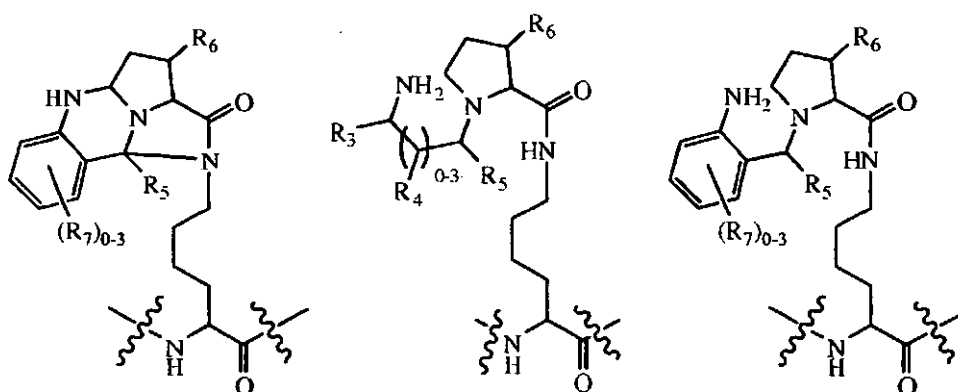




(E-1)

(F-2)

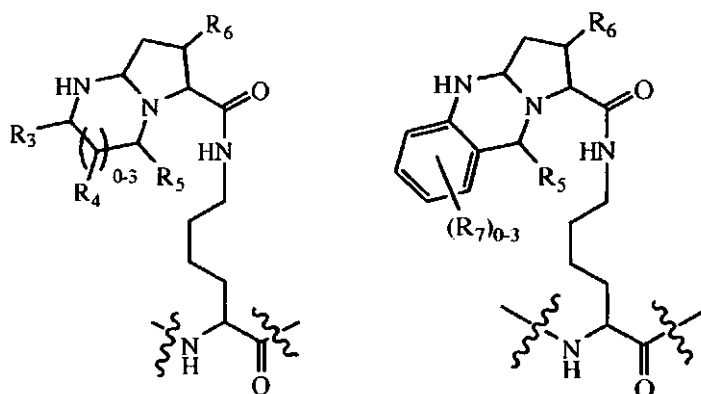
(G-1)



(H-2)

(I-1)

(J-2)



(K-1)

(L-2)

donde,

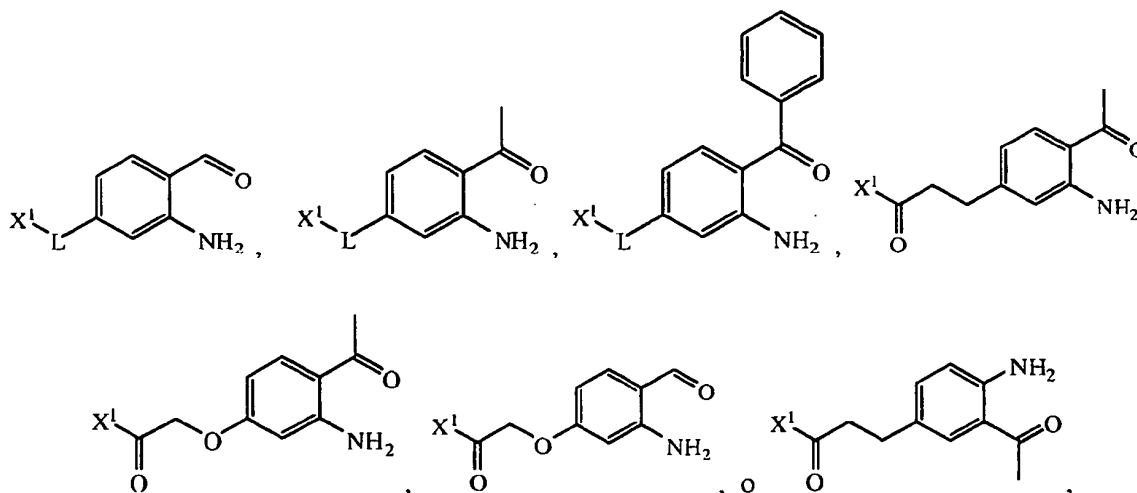
- 5 R_3 , R_5 y cada R_4 se seleccionan independientemente entre H, -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidrox-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹; R_6 es H o alquilo C₁;
- cuando está presente cada R_7 se selecciona independientemente entre -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidrox-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹;
- 10 L se selecciona entre un enlace, alquilenlo C₁₋₈, alquilenlo C₁₋₈ halo-sustituido, alquilenlo C₁₋₈ hidrox-sustituido, alquenilenlo C₂₋₈, alquenilenlo C₂₋₈ halo-sustituido, alquenilenlo C₂₋₈ hidrox-sustituido, un polialquilenglicol, un polietilenglicol), -O(CR¹¹R¹²)_k-, -S(CR¹¹R¹²)_k-, -S(O)_k(CR¹¹R¹²)_k-, -O(CR¹¹R¹²)_k-NR¹¹C(O)-,

$-O(CR^{11}R^{12})_kC(O)NR^{11}$ -, $-C(O)$ -, $-C(O)(CR^{11}R^{12})_k$ -, $-C(S)$ -, $-C(S)(CR^{11}R^{12})_k$ -, $-C(O)NR^{11}$ -, $-NR^{11}C(O)$ -, $-NR^{11}(CR^{11}R^{12})_k$ -, $-CONR^{11}(CR^{11}R^{12})_k$ -, $-N(R^{11})CO(CR^{11}R^{12})_k$ -, $-C(O)NR^{11}(CR^{11}R^{12})_k$ -, $-NR^{11}C(O)(CR^{11}R^{12})_k$ -, en donde cada R^{11} y R^{12} son independientemente H, alquilo C_{1-8} , alquilo C_{1-8} halo-sustituido, o alquilo C_{1-8} hidroxi-sustituido, y k es un número entero de 1 a 12, y

5 X^1 se ha seleccionado de una marca, un colorante, un polímero, un polímero soluble en agua, un polialquilenglicol, un poli(etilenglicol), un derivado de poli(etilenglicol), un azúcar, un lípido, un fotorreticulador, un compuesto citotóxico, un fármaco, una marca de afinidad, una marca de fotoafinidad, un compuesto reactivo; una resina, un péptido, una segunda proteína o polipéptido o análogo de polipéptido, un anticuerpo o fragmento de anticuerpo, un quelante metálico, un cofactor, un ácido graso, un hidrato de carbono, un polinucleótido, un ADN, un ARN, una sonda de la PCR, un polinucleótido antisentido, un ribooligonucleótido, un desoxirribooligonucleótido, ADN modificado con fosforotioato, ADN y ARN modificados, ácido nucleico peptídico, un sacárido, un disacárido, un oligosacárido, un polisacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial, una nanopartícula, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un grupo funcional novedoso, un grupo que interactúa de forma covalente o no covalente con otras moléculas, un resto fotoestimulado, un resto excitable mediante radiación actínica, un ligando un resto fotoisomerizable, biotina, un análogo de biotina, un resto que incorpora un átomo pesado, un grupo escindible químicamente, un grupo fotoescindible, una cadena lateral alargada, un azúcar unido a carbono, un agente redox activo, un aminotioácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo cromóforo, un grupo quimioluminiscente, un resto fluorescente, un grupo denso en electrones, un grupo magnético, un grupo intercalante, un grupo quelante, un cromóforo, un agente de transferencia de energía, un agente biológicamente activo, una marca detectable, una molécula pequeña, un ácido nucleico inhibidor, un ARNip, un radionucleótido, un agente de captura de neutrones, un derivado de biotina, punto(s) cuántico(s), un nanotransmisor, un radiotransmisor, una abzima, una enzima, un activador complejo activado, un virus, una toxina, un adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, un epítipo de linfocito T, un fosfolípido, una molécula de tipo LPS, hemocianina de lapa californiana (KLH), un hapteno inmunógeno, un aglicano, un alérgeno, una angioestatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimotopo, un receptor, una micela inversa, detergentes, potenciadores de la respuesta inmune, colorantes fluorescentes, reactivos FRET, sondas de radiografía, otra sonda de espectroscopía, profármacos, toxinas para inmunoterapia, un soporte sólido, $-CH_2CH_2-(OCH_2CH_2O)_p-OX^2$, $-O-(CH_2CH_2O)_pCH_2CH_2-X^2$, y cualquier combinación de los mismos, donde p es de 1 a 10.000 y X^2 es H, un alquilo C_{1-8} , un grupo protector o un grupo funcional terminal.

En determinadas realizaciones de los métodos anteriormente mencionados y dichas proteínas derivatizadas, R_5 es $-LX^1$. En determinadas realizaciones de los métodos anteriormente mencionados y dichas proteínas derivatizadas, X^1 es un azúcar, un polietilenglicol, un resto fluorescente, un inmunomodulador, un ácido ribonucleico, un ácido desoxirribonucleico, una proteína, un péptido, una biotina, un fosfolípido, un agonista de TLR7, un hapteno inmunógeno o un soporte sólido. En determinadas realizaciones, L es poli(alquilenglicol), un poli(etilenglicol), alquilenos C_{1-8} , alquilenos C_{1-8} halo-sustituido o alquilenos C_{1-8} hidroxi-sustituido

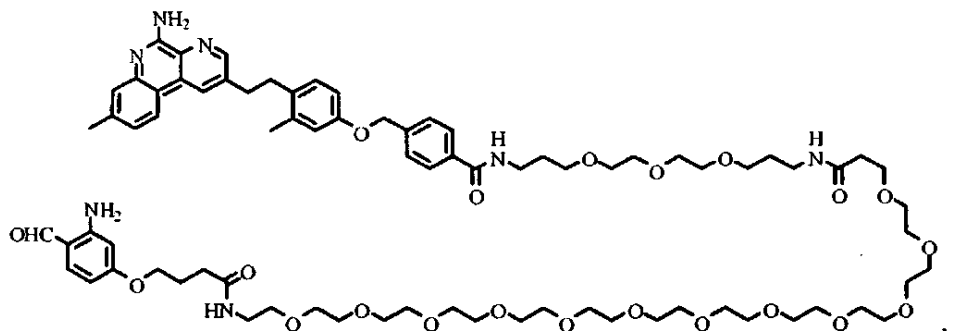
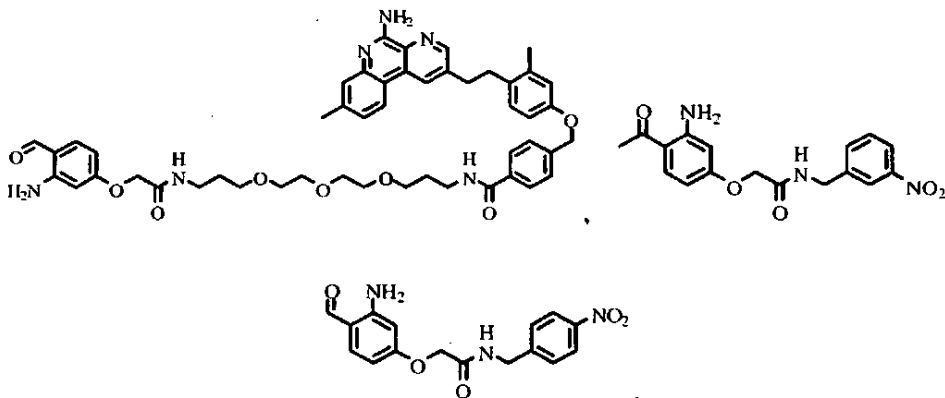
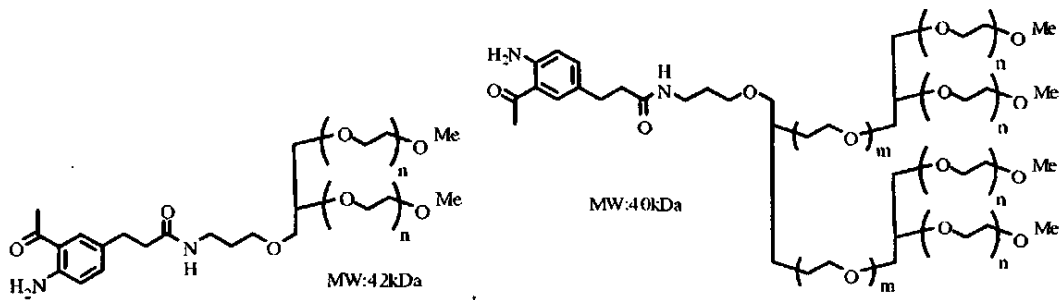
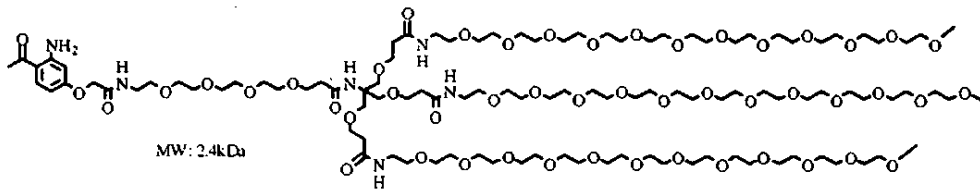
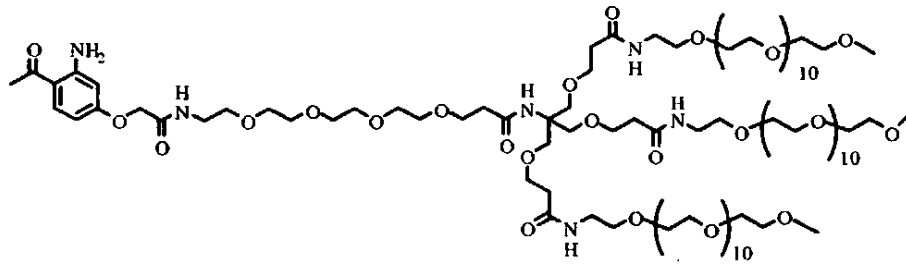
En determinadas realizaciones de los métodos anteriormente mencionados, el compuesto de Fórmula (IV) es

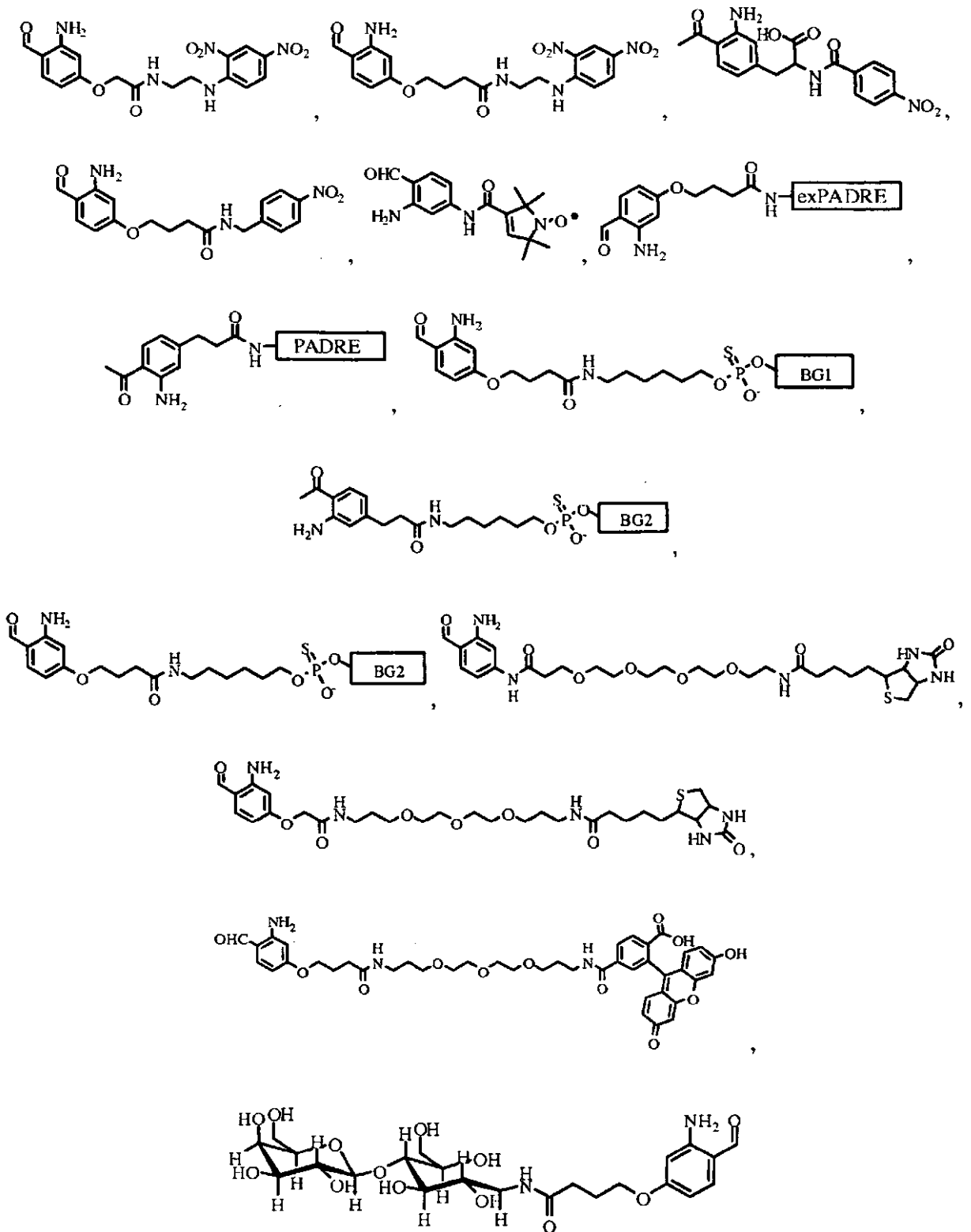


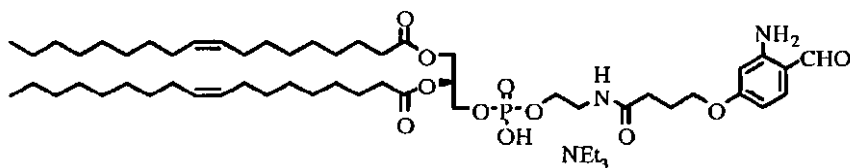
45 donde L y X^1 son como se describe en el presente documento,

En determinadas realizaciones de dichos reactivos, L es un enlace y X^1 es un polietilenglicol.

En determinadas realizaciones del reactivo de Fórmula (IV) es

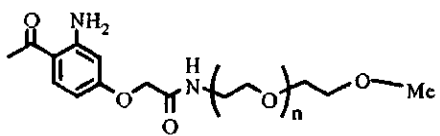




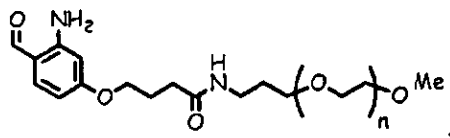


donde los compuestos que tienen uno o más restos de polietilenglicol (PEG) tienen un peso molecular promedio en el intervalo de 1000 Da a 50 kDa, y n es de 20 a 1200 y donde exPADRE es AlaGlySerArgSerGly(DAla)LysCha-ValAlaAlaTrpThrLeuLysAla(D-Ala)Gly-OH (SEC ID N°: 28), PADRE es Gly(DAla)LysChaValAlaAlaTrpThrLeuLysAla(D-Ala)Gly-OH (SEC ID N°: 29), BGI es 5'*t*C*A*T*G*A*C*G*T*T*C*T*G*A*C*G*T-T-3' (SEC ID N°: 30) y BG2 es 5*T*C*G*T*C*G*T*T*T*C*G*G*C*G*C*G*C*G*C*G-3' (SEC ID N°: 31), y donde * denota un enlace fosfotioato.

10 En determinadas realizaciones del reactivo de Fórmula (IV) es un compuesto que tiene la siguiente estructura:



15 donde el compuesto tiene un peso molecular promedio en el intervalo de 1000 Da a 30 kDa, y n es un número entero de 20 a 679. En otra realización, el reactivo de Fórmula (IV) es un compuesto que tiene la siguiente estructura:



20 donde el compuesto tiene un peso molecular promedio en el intervalo de 1000 Da a 45 kDa, y n es un número entero de 20 a 1018.

Descripción detallada de la invención

25 Se proporcionan en el presente documento métodos y composiciones usados para modificar proteínas, polipéptidos y/o péptidos de una forma específica del sitio, donde dichos métodos implican la modificación específica de sitio de restos de pirrolisina o pirrolina-carboxi-lisina (PCL) codificados genéticamente donde la pirrolisina y los aminoácidos PCL se han generado biosintéticamente. Se proporcionan en el presente documento diversos tipos de moléculas que están acopladas a proteínas, polipéptidos y/o péptidos que tienen uno o más restos PCL o pirrolisina incorporadas biosintéticamente a la anterior de una manera específica del sitio. En determinadas realizaciones, dichas modificaciones específicas de sitio se utilizan para marcar proteínas, polipéptidos y/o péptidos específicos de sitios. En determinadas realizaciones, la marca es un resto fluorescente, un resto fosforescente, un resto quimioluminiscente, un resto quelante, un resto intercalante, un resto radioactivo, un resto cromóforo, un resto radioactivo, un resto marcado con espín, un resto activo para RMN, unos reactivos para formación de imágenes IRM. En determinadas realizaciones, dichas modificaciones específicas de sitio se usan para unir inmunomoduladores a proteínas, polipéptidos y/o péptidos, En otras realizaciones, dichas modificaciones específicas de sitio se usan para unir poli(etilenglicol) (PEG) a proteínas, polipéptidos y/o péptidos, En otras realizaciones, dichas modificaciones específicas de sitio se usan para unir azúcares (glicosilados) a proteínas, polipéptidos y/o péptidos,

40 En otras realizaciones, dichas modificaciones específicas de sitio se utilizan para reticular de una forma específica del sitio proteínas, polipéptidos y/o péptidos formando por tanto hetero-oligómeros que incluyen, pero sin limitación, heterodímeros y heterotrímeros. En determinadas realizaciones, dichas modificaciones específicas de sitio se utilizan para reticular de una forma específica del sitio anticuerpos con proteínas, polipéptidos y/o péptidos, En otras realizaciones, dichas modificaciones específicas de sitio se utilizan para reticular de una forma específica del sitio proteínas, polipéptidos y/o péptidos formando por tanto conjugados de proteína-proteína, conjugados de proteína-polipéptido, conjugados de proteína-péptido, conjugados de polipéptido-polipéptido, conjugados de polipéptido-péptido o conjugados de péptido-péptido.

50 En otras realizaciones, dichas modificaciones específicas de sitio se utilizan para unir de una forma específica del sitio anticuerpos con proteínas, en el que la proteína es una proteína tóxica proporcionada en el presente documento, formando por tanto conjugados de anticuerpo-fármaco. En otras realizaciones, dichas modificaciones específicas de

sitio se utilizan para unir de una forma específica del sitio anticuerpos con proteínas, en las que los anticuerpos se acoplan a fármacos de bajo peso molecular, formando por tanto conjugados de anticuerpo-fármaco.

En otras realizaciones, dichas modificaciones específicas de sitio se utilizan para unir de una forma específica del sitio receptores-ligando a proteínas, en el que la proteína es una proteína tóxica proporcionada en el presente documento, formando por tanto conjugados de receptor-ligando-fármaco. En otras realizaciones, dichas modificaciones específicas de sitio se utilizan para unir de una forma específica del sitio receptores-ligandos con proteínas, en las que el receptor-ligando se acopla a fármacos de bajo peso molecular, formando por tanto conjugados de receptor-ligando-fármaco.

Se proporcionan en el presente documento proteínas, polipéptidos y/o péptidos que tienen pirrolina y/o PCL incorporados en el anterior usando los métodos proporcionados. Dichas proteínas incluyen, pero no se limitan a, eritropoyetina (EPO), factor de crecimiento de fibroblastos 21 (FGF21), interferón alfa (INF-a), interleuquina 2 (IL-2), interleuquina 4 (IL-4), interleuquina 6 (IL-6), interleuquina 10 (IL-10), interleuquina 17 (IL-17), factor de crecimiento de tipo insulina 1 (IGF1), e interferón beta (INF-p).

Se proporcionan en el presente documento proteínas, polipéptidos y/o péptidos que tienen pirrolisina y/o PCL incorporadas en los anteriores utilizando los métodos proporcionados en el presente documento y derivatizados adicionalmente utilizando los métodos proporcionados en el presente documento. Dicha derivatización incluye, pero sin limitación, PEGilación. Dichas proteínas incluyen, pero no se limitan a, eritropoyetina (EPO), factor de crecimiento de fibroblastos 21 (FGF21), interferón alfa (INF-a), interleuquina 2 (IL-2), interleuquina 4 (IL-4), interleuquina 6 (IL-6), interleuquina 10 (IL-10), interleuquina 17 (IL-17), factor de crecimiento de tipo insulina 1 (IGF1), e interferón beta (INF-p).

Se proporcionan adicionalmente en el presente documento proteínas, polipéptidos y/o péptidos que tienen pirrolisina y/o PCL incorporados en los anteriores, donde dichas proteínas, polipéptidos y/o péptidos se reticulan usando los métodos proporcionados en el presente documento. Dichas proteínas incluyen, pero no se limitan a, eritropoyetina (EPO), factor de crecimiento de fibroblastos 21 (FGF21), interferón alfa (INF-a), e interferón beta (INF-p).

En otras realizaciones, dichas modificaciones específicas de sitio se utilizan para producir proteínas, polipéptidos y/o péptidos donde la posición de la pirrolisina o el PCL incorporados específicamente en el sitio permite la orientación controlada y la unión de dichas proteínas, polipéptidos y/o péptidos a la superficie de un soporte sólido. En determinadas realizaciones, dicho soporte sólido es una placa de microvaloración de plástico, un porta de vidrio, una superficie de sílice, una perla de polímero, unas partículas de oro o unas nanopartículas tanto revestidas como sin revestir. En determinadas realizaciones, dicha orientación y unión controladas se usan en el análisis de proteínas, polipéptidos y péptidos mediante ELISA u otros anticuerpos de ensayo. En determinadas realizaciones, dicha orientación y unión controladas se usan para purificar y/o identificar ligandos de las proteínas, polipéptidos y/o péptidos inmovilizados. En determinadas realizaciones, dicha orientación y unión controladas se usan en el análisis de proteínas, polipéptidos y péptidos y sus interacciones mediante análisis de ondas evanescentes, incluyendo, pero sin limitarse a análisis exento de marcas utilizando análisis de resonancia de plasmón superficial. En determinadas realizaciones, dicha orientación y unión controladas sobre superficies se usan en el análisis de proteínas, polipéptidos y péptidos y sus interacciones utilizando microequilibrios, (eléctrico óptico y/o mecánico), espectroscopía infrarroja, espectroscopía Raman incluyendo espectroscopía Raman con potenciación superficial, resonancia por evanescencia, fluorescencia, interferometría, espectrometría de masas y otros métodos de espectroscopía. Dichos métodos de análisis se usan para investigar interacciones de las proteínas, polipéptidos y/o péptidos inmovilizados con otras proteínas, polipéptidos, péptidos, ácidos nucleicos, ADN, ARN, moléculas pequeñas, fármacos, metabolitos, azúcares, hidratos de carbono; oligosacáridos, polisacáridos y/u otras moléculas que incluyen cambios conformacionales inducidos por dichas interacciones. Dichos métodos de análisis se usan también para investigar interacciones de las proteínas, polipéptidos y/o péptidos inmovilizados con múltiples complejos de proteínas subunitarias, estudiar moléculas recombinantes naturales recombinantes, sintéticas o etiquetadas, descubrir nuevas moléculas de interacción en fluidos corporales, sobrenadantes de cultivos celulares o extractos brutos, estudiar la interacción de moléculas pequeñas, tales como fármacos candidatos, con sus dianas, estudiar la bioquímica de membranas o las interacciones de receptores unidos a membranas utilizando membranas naturales, membranas o vesículas artificiales, investigar la replicación, transcripción y traducción determinar las relaciones moleculares durante la formación de los complejos de proteínas y su interacción con el ADN, estudiar la hibridación del ADN y el ARN estudiar las interacciones que implican células o virus completos, estudiar los efectos de la glicosilación sobre las interacciones moleculares, y determinar las propiedades de reconocimiento específicas de los hidratos de carbono de la superficie celular.

En otras realizaciones, dichas modificaciones específicas de sitio se utilizan para unir de una forma específica del sitio ácidos nucleicos con proteínas. En determinadas realizaciones, dichas modificaciones específicas de sitio se utilizan para unir de una forma específica del sitio ácidos nucleicos con anticuerpos o fragmentos de anticuerpo. En determinadas realizaciones, el ácido nucleico unido a la proteína o anticuerpo se utiliza para inmovilizar la proteína o el anticuerpo en localizaciones definidas sobre una matriz de ADN mediante hibridación. En determinadas realizaciones, el ácido nucleico unido a la proteína o anticuerpo se utiliza para detectar la unión de la proteína o el anticuerpo mediante la PCR, amplificación por desplazamiento de hebra (SDA), reacción en cadena de la ligasa (LCR), inmuno-PCR por ligadura de proximidad, amplificación en círculo rodante, amplificación mediada por transcripción,

amplificación de la señal de la tiramida de NEN u otros métodos de amplificación de la señal. en determinadas realizaciones, dicha unión específica de sitio de una sonda de la PCR con un anticuerpo se utiliza para generar un reactivo para reacciones de inmuno-PCR (Véase, M. Adler, R. Wacker, Ch. M. Niemeyer, Sensitivity by combination: Immuno-PCR and related technologies, Analyst, 2008, 133, 702-718). En determinadas realizaciones, el ácido nucleico unido a la proteína o al anticuerpo permite la inmuno-PCR u otros inmunoensayos de muchos analitos en paralelo (inmuno-PCR multiplexada o inmunoensayos multiplexados).

En determinadas realizaciones, el ácido nucleico unido a la proteína o al anticuerpo media en la formación de homo y heterodímeros.

Definiciones

El término "alquilo", tal como se usa en el presente documento, se refiere a hidrocarburos saturados de cadena lineal o ramificada. Como se usa en el presente documento, los términos "alquilo C₁-C₃", "alquilo C₁-C₄", "alquilo C₁-C₉", "alquilo C₁-C₆", "alquilo C₁-C₇" y "alquilo C₁-C₈" se refieren a un grupo alquilo que contiene al menos 1, y como máximo 3, 4, 5, 6, 7 u 8 átomos de carbono, respectivamente. Los ejemplos no limitantes de grupos alquilo que se usan en el presente documento incluyen metilo, etilo, n-propilo, isopropilo, n-butilo, isobutilo, sec-butilo, t-butilo, n-pentilo, isopentilo, hexilo, heptilo, octilo, nonilo, decilo y similares.

El término "alquilenilo", tal como se usa en el presente documento, se refiere a un radical hidrocarburo divalente saturado de cadena lineal o ramificada, donde el radical se deriva mediante la eliminación de un átomo de hidrógeno de cada dos átomos de carbono. Como se usa en el presente documento, los términos "alquilenilo C₁-C₃", "alquilenilo C₁-C₄", "alquilenilo C₁-C₉", y "alquilenilo C₁-C₆" se refieren a un grupo alquilenilo que contiene al menos 1, y como máximo 3, 4, 5 o 6 átomos de carbono respectivamente. Los ejemplos no limitantes de grupos alquilenilo que se usan en el presente documento incluyen, metileno, etileno, n-propileno, isopropileno, n-butileno, isobutileno, sec-butileno, t-butileno, n-pentileno, isopentileno, hexileno y similares.

El término "alcoxi", tal como se usa en el presente documento, se refiere al grupo -OR_a, donde R_a es un grupo alquilo tal como se define en el presente documento. Como se usa en el presente documento, los términos "alcoxi C₁-C₃", "alcoxi C₁-C₄", "alcoxi C₁-C₉", "alcoxi C₁-C₆", "alcoxi C₁-C₇" y "alcoxi C₁-C₈" se refieren a un grupo alcoxi donde el resto alquilo contiene al menos 1, y como máximo 3, 4, 5, 6, 7 o 8, átomos de carbono. Los ejemplos no limitantes de grupos alcoxi, tal como se usa en el presente documento, incluyen metoxi, etoxi, n-propoxi, isopropoxi, n-butoxi, t-butoxi pentoxi, hexoxi, heptoxi, y similares.

El término "grupo de modificación del amino terminal" tal como se usa en el presente documento, se refiere a cualquier molécula que forma un enlace con un grupo amina terminal. A modo de ejemplo, dichos grupos amina terminales incluyen, pero no se limitan a, grupos protectores de amina, el extremo de moléculas poliméricas, donde dichas moléculas poliméricas incluyen, pero no se limitan a, polipéptidos, polinucleótidos, y polisacáridos. Los grupos de modificación del amino terminal incluyen también, pero no se limitan a, diversos polímeros, péptidos o proteínas solubles en agua. Por medio solo de ejemplo, los grupos de modificación del extremo terminal incluyen polietilenglicol o albúmina de suero. Determinados grupos de modificación del amino terminal se usan para modificar características terapéuticas de las proteínas, que incluyen, pero no se limitan a semivida en suero.

El término "arilo", tal como se usa en el presente documento, se refiere a sistemas de anillos monocíclicos, bicíclicos, y tricíclicos que tienen un total de cinco a catorce miembros en el anillo, donde al menos un anillo del sistema es aromático y donde cada anillo del sistema contiene 3 a 7 miembros. Un grupo arilo está "opcionalmente sustituido", cuando dicho grupo arilo contiene uno o más sustituyentes. Salvo que se defina de otra forma en el presente documento, los sustituyentes adecuados se seleccionan generalmente entre halógeno, -R, -OR, -SR, -NO₂, -CN, -N(R)₂, -NRC(O)R, -NRC(S)R, -NRC(O)N(R)₂, -N RC(S)N(R)₂, -NRCO₂R, -NRNRC(O)R, -NRNRC(O)N(R)₂, -NRNRCO₂R, -C(O)C(O)R, -C(O)CH₂C(O)R, -CO₂R, -C(O)R, -C(S)R, -C(O)N(R)₂, -C(S)N(R)₂, -OC(O)N(R)₂, -OC(O)R, -C(O)N(OR)R, -C(NOR)R, -S(O)₂R, -S(O)₃R, -SO₂N(R)₂, -S(O)R, -NRSO₂N(R)₂, -NRSO₂R, -N(OR)R, -C(=NH)-N(R)₂, -P(O)₂R, -PO(R)₂, -OPO(R)₂, -(CH₂)₀₋₂NHC(O)R, fenilo (Ph) opcionalmente sustituido con R, -O(Ph) opcionalmente sustituido con R, -(CH₂)₁₋₂(Ph) opcionalmente sustituido con R, o -CH=CH(Ph), opcionalmente sustituido con R, donde cada ocurrencia independiente de R se selecciona entre hidrógeno, alquilo C₁-C₆ opcionalmente sustituido, alcoxi C₁-C₆ opcionalmente sustituido, o heteroarilo no sustituido de 5-6 miembros. fenilo, -O(Ph), o -CH₂(Ph), o dos ocurrencias independientes de R, en el mismo sustituyente o diferentes sustituyentes, tomados en conjunto con el(los) átomo(s) a los cuales se une cada R, para formar un anillo bicíclico saturado sustituido de 3-12 miembros, parcialmente insaturado, o monocíclico completamente insaturado que tiene 0-4 heteroátomos seleccionados de forma independiente entre nitrógeno, oxígeno, o azufre. Los ejemplos no limitantes de grupos arilo, tal como se usa en el presente documento, incluyen fenilo, naftilo, fluorenilo, indenilo, azuleno, antraceno, fenantraceno y similares.

El término "arileno", tal como se usa significa un radical divalente derivado de un grupo arilo.

Un "enlazador bifuncional", tal como se usa en el presente documento, se denomina también como "polímero bifuncional" se refiere a un enlazador que comprende dos grupos funcionales que son capaces de reaccionar

específicamente con otros restos para formar enlaces covalentes o no covalentes. Dichos restos incluyen, pero no se limitan a, los grupos de la cadena lateral de los aminoácidos. Por medio solo de ejemplo, un enlazador bifuncional tiene un grupo funcional reactivo con un grupo en un primer péptido, y otro grupo funcional que es reactivo con un grupo en un segundo péptido, formando por tanto un conjugado que incluye el primer péptido, el enlazador bifuncional y el segundo péptido. Un enlazador bifuncional es de cualquier longitud o peso molecular deseado, y se selecciona para proporcionar una separación o conformación deseada concreta.

Un "enlazador multifuncional", tal como se usa en el presente documento, se denomina también "polímero multifuncional", se refiere a un enlazador que comprende dos o más grupos funcionales que son capaces de reaccionar con otros restos para formar enlaces covalentes o no covalentes. Dichos restos incluyen, pero no se limitan a, los grupos de la cadena lateral de los aminoácidos. Un enlazador multifuncional tiene cualquier longitud o peso molecular deseado, y se selecciona para proporcionar una separación o conformación deseada concreta.

El término "ciano", tal como se usa en el presente documento, se refiere a un grupo -CN.

El término "cicloalquilo", tal como se usa en el presente documento, se refiere a un anillo saturado o parcialmente insaturado que puede ser monocíclico, bicíclico fusionado, o tricíclico fusionado, o un conjunto de anillos policíclicos con un puente. Como se usa en el presente documento, los términos "cicloalquilo C₃-C₅", "cicloalquilo C₃-C₆", "cicloalquilo C₃-C₇", "cicloalquilo C₃-C₈", "cicloalquilo C₃-C₉" y "cicloalquilo C₃-C₁₀" se refieren a un grupo cicloalquilo donde el montaje de anillo policíclico con puente saturado o parcialmente insaturado, monocíclico, bicíclico fusionado contiene al menos 3, y como máximo 5, 6, 7, 8, 9 o 10, átomos de carbono. Los ejemplos no limitantes de grupos cicloalquilo, tal como se usa en el presente documento, incluyen ciclopropilo, ciclobutilo, ciclopentilo, ciclohexilo, cicloheptilo, ciclooctilo, ciclononilo, ciclodecilo, ciclopentenilo, ciclohexenilo, decahidronaftalenilo, 2,3,4,5,6,7-hexahidro-1H-indenilo y similares.

El término "ciclodextrina", tal como se usa en el presente documento, se refiere a hidratos de carbono cíclicos que consisten en al menos seis a ocho moléculas de glucosa en una formación de anillo. La parte externa del anillo contiene grupos solubles en agua; en el centro del anillo está una cavidad relativamente no polar capaz de acomodar moléculas pequeñas.

El término "halógeno", tal como se usa en el presente documento, se refiere a flúor (F), cloro (Cl), bromo (Br), o yodo (I).

El término "halo", tal como se usa en el presente documento, se refiere a radicales de halógeno: flúor (-F), cloro (-Cl), bromo (-Br), y yodo (-I).

El término "haloacilo", tal como se usa en el presente documento, se refiere a grupos acilo que contienen restos de halógeno, incluyendo, pero sin limitación, -C(O)CH₃, -C(O)CF₃, -C(O)CH₂OCH₃, y similares.

Los términos "haloalquilo" o "alquilo halo-sustituido" tal como se usa en el presente documento, se refiere a un grupo alquilo tal como se define en el presente documento, sustituido con al menos un grupo halo o sus combinaciones. Los ejemplos no limitantes de dichos grupos haloalquilo de cadena lineal o ramificada, tal como se usa en el presente documento, incluyen metilo, etilo, propilo, isopropilo, isobutilo y n-butilo sustituidos con uno o más grupos halo o sus combinaciones, incluyendo, pero sin limitación, trifluorometilo, pentafluoroetilo, y similares.

El término "haloalcoxi", tal como se usa en el presente documento, se refiere a un grupo alcoxi tal como se ha definido anteriormente, sustituido con uno o más grupos halo o sus combinaciones. Los ejemplos no limitantes de dichos grupos haloalquilo de cadena lineal o ramificada, tal como se usa en el presente documento, incluyen metoxi, etoxi, n-propoxi, isopropoxi, n-butoxi, t-butoxi pentoxi, hexoxi, heptoxi y similares sustituidos con uno o más grupos o sus combinaciones,

El término "heteroalquilo", tal como se usa en el presente documento, se refiere a un grupo alquilo tal como se define en el presente documento donde uno o más átomos de carbono se sustituyen de forma independiente por uno o más de oxígeno, azufre, nitrógeno, o una de sus combinaciones.

El término "heteroalquileno", tal como se usa en el presente documento, se refiere a un radical divalente derivado de un heteroalquilo,

El término "heteroarilo", tal como se usa en el presente documento, se refiere a sistemas de anillos monocíclicos, bicíclicos, y tricíclicos que tienen un total de cinco a catorce miembros en el anillo, donde al menos un anillo del sistema es aromático, al menos un anillo del sistema contiene uno o más heteroátomos seleccionados entre nitrógeno, oxígeno y azufre, y donde cada anillo del sistema contiene 3 a 7 miembros en el anillo. A no ser que se defina otra cosa anteriormente y en el presente documento, los sustituyentes adecuados en el átomo de carbono no saturado de un grupo heteroarilo se seleccionan generalmente de entre halógeno -R, -OR, -SR, -NO₂, -CN, -N(R)₂, -NRC(O)R, -NRC(S)R, -NRC(O)N(R)₂, -N RC(S)N(R)₂, -NRCO₂R, -NRNRC(O)R, -NRNRC(O)N(R)₂, -NRNRCO₂R, -C(O)C(O)R, -C(O)CH₂C(O)R, -CO₂R, C(O)R, -C(S)R, -C(O)N(R)₂, -C(S)N(R)₂, -OC(O)N(R)₂, -OC(O)R, -C(O)N(OR)R, -C(NOR)R,

-S(O)₂R, -S(O)₃R, -SO₂N(R)₂, -S(O)R, -NRSO₂N(R)₂, -NRSO₂R, -N(OR)R, -C(=NH)-N(R)₂, -P(O)₂R, -PO(R)₂, -OPO(R)₂, -(CH₂)₀₋₂NHC(O)R, fenilo (Ph) opcionalmente sustituido con R, -O(Ph) opcionalmente sustituido con R, -(CH₂)₁₋₂(Ph), opcionalmente sustituido con R, o -CH=CH(Ph), opcionalmente sustituido con R, donde cada ocurrencia independiente de R se selecciona entre hidrógeno, alquilo C₁-C₉ opcionalmente sustituido, alcoxi C₁-C₉ opcionalmente sustituido, o heteroarilo no sustituido de 5-6 miembros. fenilo, -O(Ph), o -CH₂(Ph), o dos ocurrencias independientes de R, en el mismo sustituyente o diferentes sustituyentes, tomados en conjunto con el(los) átomo(s) a los cuales se une cada R, para formar un anillo bicíclico saturado sustituido de 3-12 miembros, parcialmente insaturado, o monocíclico completamente insaturado que tiene 0-4 heteroátomos seleccionados de forma independiente entre nitrógeno, oxígeno, o azufre. Los ejemplos no limitantes de grupos heteroarilo, tal como se usa en el presente documento, incluyen benzofuranilo, benzofurazanilo, benzoxazolilo, benzopiranoilo, benzotiazolilo, benzotienilo, benzazepinilo, benzimidazolilo, benzotiopiranoilo, benzo[1,3]dioxol, benzo[b]furilo, benzo[b]tienilo, cinolinilo, furazanilo, furilo, furopiridinilo, imidazolilo, indolilo, indolizinoilo, indolin-2-ona, indazolilo, isoindolilo, isoquinolinilo, isoxazolilo, isotiazolilo, 1,8-naftiridinilo, oxazolilo, oxaindolilo, oxadiazolilo, pirazolilo, pirrolilo, ftalazinilo, pteridinilo, purinilo, piridilo, piridazinilo, pirazinilo, pirimidilo, pirimidinilo, quinoxalinilo, quinolinilo, quinazolinilo, 4H-quinolizinoilo, tiazolilo, tiadiazolilo, tienilo, triacinoilo, triazolilo y tetrazolilo.

El término "heterocicloalquilo", tal como se usa en el presente documento, se refiere a un cicloalquilo, tal como se ha definido en el presente documento, donde uno o más átomos de carbono del anillo se han sustituido por un resto seleccionado de -O-, -N=, -NR-, -C(O)-, -S-, -S(O)- o -S(O)₂-, donde R es hidrógeno, alquilo C₁-C₄ o un grupo protector de nitrógeno, con la condición de que el anillo de dicho grupo no contiene dos átomos de O o S adyacentes. Los ejemplos no limitantes de grupos heterocicloalquilo, tal como se usa en el presente documento, incluyen morfolino, pirrolidinilo, pyrrolidinil-2-ona, piperazinilo, piperidinilo, piperidinilona, 1,4-dioxa-8-aza-espiro[4.5]dec-8-ilo, 2H-pirrolilo, 2-pirrolinilo, 3-pirrolinilo, 1,3-dioxolanilo, 2-imidazolinoilo, imidazolidinilo, 2-pirazolinilo, pirazolidinilo, 1,4-dioxanilo 1,4-ditianoilo, tiomorfolino, azepanilo, hexahidro-1,4-diazepinilo, tetrahidrofuranilo, dihidrofuranilo, tetrahidrotienilo, tetrahidropiranoilo, dihidropiranoilo, tetrahidrotiopiranoilo, tioxanilo, azetidinoilo, oxetanilo, tietanilo, oxepanilo, tiepanilo, 1,2,3,6-tetrahidropiridinilo, 2H-piranoilo, 4H-piranoilo, dioxanilo, 1,3-dioxolanilo, , ditianoilo, ditioalanilo, dihidropiranoilo, dihidrotienilo, dihidrofuranilo, imidazolinoilo, imidazolidinilo, 3-azabicyclo[3.1.0]hexanilo, y 3-azabicyclo[4.1.0]heptanilo.

El término "heteroátomo", tal como se usa en el presente documento, se refiere a uno o más de oxígeno, azufre, nitrógeno, fósforo, o silicio.

El término "hidroxilo", tal como se usa en el presente documento, se refiere al grupo -OH.

El término "hidroxialquilo", tal como se usa en el presente documento se refiere a un grupo alquilo tal como se define en el presente documento sustituido con al menos un hidroxilo, siendo el hidroxilo tal como se define en el presente documento. Los ejemplos no limitantes de grupos "hidroxialquilo C₁-C₆" tal como se usan en el presente documento incluyen metilo, etilo, propilo, isopropilo, isobutilo y n-butilo sustituidos de forma independiente con uno o más grupos hidroxilo,

El término "opcionalmente sustituido", tal como se usa en el presente documento, significa que el grupo referenciado puede estar sustituido o no con uno o más grupo(s) adicionales seleccionados de forma individual e independiente a partir de alquilo, alquenoilo, alquinoilo, cicloalquilo, arilo, heteroarilo, heterocicloalquilo, hidroxilo, alcoxi, mercaptilo, ciano, halo, carbonilo, tiocarbonilo, isocianato, tiocianato, isotiocianato, nitro, perhaloalquilo, perfluoroalquilo, y amino, incluyendo grupos amino monosustituidos y disustituidos, y los derivados protegidos de los mismos. Los ejemplos no limitantes de sustituyentes opcionales incluyen, halo, -CN, -, -OR, -C(O)R, -OC(O)R, -C(O)OH, -OC(O)NHR, -C(O)N(R)₂, -SR-, -S(=O)R, -S(=O)₂R, -NHR, -N(R)₂, -NHC(O)-, NHC(O)O-, -C(O)NH-, S(=O)₂NHR, -S(O)₂N(R)₂, -NHS(=O)₂, -NHS(O)₂R, alquilo C₁-C₆, alcoxi C₁-C₆, arilo, heteroarilo, cicloalquilo, heterocicloalquilo, alquilo C₁-C₆ halo-sustituido, alcoxi C₁-C₆ halo-sustituido, donde cada R se selecciona de forma independiente entre H, halo, alquilo C₁-C₆, alcoxi C₁-C₆, arilo, heteroarilo, cicloalquilo, heterocicloalquilo, alquilo C₁-C₆ halo-sustituido, alcoxi C₁-C₆ halo-sustituido,

El término "marca de afinidad", tal como se usa en el presente documento, se refiere a una marca que se une de forma reversible o irreversible a otra molécula.

Los términos "codón ámbar", tal como se usan en el presente documento, se refieren a los sitios de incorporación de pirrolisina, PCL y otros análogos de pirrolisina y corresponden a UAG, el triplete de nucleótidos en el ARN mensajero. La secuencia de nucleótidos TAG está codificada en el ADN y se transcribe a UAG en el ARN que se traduce en una proteína. El codón TAG y el UAG se usan de forma indistinta en el presente documento para referirse al sitio de incorporación de la pirrolisina, PCL y otros análogos de pirrolisina.

El término "aminoácido", tal como se usa en el presente documento, se refiere a aminoácidos de origen natural, aminoácidos no naturales, análogos de aminoácidos y aminoácidos miméticos que funcionan de una manera similar a la de los aminoácidos de origen natural, todos en sus estereoisómeros D y L si su estructura permite dichas formas estereoisoméricas. Los aminoácidos se denominan en el presente documento por cualquiera de sus nombres, sus símbolos de tres letras comúnmente conocidos o por los símbolos de una letra recomendados por la Comisión de Nomenclatura Bioquímica de la IUPAC-IUB. Los aminoácidos de origen natural son aquellos aminoácidos que están

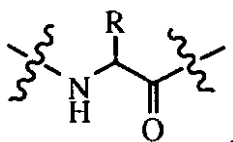
codificados por el código genético, así como aquellos aminoácidos codificados que se han modificado posteriormente. los aminoácidos de origen natural incluyen, pero no se limitan a, alanina (Ala), arginina (Arg), asparagina (Asn), ácido aspártico (Asp), cisteína (Cys), glutamina (Gln), ácido glutámico (Glu), glicina (Gly), histidina (His), isoleucina (Ile), leucina (Leu), Lisina (Lys), metionina (Met), fenilalanina (Phe), prolina (Pro), serina (Ser), treonina (Thr), triptófano (Trp), tirosina (Tyr), valina (Val), pirrolisina (Pyl), selenocisteína (Sec) y pirrolina-carboxi-lisina (PCL). Los aminoácidos codificados modificados incluyen, pero no se limitan a, hidroxiprolina, y-carboxiglutamato, O-fosfoserina, ácido azetidincarboxílico, ácido 2-aminoadípico, ácido 3-aminoadípico, beta-alanina, ácido aminopropiónico, ácido 2-aminobutírico, ácido 4-aminobutírico, ácido 6-aminocaproico, ácido 2-aminoheptanoico, ácido 2-aminoisobutírico, ácido 3-aminoisobutírico, ácido 2-aminopimélico, butilglicina terciaria, ácido 2,4-diaminoisobutírico, desmosina, ácido 2,2'-diaminopimélico, ácido 2,3-diaminopropiónico, N-etilglicina, N-metilglicina, N-etilasparagina, homoprolina, hidroxilisina, alo-hidrolisina, 3-hidroxiprolina, 4-hidroxiprolina, isodesmosina, alo-isoleucina, N-metilalanina, N-metilglicina, N-metilisoleucina N-metilpentilglicina, N-metilvalina, naftalanina, norvalina norleucina, ornitina, pentilglicina, ácido pipercolico y tioprolina. El amino terminalácido incluye también aminoácidos de origen natural que son metabolitos en determinados organismos pero que no están codificados por el código genético para la incorporación en proteínas. Dichos aminoácidos incluyen, pero no se limitan a, ornitina, D-ornitina, y D-arginina.

El término "análogo de aminoácido", tal como se usa en el presente documento, se refiere a compuestos que tienen la misma estructura química básica que un aminoácido que se produce naturalmente, a modo de ejemplo solamente, y un carbono que se une a un hidrógeno, un grupo carboxilo, un grupo amino, y un grupo R. Los análogos de aminoácidos incluyen los aminoácidos naturales y no naturales que están bloqueados químicamente, de forma reversible o irreversible, o su grupo carboxi en el extremo C, su grupo amino en el extremo N y/o sus grupos funcionales de cadena lateral están químicamente modificados. Dichos análogos incluyen, pero no se limitan a, sulfóxido de metionina, metionina sulfona, S-(carboximetil)-cisteína, S-(carboximetil)-cisteína sulfóxido, S-(carboximetil)-cisteína sulfóxido, ácido aspártico-(beta-metil éster), N-etilglicina, alanina carboxamida, homoserina, norleucina, y metionina metil sulfonio. Determinados agentes bloqueantes incluyen, pero no se limitan a, t-butiloxicarbonilo (Boc) y 9-Fluorenilmetiloxicarbonilo (Fmoc).

El término "aminoácidos miméticos", tal como se usa en el presente documento, se refiere a compuestos químicos que tienen una estructura que es diferente de la estructura química general de un aminoácido, pero que funcionan de manera similar a un aminoácido que se produce naturalmente.

El término "aminoácido no natural", tal como se usa en el presente documento, se pretende que represente estructuras de aminoácidos que no se pueden generar biosintéticamente en ningún organismo utilizando genes no modificados o modificados procedentes de cualquier organismo, ya sean iguales o diferentes. Además, se entiende que dichos "aminoácidos no naturales" requieren un ARNt modificado y una ARNt sintetasa modificada (RS) para la incorporación en una proteína. Esta pareja de ARNt/RS ortogonal "seleccionada" es específica del aminoácido no natural y se genera mediante un procedimiento de selección que ha sido desarrollado por Schultz y col., o un procedimiento similar. A modo de ejemplo, pirrolina-carboxi-lisina es un "aminoácido no natural" que está generado biosintéticamente por genes transferidos desde un organismo en las células hospedadoras y se incorpora a proteínas utilizando genes de ARNt y ARNt sintetasa naturales, a la vez que p-aminofenilalanina (véase, Generation of a bacterium with a 21 amino acid genetic code, Mehl RA, Anderson JC, Santoro SW, Wang L, Martin AB, King DS, Horn DM, Schultz PG. J Am Chem Soc. 2003 Ene 29;125(4):935-9) es un "aminoácido no natural" debido a que, aunque está generado biosintéticamente, se incorpora a proteínas mediante una pareja de ARNt sintetasa/ARNt ortogonal.

El término "resto de aminoácido", tal como se usa en el presente documento, se refiere a restos que tienen la estructura:



donde dichos restos se derivan de aminoácidos y el grupo R en la cadena lateral de cualquier aminoácido descrito en el presente documento. dichos restos de aminoácidos incluyen, pero no se limitan a alaninilo, argininilo, asparaginilo, aspartilo, cisteinilo, glutaminilo, glutamilo, glicinilo, histidinilo, isoleucinilo, leucinilo, lisinilo, metioninilo, fenilalaninilo, prolinilo, serinilo, treoninilo, triptofanilo, tirosinilo, valinilo, piroglutamato, formilmetionina, piroglicinilo y selenocisteinilo.

El término "grupo de modificación del extremo amino" se refiere a cualquier molécula que se puede unir a un grupo en el extremo N. Dichos grupos de modificación del extremo amino incluyen, pero no se limitan a, grupos protectores de amina, el extremo de moléculas poliméricas, donde dichas moléculas poliméricas incluyen, pero no se limitan a, polipéptidos, polinucleótidos, y polisacáridos. Los grupos de modificación del extremo incluyen también, pero no se limitan a, diversos polímeros, péptidos o proteínas solubles en agua. Por medio solo de ejemplo, los grupos de modificación del extremo amino incluyen polietilenglicol o albúmina de suero. Determinados grupos de modificación

del extremo amino se usan para modificar características terapéuticas de una proteína, polipéptido o péptido, incluyendo pero sin limitarse a, aumentar la semivida en suero de dichas proteínas, polipéptidos o péptidos.

El término "fragmento de anticuerpo", tal como se usa en el presente documento, se refiere a cualquier forma de un anticuerpo diferente de la que tiene longitud completa. Los fragmentos de anticuerpo en el presente documento incluyen anticuerpos que son componentes más pequeños que existen en los anticuerpos de longitud completa, y los anticuerpos que se han diseñado mediante ingeniería genética. Los fragmentos de anticuerpo incluyen, pero sin limitación, Fv, Fc, Fab, y (Fab')₂, Fv (ScFv) monocatenario, diacuerpos, triacuerpos, tetracuerpos, anticuerpos híbridos bifuncionales, CDR1, CDR2, CDR3, combinaciones de los CDR, regiones variables, regiones marco, regiones constantes, cadenas pesadas, cadenas ligeras, y regiones variables, y moléculas sin anticuerpos de estructura alternativa, anticuerpos biespecíficos, y similares. Otra subestructura funcional es un Fv (ScFv) monocatenario, comprendida por las regiones variables de las cadenas ligera y pesada de la inmunoglobulina, conectadas covalentemente por un enlazador peptídico. Estas proteínas pequeñas (PM < 25.000 Da) retienen generalmente la especificidad y la afinidad por el antígeno en un único polipéptido y pueden proporcionar un bloque de construcción conveniente para moléculas específicas de antígeno más grandes. A no ser que se indique específicamente otra cosa, las declaraciones y reivindicaciones que utilizan el término "anticuerpo" o "anticuerpos" incluyen también "fragmento de anticuerpo" y fragmentos de anticuerpos".

El término "biodisponibilidad", tal como se usa en el presente documento, se refiere a la velocidad y extensión a la cual una sustancia o su resto activo se libera desde su forma farmacéutica y llega a estar disponible en el sitio de la acción o en la circulación general. El aumento en la biodisponibilidad se refiere a aumentar la velocidad y extensión a la cual una sustancia o su resto activo se libera desde una forma farmacéutica y llega a estar disponible en el sitio de la acción o en la circulación general. A modo de ejemplo, un aumento en la biodisponibilidad puede estar indicado como un aumento en la concentración de la sustancia o su resto activo en la sangre cuando se compara con otras sustancias o restos activos.

Los términos "molécula biológicamente activa", "resto biológicamente activo" o "agente biológicamente activo" tal como se usan en el presente documento, se refieren a cualquier sustancia que puede afectar cualquier propiedad física o bioquímica de un sistema, ruta, molécula, o interacción biológica que se refiere a un organismo, incluyendo pero sin limitarse a, virus, bacterias, bacteriófago, transposón, prión, insectos, hongos, plantas, animales, y seres humanos. En particular, tal como se usa en el presente documento, las moléculas biológicamente activas incluyen, pero no se limitan a cualquier sustancia prevista para el diagnóstico, cura, mitigación, el tratamiento, o prevención de una enfermedad en seres humanos u otros animales. o para aumentar de otra manera el bienestar físico o mental de seres humanos o animales. Los ejemplos de moléculas biológicamente activas incluyen, pero no se limitan a, péptidos, proteínas, enzimas, ADN, ARN, pequeñas moléculas de fármacos, fármacos duros, fármacos blandos, polisacáridos, oligosacáridos, disacáridos, hidratos de carbono; átomos o moléculas inorgánicas colorantes, lípidos, nucleósidos, radionucleidos, oligonucleótidos, toxinas; células, virus, liposomas, micropartículas y micelas. Las clases de agentes biológicamente activos que son adecuados para el uso con los métodos y composiciones descritos en el presente documento incluyen, pero no se limitan a, fármacos, profármacos, radionucleidos, agentes de formación de imágenes, polímeros, antibióticos, fungicidas, agentes antiviricos, agentes antiinflamatorios, agentes antitumorales, agentes cardiovasculares, agentes anti ansiedad, hormonas, factores de crecimiento, agentes esteroideos, toxinas derivadas microbianamente, y similares.

El término "actividad biológica moduladora", tal como se usa en el presente documento, se refiere a aumentar o disminuir la concentración de la reactividad de una proteína, polipéptido, péptido, ADN, ARN, sacáridos, azúcares, metabolitos, precursores, cofactores u otros agentes químicos o entidades biológicamente activos, que alteran la selectividad de la proteína polipéptido, péptido, ADN, ARN, sacáridos, azúcares, metabolitos, precursores, cofactores u otros agentes químicos o entidades biológicamente activos, o aumentar o disminuir la selectividad del sustrato de la proteína, polipéptido, péptido, ADN, ARN, sacáridos, azúcares, metabolitos, precursores, cofactores u otros agentes químicos o entidades biológicamente activos.

El término "biomaterial", tal como se usa en el presente documento, se refiere a un material derivado biológicamente, incluyendo, pero sin limitarse al material obtenido a partir de biorreactores y/o de métodos y técnicas recombinantes.

El término "sonda biofísica", tal como se usa en el presente documento, se refiere a sondas que permiten la detección de o vigilar los cambios estructurales en moléculas mediante métodos de detección físicos. Dichas moléculas incluyen, pero no se limitan a, proteínas, polipéptido, péptidos, ADN o ARN. Dicha "sonda biofísica" se usa también para detectar o vigilar la interacción de las proteínas, polipéptidos, péptidos, ADN o ARN con otras moléculas, incluyendo pero sin limitarse a, macromoléculas. Los ejemplos de sondas biofísicas incluyen, pero no se limitan a, la masa molecular, espines nucleares, absorbancia del UV, fluorescencia, dicroísmo circular, capacidad calorífica, temperatura de fusión u otras propiedades moleculares intrínsecas. Los ejemplos de sondas biofísicas incluyen también marcas que se añaden a la molécula. Dichas sondas incluyen, pero no se limitan a, marcas de espín, fluoróforos, marcas de isótopos, y grupos fotoactivables.

El término "generado biosintéticamente", tal como se usa en el presente documento, se refiere a cualquier método que utiliza una célula o enzimas para generar un aminoácido. Dichos métodos incluyen el uso de al menos uno de los

siguientes componentes: un precursor y una enzima. En determinadas realizaciones, dichos aminoácidos se incorporan a continuación a una proteína. En determinadas realizaciones, la biosíntesis y la incorporación del aminoácido se produce en la misma célula, mientras que en otras realizaciones, el aminoácido se genera biosintéticamente en una célula separada ("célula alimentadora"), o en un cultivo celular separado, y el aminoácido se incorpora a una proteína en otra célula. En el último caso, el aminoácido se purifica opcionalmente a partir del cultivo celular separado, y el aminoácido purificado se añade a continuación al medio del cultivo celular que incorpora el aminoácido a la proteína

El término "análogo de biotina" o denominado también "biotina mimetizada", tal como se usa en el presente documento, es cualquier molécula, diferente de biotina, que se une con elevada afinidad a la avidina y/o a la estreptavidina.

El término "grupo de modificación del extremo carboxi" se refiere a cualquier molécula que se puede unir a un grupo en el extremo carboxi. Dichos grupos del extremo carboxi incluyen, pero no se limitan a, grupos protectores carboxilados, el extremo de moléculas poliméricas, donde dichas moléculas poliméricas incluyen, pero no se limitan a, polipéptidos, polinucleótidos, y polisacáridos. Los grupos de modificación del extremo incluyen también, pero no se limitan a, diversos polímeros, péptidos o proteínas solubles en agua. Por medio solo de ejemplo, los grupos de modificación del extremo terminal incluyen polietilenglicol o albúmina de suero. Determinados grupos de modificación del extremo carboxi se usan para modificar características terapéuticas de una proteína, polipéptidos o péptidos, incluyendo pero sin limitarse a, aumentar la semivida en suero.

El término "grupo escindible químicamente" denominado también "químicamente lábil", tal como se usa en el presente documento, se refiere a un grupo que se rompe o escinde tras la exposición a un ácido, base, agentes oxidantes, agentes reductores, iniciadores químicos, o iniciadores radicalarios.

El término "grupo quimioluminiscente", tal como se usa en el presente documento, se refiere a un grupo que emite luz como resultado de una reacción química sin la adición de calor. Por medio solo de ejemplo, luminol (5-amino-2,3-dihidro-1,4-ftalazinediona) reacciona con oxidantes del tipo del peróxido de hidrógeno (H_2O_2) en presencia de una base y un catalizador metálico para producir un producto en estado excitado (3-aminoftalato, 3-APA) dando como resultado posteriormente la liberación de luz detectable.

El término "cromóforo", tal como se usa en el presente documento, se refiere a una molécula que absorbe luz de longitudes de onda visibles, longitudes de onda UV, o longitudes de onda IR.

El término "cofactor", tal como se usa en el presente documento, se refiere a un átomo o molécula esencial para la acción de una molécula grande. Los cofactores incluyen, pero no se limitan a, iones inorgánicos, coenzimas, proteínas, o algún otro factor necesario para la actividad de las enzimas.

El término "plegado simultáneo", tal como se usa en el presente documento, se refiere a procedimientos de plegado simultáneo, reacciones, o métodos que emplean al menos dos moléculas que interactúan entre sí y dan como resultado la transformación de moléculas no plegadas o plegadas de forma inadecuada a moléculas plegadas de forma adecuada. Por medio solo de ejemplo, el "plegado simultáneo" emplea al menos dos polipéptidos que interactúan entre sí y dan como resultado la transformación de los polipéptidos no plegados o plegados de forma inadecuada en polipéptidos naturales, plegados de forma adecuada.

El término "citotóxico", tal como se usa en el presente documento, se refiere a un compuesto que perjudica a las células.

El término "agente desnaturalizante" o "desnaturalizante", tal como se usa en el presente documento, se refiere a cualquier compuesto o material que puede producir un plegado reversible de una proteína. La fuerza del agente desnaturalizante o desnaturalizante se determinará por las propiedades y la concentración del agente desnaturalizante o desnaturalizante concreto. A modo de ejemplo, los agentes desnaturalizantes o desnaturalizantes incluyen, pero no se limitan a, caótropos, detergentes, disolventes orgánicos, miscibles en agua, fosfolípidos, o una de sus combinaciones. Los ejemplos no limitantes de caótropos incluyen, pero no se limitan a, urea, guanidina, y tiocianato de sodio. Los ejemplos no limitantes de detergentes pueden incluir, pero no se limitan a, detergentes fuertes tales como dodecil sulfato de sodio, o éteres de polioxietileno (por ejemplo Tween o detergentes Triton), Sarkosyl, detergentes no iónicos suaves (por ejemplo, digitonina), detergentes catiónicos suaves tales como N-2,3-(Dioleioxi-propil-N,N,N-trimetilamonio, detergentes iónicos suaves (por ejemplo, colato de sodio o desoxicolato de sodio) o detergentes de ion híbrido que incluyen pero sin limitación, sulfo- betaínas (Zwittergent), 3-(3-clolamidopropil)dimetilamonio-1-propano sulfato (CHAPS), y 3-(3-clolamidopropil)dimetilamonio-2-hidroxi-1-propano sulfonato (CHAPSO). Los ejemplos no limitantes de disolventes orgánicos, miscibles en agua incluyen, pero no se limitan a, acetonitrilo, alcanoles inferiores (especialmente alcanoles C_2 - C_4 tales como etanol o isopropanol), o alcanodioles inferiores (alcanodioles C_2 - C_4 tales como etilenglicol) se pueden usar como desnaturalizantes. Los ejemplos no limitantes de fosfolípidos incluyen, pero no se limitan a, fosfolípidos de origen natural tales como fosfatidiletanolamina, fosfatidicolina, fosfatidilserina, y fosfatidilinositol o derivados o variantes de fosfolípidos sintéticos tales como dihexanoilfosfatidilcolina o

diheptanoilfosfatidilcolina.

- 5 El término "marca detectable", tal como se usa en el presente documento, se refiere a una marca que se puede observar utilizando técnicas analíticas que incluyen, pero sin limitación, fluorescencia, quimioluminiscencia, resonancia de espín electrónico, espectroscopía de absorción en el ultravioleta/visible, espectroscopía infrarroja, espectrometría de masas, resonancia magnética nuclear, resonancia magnética, métodos radiométricos y electroquímicos.
- 10 El término "fármaco", tal como se usa en el presente documento, se refiere a cualquier sustancia usada en la prevención, diagnóstico, alivio, el tratamiento, o cura de una enfermedad o dolencia.
- El término "colorante" tal como se usa en el presente documento, se refiere a una sustancia colorante soluble, que contiene un cromóforo.
- 15 El término "grupo denso en electrones", tal como se usa en el presente documento, se refiere a un grupo que dispersa electrones cuando se irradia con un haz de electrones. Dichos grupos incluyen, pero no se limitan a, molibdato de amonio, yoduro de cadmio subnitrito de bismuto, 99 %, carbohidrazida, cloruro férrico hexahidrato, hexametileno tetramina, 98,5 %, tricloruro de indio anhidro, nitrato de lantano, trihidrato de acetato de plomo, trihidrato de citrato de plomo, nitrato de plomo, ácido peryódico, ácido fosfomolibdico, ácido fosfotúngstico, ferricianuro de potasio, ferrocianuro de potasio, rutenio rojo, nitrato de plata, proteinato de plata (Ensayo de Ag: 8,0 -8,5 %, "Fuerte";: tetrafenilporfina de plata (S-TPPS), cloroaurato de sodio, tungstato de sodio, nitrato de talio, tiosemicarbazida (TSC), acetato de uranilo, nitrato de uranilo, y sulfato de vanadilo.
- 20 El término "agente de transferencia de energía", tal como se usa en el presente documento, se refiere a una molécula que puede tanto donar como aceptar energía de otra molécula. Por medio solo de ejemplo, la transferencia de energía de resonancia de fluorescencia (FRET) es un procedimiento de acoplamiento dipolo-dipolo por el cual la energía del estado excitado de una molécula donante de fluorescencia se transfiere de forma no radiante a una molécula aceptora no excitada que a continuación emite fluorescentemente la energía donada a una longitud de onda más larga.
- 25 Los términos "aumentar" o "que aumenta" significa aumentar o prolongar tanto la potencia como la duración de un efecto deseado. A modo de ejemplo, "potenciar" el efecto de los agentes terapéuticos se refiere a la capacidad de aumentar o prolongar, tanto la potencia como la duración, del efecto de los agentes terapéuticos durante el tratamiento de una enfermedad, trastorno o dolencia. Una "cantidad potenciadora eficaz", tal como se usa en el presente documento, se refiere a una cantidad adecuada para potenciar el efecto de un agente terapéutico en el tratamiento de una enfermedad, trastorno o dolencia. Cuando se usa en un paciente, las cantidades eficaces para este uso dependerán de la gravedad y el curso de la enfermedad, trastorno o dolencia, tratamientos anteriores, el estado de salud del paciente y la respuesta a los fármacos, y el criterio del médico a cargo del tratamiento.
- 30 El término "eucariota" se refiere a un material que se origina a partir de organismos que pertenecen al dominio filogenético Eucarya, que incluye, pero que no se limita a animales (incluyendo, pero sin limitarse a, mamíferos, insectos, reptiles, y pájaros), ciliados, plantas (incluyendo, pero sin limitarse a, monocotiledóneas, dicotiledóneas, y algas), hongos, levaduras, flagelados, microsporidia, y protistas
- 35 El término "ácido graso", tal como se usa en el presente documento, se refiere a ácidos carboxílicos con aproximadamente C₆ o una cadena lateral de hidrocarburo más larga.
- 40 El término "fluoróforo", tal como se usa en el presente documento, se refiere a una molécula que tras la excitación emite fotones y es por tanto fluorescente.
- 45 Los términos "grupo funcional", "resto activo", "grupo de activación", "grupo saliente", "sitio reactivo", "grupo químicamente reactivo" y "resto químicamente reactivo", son algo sinónimos en las técnicas químicas y se usan en el presente documento para indicar las porciones de moléculas que llevan a cabo alguna función o actividad y son reactivas con otras moléculas.
- 50 El término "idéntico", tal como se usa en el presente documento, se refiere a dos o más secuencias o subsecuencias que son iguales. Además, El término "sustancialmente idéntico", tal como se usa en el presente documento, se refiere a dos o más secuencias que tienen un porcentaje de unidades secuenciales que son iguales cuando se las compara y alinea para la correspondencia máxima sobre una ventana de comparación, o la región designada como medida usando los algoritmos de comparación o mediante alineación manual e inspección visual. Por medio solo de ejemplo,
- 55 dos o más secuencias pueden ser "sustancialmente idénticas" si las unidades secuenciales son idénticas en aproximadamente un 60 %, idénticas en aproximadamente un 65 %, idénticas en aproximadamente un 70 %, idénticas en aproximadamente un 75 %, idénticas en aproximadamente un 80 %, idénticas en aproximadamente un 85 %, idénticas en aproximadamente un 90 %, o idénticas en aproximadamente un 95 % en una región especificada. Dichos porcentajes describen el "porcentaje de identidad" de dos o más secuencias. La identidad de una secuencia puede existir en una región que tenga al menos aproximadamente 75-100 unidades secuenciales de longitud, en una región que tenga aproximadamente 50 unidades secuenciales de longitud, o bien, cuando no se especifica, para la totalidad
- 60
- 65

de la secuencia. Esta definición se refiere también al complemento de la secuencia de ensayo. Por medio solo de ejemplo, dos o más secuencias de polipéptidos son idénticas cuando los restos de aminoácidos son los mismos, mientras que dos o más secuencias de polipéptidos son "sustancialmente idénticas" si los restos de aminoácidos son idénticos en aproximadamente un 60 %, idénticos en aproximadamente un 65 %, idénticos en aproximadamente un 70 %, idénticos en aproximadamente un 75 %, idénticos en aproximadamente un 80 %, idénticos en aproximadamente un 85 %, idénticos en aproximadamente un 90 %, o idénticos en aproximadamente un 95 % en una región especificada. La identidad puede existir en una región que tenga al menos aproximadamente 75-100 aminoácidos de longitud, en una región que tenga aproximadamente 50 aminoácidos de longitud, o bien, cuando no se especifica, a lo largo de la secuencia completa de una secuencia polipeptídica. Además, a modo de ejemplo solamente, dos o más secuencias de polipéptidos son idénticas cuando los restos de ácidos nucleicos son los mismos, mientras que dos o más secuencias de polinucleótidos son "sustancialmente idénticas" si los restos de ácidos nucleicos son idénticos en aproximadamente un 60 %, idénticos en aproximadamente un 65 %, idénticos en aproximadamente un 70 %, idénticos en aproximadamente un 75 %, idénticos en aproximadamente un 80 %, idénticos en aproximadamente un 85 %, idénticos en aproximadamente un 90 %, o idénticos en aproximadamente un 95 % en una región especificada. La identidad puede existir en una región que tenga al menos aproximadamente 75-100 ácidos nucleicos de longitud, en una región que tenga aproximadamente 50 ácidos nucleicos de longitud, o bien, cuando no se especifica, a lo largo de la secuencia completa de una secuencia de polinucleótidos.

El término "inmunogenicidad", tal como se usa en el presente documento, se refiere a una respuesta a anticuerpo respecto a la administración de un fármaco terapéutico. La inmunogenicidad hacia las proteínas, polipéptidos y péptidos terapéuticos proporcionados en el presente documento se obtiene usando ensayos cuantitativos y cualitativos para la detección de anticuerpos frente a dichas proteínas, polipéptidos y péptidos terapéuticos en fluidos biológicos. Dichos ensayos incluyen, pero no se limitan a, radioinmunoensayo (RIA), enzoinmunoanálisis de adsorción (ELISA), inmunoensayo luminiscente (LIA) e inmunoensayo fluorescente (FIA). El análisis de dicha inmunogenicidad implica comparar la respuesta al anticuerpo tras la administración de proteínas, polipéptidos y péptidos terapéuticos proporcionados en el presente documento para la respuesta al anticuerpo tras la administración de una proteína, polipéptido o péptido terapéutico del control o del vehículo de administración o del tampón de administración.

El término "agente intercalante" denominado también "grupo intercalante", tal como se usa en el presente documento, se refiere a una molécula o grupo que se inserta en el espacio intramolecular de otra molécula o en el espacio intermolecular entre moléculas. A modo de ejemplo solamente, un agente o grupo intercalante puede ser una molécula que se inserta en las bases apiladas de la doble hélice del ADN.

El término "marca", tal como se usa en el presente documento, se refiere a una sustancia que se incorpora a un compuesto y es fácilmente detectable, por lo que su distribución física se puede detectar y/o vigilar.

El término "enlace", tal como se usa en el presente documento, se refiere a un enlace o resto químico formado a partir de una reacción química entre el grupo funcional de una primera molécula con el grupo funcional de una segunda molécula. Dichos enlaces incluyen, pero no se limitan a, enlaces covalentes y enlaces no covalentes, mientras que dichos restos químicos incluyen, pero no se limitan a, ésteres, carbonatos, ésteres de fosfato de iminas, hidrazonas, acetales, ortoésteres, enlaces peptídicos, enlaces de oligonucleótidos y aquellos proporcionados en la Tabla 1 en el presente documento. "Enlaces hidrolíticamente estables" significa que los enlaces son sustancialmente estables en agua y no reaccionan con agua a valores de pH útiles, incluyendo pero sin limitarse a, en condiciones fisiológicas durante un periodo extendido de tiempo. "Enlaces hidrolíticamente inestables o degradables" significa que los enlaces son degradables en agua o en disoluciones acuosas, incluyendo, por ejemplo, sangre, "Enlaces enzimáticamente inestables o degradables" significa que los enlaces son degradados por una o más enzimas, Por medio solo de ejemplo, determinados PEG y polímeros relacionado incluyen enlaces degradables en la estructura polimérica o en un grupo enlazador entre la estructura del polímero de PEG y uno o más grupos funcionales terminales de la proteína, polipéptido o péptido proporcionado en el presente documento. Dichos enlaces degradables incluyen, pero no se limitan a, enlaces éster formados mediante la reacción de ácidos carboxílicos de PEG o ácidos carboxílicos activados con PEG con grupos alcohólicos de un agente biológicamente activo, donde dichos grupos éster se hidrolizan generalmente en condiciones fisiológicas para liberar el agente biológicamente activo. Otros enlaces hidrolíticamente degradables incluyen, pero no se limitan a enlaces carbonato; enlaces imina resultado de la reacción de una amina y un aldehído; enlaces de ésteres de fosfato haciendo reaccionar un alcohol con un grupo fosfato; enlaces hidrazona que son un producto de reacción entre una hidrazida y un aldehído; enlaces acetal que son un producto de reacción entre un aldehído y un alcohol; enlaces ortoéster que son el producto de reacción entre un formiato y un alcohol; enlaces peptídicos formados por un grupo amina, incluyendo pero sin limitarse a, en un extremo de un polímero tal como PEG, y un grupo carboxilo de un péptido; y enlaces de oligonucleótidos formados por un grupo fosforamídita, incluyendo pero sin limitarse a, al extremo de un polímero, y un grupo 5' hidroxilo de un oligonucleótido.

Los términos "medio" o "medios", tal como se usan en el presente documento, se refieren a cualquier medio de cultivo utilizado para hacer crecer y cosechar células y/o productos expresados y/o secretados por dichas células. Dicho "medio" o "medios" incluye, pero no se limitan a, solución, sólido, semisólido, o soportes rígidos que pueden soportar o contener cualquier célula hospedadora, incluyendo, a modo de ejemplo, células hospedadoras bacterianas, células hospedadoras de levaduras, células hospedadoras de insectos, células hospedadoras de plantas, células

- 5 hospedadoras eucariotas, células hospedadoras de mamíferos, células CHO, células hospedadoras procariotas, *Escherichia coli*, o células hospedadoras de *Pseudomonas*, y los contenidos celulares. Dicho "medio" o "medios" incluye, pero sin limitación, medio o medios en los que se ha hecho crecer la célula hospedadora o en los que se ha secretado un polipéptido, incluyendo el medio tanto antes como después de una etapa de proliferación. Dicho "medio" o "medios" incluye también, pero sin limitación, tampones o reactivos que contienen lisados de células hospedadoras, a modo de ejemplo, un polipéptido producido intracelularmente y las células hospedadoras se lisan o perturban para liberar el polipéptido.
- 10 El término "quelante metálico", tal como se usa en el presente documento, se refiere a una molécula que forma un complejo con iones metálicos. A modo de ejemplo, dichas moléculas forman dos o más enlaces de coordinación con un ion metálico central y forma opcionalmente estructuras de anillo.
- 15 El término "resto contenedor de metal", tal como se usa en el presente documento, se refiere a un grupo que contiene un ion, átomo o partícula metálica. Dichos restos incluyen, pero no se limitan a, cisplatino, iones metálicos quelados (tales como níquel, hierro, y platino), y nanopartículas metálicas (tales como níquel, hierro, y platino).
- 20 El término "resto que incorpora un átomo pesado", tal como se usa en el presente documento, se refiere a un grupo que incorpora un ion de un átomo que es usualmente más pesado que el carbono. Dichos iones o átomos incluyen, pero no se limitan a, silicón, tungsteno, oro, plomo, y uranio.
- 25 El término "modificado", tal como se usa en el presente documento se refiere a la presencia de un cambio de un aminoácido o resto de aminoácido, donde dichos cambios, o modificaciones, se obtienen mediante procedimientos químicos o bioquímicos.
- 30 Como se usa en el presente documento, el término "semivida en suero modulada" se refiere a cambios positivos o negativos en la semivida en circulación de una molécula biológicamente activa modificada con respecto a su forma no modificada. A modo de ejemplo, Las moléculas biológicamente activas modificadas incluyen, pero no se limitan a, los compuestos proporcionados en el presente documento. A modo de ejemplo, se midió la semivida en suero tomando muestras de sangre en diversos puntos temporales tras la administración de la molécula biológicamente activa o la molécula biológicamente activa modificada, y determinando la concentración de esta molécula en cada muestra. La correlación de la concentración en suero con el tiempo permite el cálculo de la semivida en suero. A modo de ejemplo, la semivida en suero modulada puede ser una semivida en suero aumentada, que puede permitir regímenes de dosificación mejorados o evitar efectos tóxicos. Dichos aumentos en suero pueden ser al menos aproximadamente de dos veces, al menos aproximadamente tres veces, al menos aproximadamente cinco veces, o al menos aproximadamente diez veces.
- 35 El término "nanopartícula", tal como se usa en el presente documento, se refiere a una partícula que tiene un tamaño de partícula entre aproximadamente 500 nm y aproximadamente 1 nm.
- 40 El término "casi estequiométrico", tal como se usa en el presente documento, se refiere a la relación de los moles de compuestos que participan en la reacción química que es aproximadamente de 0,75 a aproximadamente 1,5.
- 45 Como se usa en el presente documento, el término "no eucariota" se refiere a organismos no eucarióticos. A modo de ejemplo, un organismo no eucariótico pertenece al dominio filogenético de las Eubacterias, que incluye, pero sin limitarse a, *Escherichia coli*, *Thermus thermophilus*, o *Bacillus stearothermophilus*, *Pseudomonas fluorescens*, *Pseudomonas aeruginosa*, *Pseudomonas putida*, o el dominio filogenético Archaea, que incluye, pero sin limitación, *Methallobacillus jannaschii*, *Methanobacterium thermoautotrophicum*, *Archaeoglobus fulgidus*, *Pyrococcus furiosus*, *Pyrococcus horikoshii*, *Aeuro-pyrum pernix*, o *Halobacterium* tal *Haloferax volcanii* y especies de *Halobacterium* NRC-1.
- 50 El término "ácido nucleico", tal como se usa en el presente documento, se refiere a desoxirribonucleótidos, desoxirribonucleósidos, ribonucleósidos o ribonucleótidos y sus polímeros en forma tanto monocatenaria como bicatenaria. Por medio solo de ejemplo, dichos ácidos nucleicos y polímeros de ácidos nucleicos incluyen, pero no se limitan a, (i) análogos de nucleótidos naturales que tienen propiedades similares que un ácido nucleico de referencia; (ii) análogos de oligonucleótidos que incluyen, pero no se limitan a, PNA (ácido peptidonucleico), Análogos de ADN utilizados en tecnología antisentido (fosforotioatos, fosforoamidatos, y similares); (iii) sus variantes modificadas conservativamente (incluyendo, pero sin limitarse a, sustituciones de codones degenerados) y secuencias complementarias y secuencias indicadas explícitamente. A modo de ejemplo, se pueden conseguir sustituciones de codones degenerados generando secuencias en las que la tercera posición de uno o más codones seleccionados (o todos) se sustituye con una base mixta y/o restos de desoxiinosina.
- 60 El término "agente oxidante", tal como se usa en el presente documento, se refiere a un compuesto o material que es capaz de eliminar un electrón de un compuesto que se ha oxidado. A modo de ejemplo, los agentes oxidantes incluyen, pero no se limitan a, glutatona oxidada, cistina, cistamina, ditiotreitól oxidado, eritreitól oxidado y oxígeno. una amplia variedad de agentes oxidantes son adecuados para el uso en los métodos y composiciones descritos en el presente documento.
- 65

El término "marca de fotoafinidad", tal como se usa en el presente documento, se refiere a una marca con un grupo, que, tras la exposición a la luz, forma un enlace con una molécula por la cual la marca tiene una afinidad. Por medio solo de ejemplo, dicho enlace puede ser covalente o no covalente.

- 5 El término "resto fotoestimulado", tal como se usa en el presente documento, se refiere a un grupo que, tras la iluminación a determinadas longitudes de onda, se une de forma covalente o no covalente a iones u otras moléculas.

El término "grupo fotoescindible", tal como se usa en el presente documento, se refiere a un grupo que se rompe tras la exposición a ácidos

- 10 El término "fotorreticulador", tal como se usa en el presente documento, se refiere a un compuesto que comprende dos o más grupos funcionales que, tras la exposición a la luz, son reactivos y forman un enlace covalente o no covalente con dos o más moléculas monoméricas o poliméricas.

- 15 El término "resto fotoisomerizable", tal como se usa en el presente documento, se refiere a un grupo donde tras la iluminación la luz cambia de una forma isomérica a otra

- 20 El término "polialquilenglicol", tal como se usa en el presente documento, se refiere a poliéter polioles poliméricos de cadena lineal o ramificada. tales como polialquilenglicoles, incluyendo, pero no se limitan a, polietilenglicol, propilenglicol, polibutilenglicol, y sus derivados.

- 25 El término "polímero", tal como se usa en el presente documento, se refiere a una molécula compuesta de subunidades repetidas. Dichas moléculas incluyen, pero no se limitan a, proteínas, polipéptidos, péptidos, polinucleótidos, o polisacáridos o polialquilenglicoles.

- 30 Los términos "polipéptido", "péptido" y "proteína" se usan de forma indistinta en el presente documento para referirse a un polímero de restos de aminoácidos. Esto es, una descripción dirigida a un polipéptido se aplica igualmente a la descripción de un péptido y la descripción de una proteína, y viceversa. Los restos de aminoácidos incluyen los restos resultantes de aminoácidos naturales y no naturales. Los términos se aplican a polímeros de aminoácidos de origen natural así como polímeros de aminoácidos en los que uno o más restos de aminoácidos es un compuesto proporcionado en el presente documento. Adicionalmente, dichos "polipéptidos", "péptidos" y "proteínas" incluyen cadenas de aminoácidos de cualquier longitud, incluyendo proteínas de longitud completa, donde los restos de aminoácidos se unen mediante enlaces covalentes peptídicos u otros enlaces.

- 35 El término "modificado después de la traducción" se refiere a cualquier modificación de un aminoácido que se produce después que dicho aminoácido se ha incorporado traduccionalmente a una cadena polipeptídica. Dichas modificaciones incluyen, pero no se limitan a, modificaciones posteriores a la traducción in vivo, y modificaciones posteriores a la traducción in vitro.

- 40 El término "protegido", tal como se usa en el presente documento, se refiere a la presencia de un "grupo protector" o resto que evita la reacción del grupo funcional químicamente reactivo en determinadas condiciones de reacción. El grupo protector variará dependiendo del tipo de grupo químicamente reactivo que se está protegiendo. Por medio solo de ejemplo, (i) si el grupo químicamente reactivo es una amina o una hidrazida, el grupo protector se puede seleccionar entre terc-butiloxicarbonilo (t-Boc) y 9- fluorenilmetoxicarbonilo (Fmoc); (ii) si el grupo químicamente reactivo es un tiol, el grupo protector puede ser ortopiridildisulfuro; y (iii) si el grupo químicamente reactivo es un ácido carboxílico, tal como ácido butanoico o propiónico, o un grupo hidroxilo, el grupo protector puede ser bencilo o un grupo alquilo tal como metilo, etilo, o terc-butilo. Por medio solo de ejemplo, grupos bloqueantes/protectores se seleccionan también entre: acetamida, aliloxicarbonilo, alil éter bencilo, bencilamina, bencilidenamina, bencil carbamato, ésteres de bencilo, éster de metilo, éster de t-butilo, éster de S-t-butilo, 2-alkil-1,3-oxazolona, dimetil acetal, 1,3-dioxano, 1,3-ditiano, N,N-dimetilhidrazona, ftalimida, tritilo, éter de p-metoxibencilo, Carbobenciloxi, p-toluensulfonamida, trifluoracetamida, trifenilmetilamina, éter de n-butilo, éter de bencilo, éter de butildimetilsililo, éter de t-butildifenilsililo, 2-(trimetilsilil) etoxicarbonilo, éster de ácido acético, éster de ácido piválico, éster de ácido benzoico, acetona, benciliden acetal, y grupos fotolábiles tales como Nvoc y MeNvoc.

- 55 El término "resto radioactivo", tal como se usa en el presente documento, se refiere a un grupo cuyos núcleos liberan espontáneamente radiación nuclear, tales como partículas alfa, o beta, o radiación gamma.

El término "compuesto reactivo", tal como se usa en el presente documento, se refiere a compuesto que en condiciones adecuadas es reactivo hacia otro átomo, molécula o compuesto.

- 60 El término "célula hospedadora recombinante" denominado también "célula hospedadora", se refiere a una célula que incluye un polinucleótido exógeno, donde el método usado para insertar el polinucleótido exógeno en una célula incluye, pero no se limitan a, captación directa, transducción, correspondencia-f, u otros métodos conocidos en la materia para crear células hospedadoras recombinantes. Por medio solo de ejemplo, dicho polinucleótido exógeno puede ser un vector no integrado, incluyendo, pero sin limitarse a un plásmido, o se puede integrar en el genoma hospedador.

- 65

El término "agente redox activo", tal como se usa en el presente documento, se refiere a una molécula que oxida o reduce otra molécula, por lo que el agente redox activo llega a reducirse u oxidarse. Los ejemplos de agente redox activo incluyen, pero no se limitan a, ferroceno, quinonas, Complejos $Ru^{2+/3+}$, Complejos $Co^{2+/3+}$, y complejos $Os^{2+/3+}$.

El término "agente reductor", tal como se usa en el presente documento, se refiere a un compuesto o material que es capaz de añadir un electrón a un compuesto que se ha reducido. A modo de ejemplo, los agentes reductores incluyen, pero no se limitan a, ditiotreitól (DTT), 2-mercaptoetanol, ditiotritól, cisteína, cisteamina (2-aminoethanotiol), y glutatión reducido. Dichos agentes reductores se pueden utilizar, a modo de ejemplo solamente, para mantener los grupos sulfidriló en el estado reducido y reducir los enlaces disulfuro intra o intermoleculares.

El término "replegado", tal como se usa en el presente documento describe cualquier procedimiento, reacción o método que transforma un estado plegado inadecuadamente o no plegado a una conformación natural o plegada adecuadamente. Por medio solo de ejemplo, el replegado transforma el enlace disulfuro que contiene las proteínas o polipéptidos de un estado plegado inadecuadamente o un estado no plegado a una conformación natural o plegada adecuadamente con respecto a los enlaces disulfuro. Dichas proteínas o polipéptidos que contienen un enlace disulfuro incluyen proteínas o polipéptidos que tiene incorporado en el anterior compuestos proporcionados en el presente documento. El replegado se inicia a menudo mediante la eliminación de agentes caótrópos tales como urea o clorhidrato de guanidinio añadido anteriormente a las disoluciones de proteína a fin de solubilizar y desplegar dichas proteínas.

El término "resina", tal como se usa en el presente documento, se refiere a perlas de polímeros insolubles de elevado peso molecular. Por medio solo de ejemplo, dichas perlas se pueden usar como soportes para síntesis de péptidos en fase sólida, o sitios para la unión de moléculas antes de la purificación.

El término "sacárido", tal como se usa en el presente documento, se refiere a una serie de hidratos de carbono que incluye, pero no se limita a azúcares, monosacáridos, oligosacáridos, y polisacáridos.

El término "marca de de espín", tal como se usa en el presente documento, se refiere a moléculas que contienen un átomo o un grupo de átomos que presentan un espín electrónico desemparejado (es decir, un grupo paramagnético estable) que se puede detectar mediante espectroscopía de resonancia de espín electrónico y se puede unir a otra molécula. Dichas moléculas de marca de espín incluyen, pero no se limitan a, radicales nitrilo y nitroxidos, y pueden ser marcas de espín único o marcas de espín doble.

El término "estequiométrico", tal como se usa en el presente documento, se refiere a la relación de los moles de compuestos que participan en la reacción química que es aproximadamente de 0,9 a aproximadamente 1,1.

El término "sustancialmente purificado", tal como se usa en el presente documento, se refiere a un componente de interés que está sustancialmente o esencialmente exento de otros componentes que acompañan o interactúan normalmente con el componente de interés antes de la purificación. Por medio solo de ejemplo, un componente de interés puede estar "sustancialmente purificado" cuando la preparación del componente de interés contiene menos de aproximadamente 30 %, menor de aproximadamente 25 %, menor de aproximadamente 20 %, menor de aproximadamente 15 %, menor de aproximadamente 10 %, menor de aproximadamente 5 %, menor de aproximadamente 4 %, menor de aproximadamente 3 %, menor de aproximadamente 2 %, o menos de aproximadamente 1 % (en peso seco) de componentes contaminantes. Por tanto, un "componente sustancialmente purificado" de interés puede tener un nivel de pureza de aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90 %, aproximadamente 95 %, aproximadamente 96 %, aproximadamente 97 %, aproximadamente 98 %, aproximadamente 99 % o superior. Por medio solo de ejemplo, una proteína, polipéptido o péptido que contiene un compuesto proporcionado en el presente documento está purificado a partir de una célula natural o una célula hospedadora. Por medio solo de ejemplo, una preparación de una proteína, polipéptido o péptido que contiene un compuesto proporcionado en el presente documento está "sustancialmente purificado" cuando la preparación contiene menos de aproximadamente 30 %, menor de aproximadamente 25 %, menor de aproximadamente 20 %, menor de aproximadamente 15 %, menor de aproximadamente 10 %, menor de aproximadamente 5 %, menor de aproximadamente 4 %, menor de aproximadamente 3 %, menor de aproximadamente 2 %, o menos de aproximadamente 1 % (en peso seco) de material contaminante. Por medio solo de ejemplo, cuando una proteína, polipéptido o péptido que contiene un compuesto proporcionado en el presente documento se produce de manera recombinante por células hospedadoras, la proteína, polipéptido o péptido que contiene un compuesto proporcionado en el presente documento está presente en aproximadamente un 30 %, aproximadamente 25 %, aproximadamente 20 %, aproximadamente 15 %, aproximadamente 10 %, aproximadamente 5 %, aproximadamente 4 %, aproximadamente 3 %, aproximadamente 2 %, o aproximadamente un 1 % o menos del peso seco de las células. Por medio solo de ejemplo, cuando una proteína, polipéptido o péptido que contiene un compuesto proporcionado en el presente documento se produce de manera recombinante por células hospedadoras, la proteína, polipéptido o péptido que contiene un compuesto proporcionado en el presente documento está presente en el medio de cultivo a aproximadamente 5 g/l, aproximadamente 4 g/l, aproximadamente 3 g/l, aproximadamente 2 g/l, aproximadamente 1 g/l, aproximadamente 750 mg/l, aproximadamente 500 mg/l, aproximadamente 250 mg/l, aproximadamente 100 mg/l, aproximadamente 50 mg/l,

aproximadamente 10 mg/l, o aproximadamente 1 mg/l o menos. Por medio solo de ejemplo, una proteína "sustancialmente purificada", polipéptido o péptido que contiene un compuesto proporcionado en el presente documento tiene un nivel de pureza de aproximadamente un 30 %, aproximadamente 35 %, aproximadamente 40 %, aproximadamente 45 %, aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 65 %, aproximadamente 70 %, aproximadamente 75 %, aproximadamente 80 %, aproximadamente 85 %, aproximadamente 90, aproximadamente 95 %, aproximadamente 99 % o más tal como se determina mediante los métodos adecuados, incluyendo, pero sin limitación, análisis SDS/PAGE, RP-HPLC, SEC, y electroforesis capilar,

El término "resto tóxico", tal como se usa en el presente documento, se refiere a un compuesto que puede producir perjuicio o muerte.

los términos "trata", "que trata", o "tratamiento", tal como se usan en el presente documento, incluye aliviar, disminuir o mejorar los síntomas de una enfermedad o dolencia, evitar síntomas adicionales, mejorar o evitar las causas metabólicas subyacentes de los síntomas, inhibir la enfermedad o dolencia, por ejemplo, detener el desarrollo de la enfermedad o dolencia, aliviar la enfermedad o dolencia, producir la regresión de la enfermedad o dolencia, aliviar una dolencia producida por la enfermedad o dolencia, o detener los síntomas de la enfermedad o dolencia, los términos "trata", "que trata", o "tratamiento", incluyen, pero no se limitan a, tratamientos profilácticos y/o terapéuticos.

Como se usa en el presente documento, el término "polímero soluble en agua" se refiere a cualquier polímero que es soluble en disolventes acuosos. Dichos polímeros solubles en agua incluyen, pero no se limitan a, polietilenglicol, polietilenglicol propionaldehído, monoalcoxi C₁-C₁₀ o derivados de ariloxi de los mismos monometoxi-polietilenglicol, polivinilpirrolidona, poli(alcohol vinílico), poliaminoácidos, divinil éter de anhídrido maleico, N-(2-Hidroxipropil)-metacrilamida, dextrano, los derivados de dextrano incluyen sulfato de dextrano, propilenglicol, copolímero de óxido de polipropileno/óxido de etileno, polioli polioxiethylado, heparina, fragmentos de heparina, polisacáridos, oligosacáridos, glicanos, celulosa y derivados de celulosa, incluyendo, pero sin limitarse a metilcelulosa y carboximetil celulosa, albúmina de suero, almidón y derivados de almidón, polipéptidos, polialquilenglicol y sus derivados, copolímeros de polialquilenglicoles y sus derivados, éteres de polivinil etilo, y alfa-beta-poli[(2-hidroxietil)-DL-aspartamida, y similares, o mezclas de los mismos. Por medio solo de ejemplo, el acoplamiento de dichos polímeros solubles en agua con proteínas, polipéptidos o péptidos que contienen un compuesto proporcionado en el presente documento dan como resultado cambios que incluyen, pero sin limitación, solubilidad en agua aumentada, semivida en suero aumentada o modulada, semivida terapéutica aumentada o modulada con respecto a la forma sin modificar, biodisponibilidad aumentada, actividad biológica modulada, tiempo de circulación extendido, inmunogenicidad moduladas, las características de asociación física modulada que incluyen, pero sin limitación, agregación y formación de multímeros, unión al receptor alterada, unión alterada con uno o más ligandos, y dimerización o multimerización del receptor alterada.

Otros objetos, características y ventajas de los métodos, composiciones y combinaciones descritos en el presente documento llegaran a ser evidentes a partir de la siguiente descripción detallada. Deberá entenderse, sin embargo, que la descripción detallada y los ejemplos específicos, indicando a la vez las realizaciones específicas, se proporcionan con fines meramente ilustrativos.

Incorporación específica de sitio de pirrolisina y PCL generadas biosintéticamente.

La pirrolisina (PYL) es el aminoácido 22^o natural codificado genéticamente encontrado en determinadas Archaea metanogénicas de la familia Methanosarcocinae y dos especies bacterianas no relacionadas. Específicamente, la pirrolisina se encuentra en MtmB1, la monometilamina (MMA) metiltransferasa que inicia la formación del metano en dichas bacterias Archaea, (véase Srinivasan, G., James, C. M., y Krzycki, J. A., (2002), "Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA," *Science*, 296, 1459-62; Soares, J. A., Zhang, L., Pitsch, R. L., Kleinholz, N. M., Jones, R. B., Wolff, J. J., Amster, J., Green-Church, K. B., y Krzycki, J. (2005). "The residue mass of L-pyrrolysine in three distinct methylamine methyltransferases," *Journal of Biological Chemistry*, 280, 36962-9; Hao, B., Gong, W., Ferguson, T. K., James, C. M., Krzycki, J. A., y Chan, M. K., (2002), "A new UAG-encoded residue in the structure of a methanogen methyltransferase", *Science*, 296, 1462-6; Krzycki, J. A., (2005), "The direct genetic encoding of pyrrolysine," *Current Opinion in Microbiology*, 8, 706-12; Krzycki, J. A., (2004), "Function of genetically encoded pyrrolysine in corrinoid- dependent methylamine methyltransferases," *Current Opinion in Chemical Biology*, 8, 484-91, y Ambrogelly, A., Palioura, S., y Soll, D., (2007), "Natural expansion of the genetic code," *Nature Chemical Biology* 3, 29-35). Pyrrolysine is considered a dipeptide wherein the α -amine of lysine is linked to the D-isomer of 4-methyl-pyrroline-5-carboxylate via an amide bond (see, Polcarpo, C. R., Herring, S., Berube, A., Wood, J. L., Soll, D., y Ambrogelly, A., (2006), "Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase," *FEBS Letters*, 580, 6695-700). La estructura de la pirrolisina (Figura 1) se dedujo de la estructura cristalina de MtmB1 t de la masa residual (véase, *J. Biol. Chem.* 2005, 44, 36962-36969; PNAS, 2007, 104, 1021-1026).

El gen mtmB1 que codifica MtmB 1 tiene un codón ámbar en marco (TAG) codon, que es habitualmente un codón de detención en toda regla. Sin embargo, en el ARNm de mtmB1, el codón UAG, codificado como TAG en el ADN, no finaliza la traducción durante la producción de la proteína MtmB 1, sino que en su lugar, el codón UAG codifica pirrolisina que se incorpora a la proteína. La pirrolisina se sintetiza de manera endógena y a continuación se incorpora mediante traducción simultánea como codones UAG en el marco como el aminoácido libre.

La biosíntesis e incorporación de pirrolisina se facilita por los genes naturales *pylT*, *pylS*, *pylB*, *pylC* y *pylD*. *pylT* codifica el ARNt de pirrolisilo, *pylS* codifica la sintetasa del ARNt de pirrolisilo mientras que *pylB*, *pylC* y *pylD* codifican proteínas necesarias para la biosíntesis de pirrolisina. Estos genes se han derivado de *Methanosarcina mazei*, (véase, Longstaff, D. G., Larue, R. C., Faust, J. E., Mahapatra, A., Zhang, L., Green-Church, K. B., y Krzycki, J. A., (2007), "A natural genetic code expansion cassette enables transmissible biosynthesis and genetic encoding of pyrrolysine," Proceedings of the National Academy of Sciences of the United States of America, 104, 1021-6, y Namy, O., Zhou, Y., Gundllapalli, S., Polycarpo, C. R., Denise, A., Rousset, J. P., Soll, D., y Ambrogelly, A., (2007), "Adding pyrrolysine to the *Escherichia coli* genetic code," FEBS Letters, 581, 5282-8). Los genes *pylT* y *pylS* junto con los genes *pylB*, *pylC* y *pylD* forman una agrupación de genes *pylTSBCD* que es un casete de expansión del código genético natural cuya transferencia permite al codón UAG traducirse como pirrolisina, un aminoácido libre sintetizado de forma endógena, que se incorpora a una proteína en el sitio UAG.

Se ha sugerido que la biosíntesis de pirrolisina se ve facilitada por los productos génicos de los genes naturales *pylB*, *pylC* y *pylD*, donde se ha propuesto a D-ornitina como precursor (véase, Namy, O., Zhou, Y., Gundllapalli, S., Polycarpo, C. R., Denise, A., Rousset, J. P., Soll, D., y Ambrogelly, A., (2007), "Adding pyrrolysine to the *Escherichia coli* genetic code," FEBS Letters, 581, 5282-8). Aunque se han propuesto varios precursores para la pirrolisina, tal como D-glutamato, D- isoleucina, D-prolina y D-ornitina, se ha indicado que D-ornitina es el precursor más eficaz de la biosíntesis de pirrolisina en *Escherichia coli* transformada con un plásmido que transporta los genes naturales *pylT*, *pylS*, *pylB*, *pylC* y *pylD*. Este estudio utilizó la lectura en la incorporación de TAG o la señal de truncado, es decir, la producción de la proteína de longitud completa, como demostración de que la pirrolisina se biosintetiza y se incorpora a las proteínas producidas en *Escherichia coli* transformada con un plásmido que transporta los genes naturales *pylT*, *pylS*, *pylB*, *pylC* y *pylD* y usando D-ornitina como precursor. La incorporación de pirrolisina no se verificó directamente por espectrometría de masas, pero la conclusión viene respaldada por datos de espectrometría de masas anteriores que, en ausencia de adición de D-ornitina, la pirrolisina se biosintetiza y se incorpora, si bien a niveles bajos, a proteínas producidas en *Escherichia coli* transformada con un plásmido que transporta los genes naturales *pylT*, *pylS*, *pylB*, *pylC* y *pylD* (véase, Longstaff, D. G., Larue, R. C., Faust, J. E., Mahapatra, A., Zhang, L., Green-Church, K. B., y Krzycki, J. A., (2007), "A natural genetic code expansion cassette enables transmissible biosynthesis and genetic encoding of pyrrolysine," Proceedings of the National Academy of Sciences of the United States of America, 104, 1021-6).

Sin embargo, tal como se proporciona en el presente documento, se ha descubierto que la introducción de los genes *pylT*, *pylS*, *pylB*, *pylC* y *pylD* en *Escherichia coli* o células de mamíferos, y la adición de D-ornitina en el medio de crecimiento dio como resultado la biosíntesis y la incorporación de una "pirrolisina desmetilada" (Figura 1, PCL-A y PCL-B), tal como se identifica mediante espectrometría de masas. Esta observación es sorprendente y no se podía predecir a la vista de los resultados presentados por Longstaff y col. Así, se proporciona en el presente documento un análogo de pirrolisina, pirrolina-carboxi-lisina (PCL) que se codifica de forma natural, generado de manera biosintética e incorporado a las proteínas mediante el uso de los genes naturales *pylT*, *pylS*, *pylB*, *pylC* y *pylD*, y D-ornitina como precursor. En otras realizaciones, como la D-arginina es un precursor de la D-ornitina, también se proporciona en el presente documento un análogo de pirrolisina que se codifica de forma natural, generado de manera biosintética e incorporado a las proteínas mediante el uso de los genes naturales *pylT*, *pylS*, *pylB*, *pylC* y *pylD*, y D-arginina como precursor.

La formación del anillo de pirrolina de la pirrolisina a partir de la D-ornitina se creía inicialmente que era análogo a la formación de la prolina a partir de L-ornitina (véase, Longstaff, D. G., Larue, R. C., Faust, J. E., Mahapatra, A., Zhang, L., Green-Church, K. B., y Krzycki, J. A., (2007). "A natural genetic code expansion cassette enables transmissible biosynthesis and genetic encoding of pyrrolysine," Proceedings of the National Academy of Sciences of the United States of America, 104, 1021-6). Numerosas bacterias convierten L-ornitina en L-prolina mediante el intermedio 1-pirrolina-5-carboxilato. La formación de 1-pirrolina-5-carboxilato a partir de ornitina es el primer paso de un proceso en dos etapas para la biosíntesis de L-prolina. La formación de 1-pirrolina-5-carboxilato a partir de L-ornitina puede suceder mediante dos rutas; la primera cuando la formación de 1-pirrolina-5-carboxilato es el resultado de la ciclodeaminación de L-ornitina, y una segunda mediante la formación de L-glutamato semialdehído mediante L-ornitina aminotransferasa. El glutamato semialdehído forma a continuación 1-pirrolina-5-carboxilato. La reducción de 1-pirrolina-5-carboxilato mediante pirrolina-5-carboxilato reductasa da como resultado la formación de L-prolina.

En la actualidad no se ha identificado un homólogo de la L-ornitina aminotransferasa en la formación de pirrolisina en las archaea metanogénicas. Sin embargo, se anticipa que implica una enzima análoga. Se cree que la biosíntesis de pirrolisina en un proceso concertado que utiliza un precursor celular y la acción concertada de los productos de los genes *pylB*, *pylC* y *pylD*, donde el esquema propuesto para la biosíntesis de la pirrolisina a partir de D-prolina mediante 1-pirrolino-5-carboxilato se muestra en la Figura 2A (véase, Longstaff, D. G., Larue, R. C., Faust, J. E., Mahapatra, A., Zhang, L., Green-Church, K. B., y Krzycki, J. A., (2007), "A natural genetic code expansion cassette enables transmissible biosynthesis and genetic encoding of pyrrolysine," Proceedings of the National Academy of Sciences of the United States of America, 104, 1021-6). Los productos génicos de *pylB*, *pylC* y *pylD* son miembros de varias familias de proteínas. *PyIB* tiene restos de firma del radical Fe-S de la familia de enzimas SAM conocidas por medical reacciones catalizadas por radicales. Asimismo, *PyIB* comparte cierta homología de secuencia con la biotina sinatasa, por tanto se ha sugerido que *PyIB* está implicado en la formación o metilación del anillo de pirrolina de la pirrolisina. El producto *PyID* del gen *pylD* tiene el dominio de unión a NADH de varias familias de deshidrogenasas, lo que sugiere

que PylD está implicado en la formación del enlace imina de 1-pirrolina. PylC es similar a carbamoilfosfato sintetasa y D-alanil-D-alanina ligasas, lo que sugiere un papel en la formación del enlace amida de pirrolisina entre la lisina y el isómero D de 4-metil-1-pirrolina-5-carboxilato. De este modo, los posibles productos génicos de pylB, pylD y pylD tienen la similitud respectiva con las proteínas SAM radicalarias, proteínas formadoras de aminas y aminoácido deshidrogenasas, y por tanto se cree que participan en rutas similares en la biosíntesis de pirrolisina.

Sin embargo, tal como se proporciona en el presente documento, los intentos de biosintetizar pirrolisina en *Escherichia coli* y células HEK293F usando D-ornitina como precursor no dio como resultado la formación de pirrolisina, sino en su lugar de una "pirrolisina desmetilada" que se denomina en el presente documento como pirrolina-carboxilisina (PCL) (PCL-A o PCL-B: véase la Figura 1). Tal como se indica en el presente documento, la biosíntesis de PCL-A o PCL-B no requiere la presencia del gen pylB y por tanto un esquema posible para la biosíntesis de PCL-A o PCL-B se muestra en la Figura 2B. Sin desear quedar vinculado por teoría alguna, esta posible ruta implica la conversión de D-ornitina en ácido 5-amino-2-oxopentanoico mediante una D-aminoácido oxidasa endógena (E.C. 1.4.3.3), y la ciclación espontánea del ácido 5-amino-2-oxopentanoico a 1-pirrolina-2-carboxilato y agua. Este precursor, que está disponible para la mayoría de organismos, se puede convertir en D-1-pirrolina-5-carboxilato mediante reordenación del doble enlace, posiblemente ayudado por la enzima PylD. La unión de D-1-pirrolina-5-carboxilato a la amina épsilon de la L-lisina mediante PylC y ATP daría como resultado la pirrolina-carboxi-lisina (PCL: PCL-A). Alternativamente, la unión de 1-pirrolina-2-carboxilato daría como resultado PCL-B que tiene una masa molecular igual a la de PCL-A. La observación de que tanto PylD como PylC son necesarios para incorporación de PCL y la incapacidad de PylS para aceptar como sustrato análogos de pirrolisina con un carbono sp² en posiciones equivalentes a la posición C-5 position del anillo de 1-pirrolina de PCL (como se presenta en el presente documento), sugirió inicialmente que PCL se incorporaría a las proteínas principalmente en forma de PCL-A. Sin desear quedar vinculado por teoría alguna, se cree que la PCL-A o PCL-B desmetiladas se obtienen en presencia de la D-ornitina añadida como resultado bien de la desactivación de PylB, la ausencia de un sustrato PylB donante de metilo, la ausencia del cofactor o cofactores necesarios, o combinaciones de los mismos. Sin embargo, esto no impide la presencia de un mecanismo alternativo. De hecho, los ensayos de incorporación con varios intermedios (como se indica en el presente documento) sugiere una ruta biosintética diferente para PCL-A o PCL-B de lo que se sugiere en la Figura 2B. Esta ruta alternativa se indica en el presente documento.

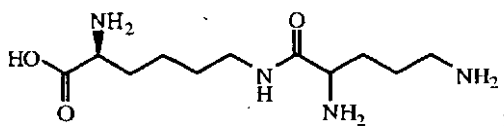
La incorporación de pirrolisina a una proteína en respuesta al codón UAG se ha demostrado que se facilita mediante los genes naturales pylT y pylS, donde la traducción de UAG como pirrolisina requiere la aminoacilación del ARNt^{pyl} decodificador de ámbar con pirrolisina mediante la pirrolisil-AENt sintetasa (PylS) (Figura 3A). El gen pylT codifica el supresor natural específico ARNt^{pyl} (PylT), cuyo anticodón CUA complementa el codón UAG de sentido directo correspondiente a la pirrolisina. El gen pylS codifica la pirrolisil-ARNt sintetasa (PylS), una sintetasa de ARNt de clase II específica que carga ARNt^{pyl} con pirrolisina (sintetizada de forma química o bioquímica) y también lleva a cabo reacciones de intercambio ATP:pirofosfato dependientes de pirrolisina. De igual forma, la incorporación de los análogos de pirrolisina PCL-A o PCL-B a una proteína en respuesta al codón UAG se cree que se facilita mediante los genes naturales pylT y pylS (Figura 3B).

Biosíntesis e incorporación específica del sitio de pirrolisina y PCL a proteínas expresadas por células procariontas y eucariotas

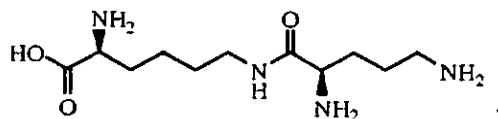
Se proporcionan en el presente documento métodos para la incorporación específica del sitio de pirrolisina y/o pirrolina-carboxi-lisina (ácido fSJ-2-amino-6-(3,4-dihidro-2H-pirrol-2-carboxamido) hexanoico (PCL-A) o ácido (S)-2-amino-6-(3,4-dihidro-2H-pirrol-5-carboxamido)hexanoico (PCL-B)) generados biosintéticamente. Los análogos de pirrolisina PCL-A y PCL-B, denominados conjuntamente en el presente documento como PCL, carecen del grupo metilo de la pirrolisina (PYL) (Figura 1). En determinadas realizaciones de dichos métodos, la célula eucariota es una célula de mamífero, una célula de levadura, una célula de insecto, una célula fúngica o una célula vegetal. En otras realizaciones, las células de mamífero utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células de riñón embrionario humano (HEK293F), células de carcinoma epiteloide humano (HeLa y GH3), células de riñón de mono (COS), células C6 de glioma de rata, células de riñón de cría de hámster (BHK-21) y células de ovario de hámster chino (CHO). En determinadas realizaciones, las células de levadura utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células de *Saccharomyces cerevisiae* y *Pichia pastoris*. En otras realizaciones, las células de insecto utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células de *Spodoptera frugiperda* (sf9 y sf21), células de *Trichoplusia ni* (BTI TN-5B1-4 o High-Five™) y células de *Mammestra brassicae*. En determinadas realizaciones, la célula procarionta es una bacteria, mientras que en otras realizaciones, la bacteria utilizada en los métodos proporcionados en el presente documento incluye, pero no se limitan a, *Escherichia coli*, *Mycobacterium smegmatis*, *Lactococcus lactis* y *Bacillus subtilis*.

En determinadas realizaciones, dichos métodos para la incorporación específica del sitio de la pirrolisina y PCL generados biosintéticamente implican introducir los genes pylT, pylS, pylB, pylC y pylD, y el gen de la proteína deseada, en células procariontas y/o células eucariotas, y opcionalmente añadir un precursor de pirrolisina o PCL a los medios de crecimiento de las células transfectedas. En determinadas realizaciones, el precursor es D-ornitina, mientras que en otras realizaciones el precursor es L-ornitina. En determinadas realizaciones, el precursor es D,L-ornitina. En determinadas realizaciones, el precursor es D-arginina, mientras que en otras realizaciones el

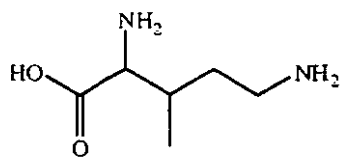
precursor es L-arginina. En determinadas realizaciones, el precursor es D,L-arginina. En determinadas realizaciones, el precursor es



5 ácido (2S)-2-amino-6-(2,5-diaminopentanamido)hexanoico. En determinadas realizaciones, el precursor es



10 ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico. En determinadas realizaciones, el precursor es



15 ácido 2,5-diamino-3-metilpentanoico. En determinadas realizaciones, el precursor es el ácido (2R,3S)-2,5-diamino-3-metilpentanoico. En determinadas realizaciones, la célula eucariota es una célula de mamífero, una célula de levadura, una célula de insecto, una célula fúngica o una célula vegetal. En otras realizaciones, las células de mamífero utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células HEK293F de riñón embrionario humano, células HeLa y GH3 de carcinoma epitelioide humano, células COS de riñón de mono, células C6 de glioma de rata, células BHK-21 de riñón de cría de hámster y células CHO de ovario de hámster chino. En determinadas realizaciones, las células de levadura utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células de *Saccharomyces cerevisiae* y *Pichia pastoris*. En otras realizaciones, las células de insecto utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células de *Spodoptera frugiperda* sf9 y sf21, células de *Trichoplusia ni* (BTI TN-5B1-4 o High-Five™) y células de *Mamestra brassicae*. En determinadas realizaciones, la célula procarionta es una bacteria, mientras que en otras realizaciones, la bacteria utilizada en los métodos proporcionados en el presente documento incluye, pero no se limitan a, *Escherichia coli*, *Mycobacterium smegmatis*, *Lactococcus lactis* y *Bacillus subtilis*.

En determinadas realizaciones, dichos métodos para la incorporación específica del sitio de la pirrolisina y PCL generados biosintéticamente implican introducir los genes *pylT*, *pylS*, *pylB*, *pylC* y *pylD*, y el gen de la proteína deseada, en células procariontas y/o células eucariotas, y añadir un precursor de la pirrolisina o la PCL al medio de crecimiento de las células transfectadas. En determinadas realizaciones, el precursor es D-ornitina, mientras que en otras realizaciones el precursor es L-ornitina. En determinadas realizaciones, el precursor es D-arginina, mientras que en otras realizaciones el precursor es L-arginina. En determinadas realizaciones, el precursor es D,L-arginina. En determinadas realizaciones, el precursor es el ácido (2S)-2-amino-6-(2,5-diaminopentanamido)hexanoico. En determinadas realizaciones, el precursor es el ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico. En determinadas realizaciones, el precursor es el ácido 2,5-amino-3-metilpentanoico. En determinadas realizaciones, el precursor es el ácido (2R,3S)-2,5-diamino-3-metilpentanoico. En determinadas realizaciones, la célula eucariota es una célula de mamífero, una célula de levadura, una célula de insecto, una célula fúngica o una célula vegetal. En otras realizaciones, las células de mamífero utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células HEK293F de riñón embrionario humano, células HeLa y GH3 de carcinoma epitelioide humano, células COS de riñón de mono, células C6 de glioma de rata, células BHK-21 de riñón de cría de hámster y células CHO de ovario de hámster chino. En determinadas realizaciones, las células de levadura utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células de *Saccharomyces cerevisiae* y *Pichia pastoris*. En otras realizaciones, las células de insecto utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células de *Spodoptera frugiperda* sf9 y sf21, células de *Trichoplusia ni* (BTI TN-5B1-4 o High-Five™) y células de *Mamestra brassicae*. En determinadas realizaciones, la célula procarionta es una bacteria, mientras que en otras realizaciones, la bacteria utilizada en los métodos proporcionados en el presente documento incluye, pero no se limitan a, *Escherichia coli*, *Mycobacterium smegmatis*, *Lactococcus lactis* y *Bacillus subtilis*.

50

En determinadas realizaciones, dichos métodos para la incorporación específica al sitio de la PCL generada biosintéticamente implica introducir los genes *pylT*, *pylS*, *pylC* y *pylD*, y el gen de la proteína deseada, en células procariontas y/o células eucariotas, y añadir D-ornitina al medio de crecimiento como un precursor de la PCL. En determinadas realizaciones, la célula eucariota es una célula de mamífero, una célula de levadura, una célula de insecto, una célula fúngica o una célula vegetal. En otras realizaciones, las células de mamífero utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células HEK293F de riñón embrionario humano, células HeLa y GH3 de carcinoma epitelioide humano, células COS de riñón de mono, células C6 de glioma de rata, células BHK-21 de riñón de cría de hámster y células CHO de ovario de hámster chino. En determinadas realizaciones, las células de levadura utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células de *Saccharomyces cerevisiae* y *Pichia pastoris*. En otras realizaciones, las células de insecto utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células de *Spodoptera frugiperda* sf9 y sf21, células de *Trichoplusia ni* (BTI TN-5B1-4 o High-Five™) y células de *Mamestra brassicae*. En determinadas realizaciones, la célula procarionta es una bacteria, mientras que en otras realizaciones, la bacteria utilizada en los métodos proporcionados en el presente documento incluye, pero no se limitan a, *Escherichia coli*, *Mycobacterium smegmatis*, *Lactococcus lactis* y *Bacillus subtilis*. En dichas realizaciones donde se utilizan los genes naturales *pylT*, *pylS*, *pylC* y *pylD*, más que generar e incorporar biosintéticamente pirrolisina, en su lugar se genera e incorpora biosintéticamente la PCL análoga de la pirrolisina desmetilada.

En determinadas realizaciones, dichos métodos para la incorporación específica al sitio de la pirrolisina y la PCL generada biosintéticamente implica introducir los genes *pylT*, *pylS*, *pylC* y *pylD*, y el gen de la proteína deseada, en células procariontas y/o células eucariotas, y añadir D-ornitina, L-ornitina, D,L-ornitina, D-arginina, L-arginina, D,L-arginina, ácido (2S)-2-amino-6-(2,5-diaminopentanamido)hexanoico, ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico o ácido 2,5-diamino-3-metilpentanoico o ácido (2R,3R)-2,5-diamino-3-metilpentanoico al medio de crecimiento como un precursor de la pirrolisina y/o la PCL. En determinadas realizaciones, la célula eucariota es una célula de mamífero, una célula de levadura, una célula de insecto, una célula fúngica o una célula vegetal. En otras realizaciones, las células de mamífero utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células HEK293F de riñón embrionario humano, células HeLa y GH3 de carcinoma epitelioide humano, células COS de riñón de mono, células C6 de glioma de rata, células BHK-21 de riñón de cría de hámster y células CHO de ovario de hámster chino. En determinadas realizaciones, las células de levadura utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células de *Saccharomyces cerevisiae* y *Pichia pastoris*. En otras realizaciones, las células de insecto utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células de *Spodoptera frugiperda* sf9 y sf21, células de *Trichoplusia ni* (BTI TN-5B1-4 or High-Five™) y células de *Mamestra brassicae*. En determinadas realizaciones, la célula procarionta es una bacteria, mientras que en otras realizaciones, la bacteria utilizada en los métodos proporcionados en el presente documento incluye, pero no se limitan a, *Escherichia coli*, *Mycobacterium smegmatis*, *Lactococcus lactis* y *Bacillus subtilis*. En dichas realizaciones donde se utilizan los genes naturales *pylT*, *pylS*, *pylC* y *pylD*, más que generar e incorporar biosintéticamente pirrolisina, en su lugar se genera e incorpora biosintéticamente la PCL análoga de la pirrolisina desmetilada.

En determinadas realizaciones, dichos métodos para la incorporación específica al sitio de la pirrolisina generada biosintéticamente implican introducir los genes *pylT*, *pylS*, *pylC* y *pylD*, y el gen de la proteína deseada, en células procariontas y/o células eucariotas, y añadir ácido 2,5-diamino-3-metilpentanoico o ácido (2R,3R)-2,5-diamino-3-metilpentanoico al medio de crecimiento como un precursor de la pirrolisina. En determinadas realizaciones, la célula eucariota es una célula de mamífero, una célula de levadura, una célula de insecto, una célula fúngica o una célula vegetal. En otras realizaciones, las células de mamífero utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células HEK293F de riñón embrionario humano, células HeLa y GH3 de carcinoma epitelioide humano, células COS de riñón de mono, células C6 de glioma de rata, células BHK-21 de riñón de cría de hámster y células CHO de ovario de hámster chino. En determinadas realizaciones, las células de levadura utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células de *Saccharomyces cerevisiae* y *Pichia pastoris*. En otras realizaciones, las células de insecto utilizadas en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, células de *Spodoptera frugiperda* sf9 y sf21, células de *Trichoplusia ni* (BTI TN-5B1-4 o High-Five™) y células de *Mamestra brassicae*. En determinadas realizaciones, la célula procarionta es una bacteria, mientras que en otras realizaciones, la bacteria utilizada en los métodos proporcionados en el presente documento incluye, pero no se limitan a, *Escherichia coli*, *Mycobacterium smegmatis*, *Lactococcus lactis* y *Bacillus subtilis*.

La incorporación específica del sitio de la PCL generada biosintéticamente en los sitios codificados para TAG en una proteína modelo (hRBP4) se ha llevado a cabo en células de mamíferos utilizando los genes naturales *pylT*, *pylS*, *pylB*, *pylC* y *pylD* y con adición de D-ornitina al medio de crecimiento como precursor (véase el Ejemplo 2). En una realización, la incorporación específica del sitio de la PCL en la proteína modelo, hRBP4, se consiguió transfectando simultáneamente células HEK293F con ADN que codifica el ARNt *PyIT* que reconoce UAG, su aminoacil ARNt sintetasa *PyIS* específica, los genes biosintéticos *pylB*, *pylC* y *pylD* así como el ADN que codifica la proteína diana con el sitio de incorporación codificado por TAG. En la Figura 4A se muestran las construcciones génicas utilizadas para dicha biosíntesis in vivo y la incorporación en células de mamífero en la proteína modelo hRBP4 en sitios únicos

especificados por codones TAG. En la Figura 4B se muestran otras construcciones génicas para su uso en mamíferos para la expresión de diversos mutantes TAG de sitio único de hRBP4, mEPO, hEPO, y el dominio Fc de mIgG1. En la Figura 5 se muestra un plásmido que transporta *pylT*, *pylS*, *pylB*, *pylC* y *pylD* para la incorporación de PCL derivada biosintéticamente o pirrolisina en células de *Escherichia coli*.

5 Cuando se añade D-ornitina, el presunto precursor de la biosíntesis de la pirrolisina, al medio de cultivo de células HEK293F transfectadas con *pylT*, *pylS*, *pylB*, *pylC* y *pylD* así como el ADN que codifica la proteína diana, PCL se incorpora a hRBP4 en el sitio del codón TAG, en lugar de la pirrolisina (véase el Ejemplo 2, Figura 6A). La eficacia de incorporación varía para los diferentes sitios de incorporación como se observa en la Figura 6A.

10 La incorporación específica del sitio de la PCL generada biosintéticamente en los sitios codificados para TAG en una proteína modelo (hRBP4) se ha llevado a cabo en células de mamíferos utilizando los genes naturales *pylT*, *pylS*, *pylC* y *pylD* y adición de D-ornitina al medio de crecimiento como precursor (véase el Ejemplo 2). En una realización, la incorporación específica del sitio de la PCL en la proteína modelo, hRBP4, se consiguió transfectando simultáneamente células HEK293F con ADN que codifica el ARNt de *PylT* que reconoce UAG, su aminoacil ARNt sintetasa *PylS* específica, los genes biosintéticos *pylC* y *pylD* así como el ADN que codifica la proteína diana con el sitio de incorporación codificado por TAG. Cuando se añade D-ornitina, el presunto precursor de la biosíntesis de la pirrolisina, al medio de cultivo de células HEK293F transfectadas, PCL se incorpora a hRBP4 en el sitio del codón TAG, en lugar de la pirrolisina. La Figura 6B muestra el SDS-PAGE y la Figura 6C muestra el espectro de masas de hRBP4 Phe62PCL purificada (mutante nº 2) producido en células HEK293F en ausencia (B, hilera 1) o presencia (B, hilera 2) de D-ornitina. La flecha indica la hRBP4 de longitud completa y la masa obtenida fue 23166,0 Da, cercana a la masa esperada de 23168 Da. Los datos de espectrometría de masas que se muestran en las Figuras 7-9 ilustran adicionalmente que la PCL, y no la pirrolisina, se incorpora en el sitio TAG, en el resto 62 de hRBP4 (véase el Ejemplo 2).

25 La Figura 1 muestra las estructuras de la pirrolisina () y del análogo de pirrolisina desmetilada, pirrolina-carboxi-lisina (PCL) (estructura PCL-A o la estructura alternativa PCL-B). La estructura de PCL-A (o la estructura alternativa PCL-B) se distinguieron de la pirrolisina con el uso del análisis por espectrometría de masas de alta precisión de un fragmento de péptido de un sitio de incorporación (Figuras 7-9). Además, se distinguió la estructura de la PCL procedente de la pirrolisina mediante la detección de la PCL por espectrometría de masas como aminoácido libre en el lisado celular (Figura 10, véase el Ejemplo 3). Asimismo, la observación de la PCL en el lisado celular demuestra que la PCL se generó biosintéticamente como aminoácido libre, en lugar de formarse mediante una modificación posterior a la traducción. Las Figuras 10A y 10B muestran la detección de la PCL, en el lisado de células HEK293F biosintetizadas a partir de D-ornitina. La Figura 10A, es la traza de la HPLC del lisado de células transfectadas con los genes biosintéticos *pylB*, *pylC* y *pylD* y que crecen en presencia de D-ornitina (traza de la parte inferior) y en ausencia de D-ornitina (traza de la parte superior). El cromatograma del lisado obtenido de células que crecen en presencia de D-ornitina muestra un pico a un tiempo de elución de 4,13 min (marcado con un asterisco) que está ausente en el lisado de células idénticas cultivadas en ausencia de D-ornitina. La Figura 10C es el espectro de masas de barrido completo del pico de HPLC a los 4,13 min, donde la masa obtenida ($m+1$) es consistente con la masa teórica para PLC. La Figura 10B es un cromatograma de HPLC que muestra que la lisina es igualmente abundante en ambas muestras y un espectro de masas completo ($m+1$) de la lisina de HPLC a los 1,44 min (Figura 10D) ilustra la precisión del método.

45 La Figura 11 muestra la incorporación de N-s-ciclopentiloxicarbonil-L-lisina (CYC) a varias proteínas hRBP4 TAG mutantes en células HEK293F. La Figura 12 muestra la verificación espectrométrica de masas de la incorporación de CYC en el sitio TAG de la Phe62TAG mutante de hRBP4. La Figura 13A y la Figura 13B (véase el Ejemplo 4) muestra la incorporación de PCL en función de diversos precursores relacionados en la tabla y mostrados en la Figura 14A. La Figura 13A y la Figura 13B muestran también la incorporación directa de diversos análogos de pirrolisina (que incluyen CYC) en una proteína mutante hRBP4 TAG que utiliza células HEK293F. Para demostrar que D-ornitina es un precursor de la biosíntesis de PCL, se añadieron precursores potenciales de PCL (Figura 14A) al medio de crecimiento de las células HEK293F transfectadas con los genes de la pirrolisina biosintéticos *pylB*, *pylC* y *pylD*, la pareja *pylT/pylS* ARNt/pirrolisil-ARNt sintetasa y el ADN mutante de hRBP4 TAG. Además, los diversos análogos de pirrolisina, añadidos al medio de crecimiento de células HEK293F transfectadas solo con la pareja de *pylT/pylS* ARNt/aa-ARNt sintetasa y el ADN mutante de hRBP4 TAG, se muestran en la Figura 14B. La Figura 13A es una transferencia Western de la proteína hRBP4 de longitud completa con un anticuerpo dirigido contra la etiqueta His, mientras que la Figura 13B es un SDS-PAGE de las mismas muestras después de la purificación Ni-NTA. Tal como se muestra, la D-ornitina (banda 2) es el mejor precursor de la biosíntesis de PCL, aunque la D-arginina (banda 4), que se puede metabolizar a D-ornitina, proporciona una producción de proteína por encima del fondo. De los análogos de pirrolisina, solamente CYC da como resultado la producción de proteína a un nivel similar a la D-ornitina, mientras que el análogo del ácido 3-oxobutanoico TU3000-016 muestra una incorporación medible pero mucho más baja. En el Ejemplo 33 se proporciona la descripción de la síntesis de diversos análogos. Todas las bandas muestran una producción de bajo nivel de la proteína de longitud completa debido posiblemente niveles endógenos bajos de D-ornitina y de metabolitos o componentes del medio aceptables como sustratos *PylS*.

65 Se evaluó la incorporación de PCL utilizando diferentes combinaciones de los genes biosintéticos *pylB*, *pylC* y *pylD* (Figura 13C). Los datos muestran que solo los genes *pylC* y *pylD* son esenciales para la biosíntesis de la PCL y la

posterior incorporación de proteína. De forma más notable la D-prolina, así como otros análogos de D-prolina y ácido D-glutámico (Figura 13 y 14) no dieron como resultado la producción de proteína de longitud completa por encima del fondo. Esto sugiere que la biosíntesis de PCL-A y PCL-B no sigue la ruta sugerida en la Figura 2A. Sin embargo, los ensayos de incorporación de 3,4-dihidro-2H-pirrol-5-carboxilato (P2C) y 1-pirrolina-5-carboxilato (P5C) fracasaron también en producir proteína de longitud completa en *Escherichia coli*, mientras que los sintéticos PCL-A y PCL-B se incorporaron con eficacia elevada (Figura 15A, Ejemplo 15). Por tanto, la ruta biosintética sugerida en la Figura 2B tampoco es, probablemente, la ruta prominente.

La incorporación del precursor del ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico (denominado también Lys-Ns-D-orn) dio como resultado cantidades sustanciales de proteína de longitud completa (Figura 15A, Ejemplo 15). Además, se ha encontrado que la incorporación de ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico requiere solo la presencia de los genes *pylS*, *pylT* y *pylD* (Figura 15B, Ejemplo 15). En su conjunto, estas observaciones sugieren las rutas alternativas que se muestran en la Figura 16A y 16B. Estas rutas se basan en el concepto de que la D-ornitina, con la ayuda de *pylC*, un presunta D-alanil-D-alanina ligasa, se acopla en primer lugar al grupo amino épsilon de la L-lisina para formar el ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico. En la Figura 16A, el ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico intermedio se activa mediante una D-ornitina: oxígeno oxidoreductasa (EC 1.4.3.3) o D-amino-transaminasa (EC 2.6.1.21) dando como resultado la ciclación espontánea para formar PCL-B. Alternativamente, en la Figura 16B, el ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico intermedio se activa mediante una D-ornitina:2-oxoglutarato 5-transaminasa (EC 2.6.1.13) dando como resultado la ciclación espontánea para formar PCL-A. Las reacciones de ciclación son similares a las del D-glutamato 5-semialdehído. Cualquiera de las etapas de activación previas a la ciclación podría presuntamente ser catalizada por *pylD*, la segunda enzima requerida para la biosíntesis de PCL. Tal como se muestra en la Figura 16, se propone que la metilación de PCL-A por *pylB* completa la biosíntesis de PYL. Señalar que en el caso de la ruta en la Figura 16A, esto requeriría probablemente la presencia de otra enzima *Pyl*, aún por descubrir, que formara PCL-A a partir de PCL-B. Sin embargo, esta enzima adicional no se requeriría para la ruta alternativa que se muestra en la Figura 16B. La PCL-A formada tras la activación por *pylD* y la ciclación espontánea podría metilarse por *pylB* directamente para formar pirrolisina. Alternativamente, cualquiera de los intermedios o la D-ornitina podría ser sustrato para la metilación por *pylB* si las subsiguientes enzimas toleraran dicha modificación del sustrato.

Se ha demostrado que la pirrolisil-ARNt sintetasa *PylS* acepta algunos análogos de pirrolisina para la carga del ARNt^{Pyl}. Sin embargo, dichos estudios indican también la importancia del estereocentro C-5 del anillo de pirrolisina para el reconocimiento de la sintetasa *PylS*, cuando se requiere la conformación-D (véase, Polycarpo, C. R., Herring, S., Berube, A., Wood, J. L., Soll, D., y Ambrogelly, A., (2006), "Pyrrolysine analogues as substrates for pyrrolysyl-tRNA synthetase," *FEBS Letters*, 580, 6695-700). Tal como se proporciona en el presente documento, y consistente con esta interpretación, los intentos de incorporar diversos análogos de pirrolisina (Figura 13), incluyendo anillos aromáticos de cinco y seis miembros análogos de pirrolisina, demostraron que dichos anillos aromáticos análogos de cinco y seis miembros no son sustratos aceptables para *PylS* (Figura 14). Esto es posiblemente debido a la ausencia del centro quiral C-5, o, en determinados casos, debido al tamaño más grande de los análogos. Como se ha indicado, N-s-ciclopentiloxicarbonil-L-lisina (CYC), un sustrato conocido de *PylS*, y el análogo del ácido 3-oxobutanoico, TU3000-016, se incorporaron usando la pareja *pylT/pylS* ARNt/pirrolisil-ARNt sintetasa (Figura 14). Sin embargo, los análogos evaluados que tienen un tamaño y volumen similar al de la N-s-ciclopentiloxicarbonil-L-lisina (CYC) no se incorporaron, posiblemente debido a que carecían del estereocentro C-5. Por tanto, *PylS* parece ser selectivo para la quiralidad en el punto unión del anillo de pirrolina. Sin embargo, se incorporó también la PCL-B sintética con eficacia elevada sugiriendo que el carbono sp² acquiral en la posición C-5 no es en todos los casos una incorporación desfavorable. Actualmente, la baja reactividad de la PCL-B sintética con 2-ABA cuando se compara con la PCL-A sintética (Ejemplo 44, Figuras 56 y 58) y la ausencia de una enzima conocida para la conversión de PCL-B a PCL-A en la Figura 16A favorece la asignación del análogo de pirrolisina incorporado como PCL-A.

Tal como se proporciona en el presente documento, en determinadas realizaciones se ha encontrado que, cuando D-ornitina se añade al medio de crecimiento como precursor, el uso de los genes naturales *pylB*, *pylC* y *pylD* para la biosíntesis de la pirrolisina en células de mamíferos (Ejemplos 2) y células bacterianas (Ejemplos 8, 9 y 10) dio como resultado la generación de PCL y no de pirrolisina. Además, la evaluación de la biosíntesis de PCL utilizando cualquiera de los genes naturales *pylB* y *pylC*, los genes naturales *pylB* y *pylD* o los genes naturales *pylC* y *pylD*, dio como resultado la generación de PCL solamente cuando los productos génicos *pylC* y *pylD* estuvieron presentes. Esto sugiere que el producto génico de *pylB* no se necesita en la biosíntesis de PCL, cuando se añade D-ornitina al medio de crecimiento como precursor (Figura 13C, véase el Ejemplo 4). Esto se ve respaldado por la incorporación de la PCL en las tres proteínas modelo, hRBP4 (Figura 6, véase el Ejemplo 2), el dominio mlgG1 Fc (Figura 17A, véase el Ejemplo 5) y mEPO (Figura 17B, véase el Ejemplo 6), que se ha llevado a cabo sin transfección simultánea del gen *pylB* en células de mamíferos. En *Escherichia coli*, FAS-TE, FGF21 y FKBP son ejemplos donde el uso de los genes naturales *pylB*, *pylC* y *pylD*, dio como resultado la generación de PCL exclusivamente (Figuras 18B, 19 y 20, Ejemplos 8, 9 y 10).

La biosíntesis de PCL requiere la presencia de los genes biosintéticos *pylC* y *pylD*, pero no de *pylB*, a las células hospedadoras. En la biosíntesis de pirrolisina en *Methanosarcina*, se ha sugerido que *pylD* contiene el dominio de unión a NADH de las deshidrogenasas y genera por tanto D-1-pirrolina-5-carboxilato a partir de D-prolina. Sin

embargo, la adición de D-prolina al medio de crecimiento no da como resultado incorporación significativa de PCL (Figura 13). La adición de 3,4-dihidro-2H- pirrol-5-carboxilato (denominado también en el presente documento como 1-pirrolina-2-carboxilato; P2C) y D-1-pirrolina-5-carboxilato (P5C) al medio de crecimiento fracasa también en producir proteínas de longitud completa mientras que el ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico da como resultado proteína que contiene PCL. PylC tiene homología de secuencia con D-alanil-D-alanina ligasas y, en la biosíntesis de PCL o pirrolisina, catalizaría la unión de la D-ornitina al grupo amino-épsilon de la lisina para dar el ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico. Por tanto, se postula en el presente documento y sin pretender quedar vinculado por teoría alguna, que la biosíntesis de PCL a partir de D-ornitina en células de mamífero de *Escherichia coli* implica probablemente la conversión de la D-ornitina a ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico mediante PylC (Figura 16B). PylD está implicada probablemente en la activación del ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico en el semialdehído ((S)-2-amino-6-((R)-2-amino-5-oxopentanamido)hexanoico) que cicla espontáneamente a PCL-A tal como se muestra en la Figura 16B. Alternativamente, se podría formar espontáneamente PCL-B si PylD tiene actividad del tipo D-aminoácido transaminasa or D-ornitina: oxígeno oxidoreductasa (Figura 16A).

En otras realizaciones, la expresión de EGF de ratón y de TNF- α de ratón en *Escherichia coli* (Figura 21 y Ejemplos 13 y 14) dio como resultado mezclas de proteínas tanto con PCL como pirrolisina incorporada en el sitio TAG. La incorporación de pirrolisina fue dependiente de la presencia del gen *pylB*, mientras que se observaron preparaciones homogéneas de PCL que contenían proteínas en ausencia de *pylB* (Figura 21 y Ejemplos 11 y 12). Esto verifica por tanto experimentalmente *pylB* como la metiltransferasa en la biosíntesis de pirrolisina. Incluso en presencia del gen *pylB*, las cantidades relativas de PCL y pirrolisina que contienen proteínas variadas de fermentación a fermentación siendo normalmente más prominente la proteína PCL. Estas observaciones sugieren que la actividad metiltransferasa del sustrato donante de metilo o de los cofactores requeridos son limitantes de la biosíntesis eficaz de pirrolisina en *Escherichia coli* y células de mamíferos.

Además, el producto génico de *pylB* no se expresa eficazmente. Por lo tanto, en determinadas realizaciones, los genes *pylB* modificados se usan en la biosíntesis de pirrolisina u otros análogos de pirrolisina. Por tanto, se proporcionan en el presente documento métodos para la biosíntesis de pirrolisina, PCL y otros análogos de pirrolisina en *Escherichia coli*, células de mamífero y otras células hospedadoras donde uno o más de los genes *pylB*, *pylC* y *pylD* están modificados. Dichas modificaciones pueden incluir utilizar genes homólogos procedentes de otros organismos, incluyendo, pero sin limitarse a otras especies de *Methanosarcinae*, o genes mutados. En determinadas realizaciones, se usa la mutagénesis dirigida al emplazamiento, mientras que en otras realizaciones se usa la mutagénesis aleatoria combinada con la selección. Dichos métodos incluyen también la adición de ADN de la proteína deseada y la inclusión de genes *pylT* y *pylS* para incorporar la pirrolisina, la PCL o los análogos de pirrolisina en la proteína. En una realización determinada, se utilizan un gen *pylB* modificado, los genes *pylC*, *pylD*, *pylT* y *pylS* naturales para generar e incorporar biosintéticamente pirrolisina. En otras realizaciones, se utilizan los genes *pylB* y *pylC*, los genes *pylD*, *pylT* y *pylS* naturales para generar para generar e incorporar biosintéticamente un análogo de pirrolisina diferente de PCL. En otras realizaciones se utilizan los genes *pylB* and *pylD* modificados, los genes *pylC*, *pylT* y *pylS* naturales para generar para generar e incorporar biosintéticamente un análogo de pirrolisina diferente de PCL. En otras realizaciones, los genes *pylT*, *pylS*, *pylB*, *pylC* y *pylD* están modificados opcionalmente, para mejorar la incorporación de la pirrolisina, PCL u otros análogos de pirrolisina en proteínas.

Además, para determinadas realizaciones, la formación de intermedios en la biosíntesis de la pirrolisina, la PCL y/u otros análogos de pirrolisina o la biosíntesis de pirrolisina puede estar limitada por la función de las enzimas y proteínas hospedadoras. En determinadas realizaciones, la actividad o concentración baja de una o más enzimas hospedadoras puede ser limitante de la formación de los intermedios requeridos en la biosíntesis de la pirrolisina, la PCL u otros análogos de pirrolisina. En determinadas realizaciones, la actividad de las enzimas hospedadoras puede desviar los intermedios de la ruta que conduce a la pirrolisina, la PCL u otros análogos de pirrolisina a otras rutas metabólicas, o puede inhibir la formación de dichos intermedios. Por tanto, se proporcionan en el presente documento métodos para la biosíntesis de pirrolisina, PCL y otros análogos de pirrolisina en *Escherichia coli*, células de mamífero y otras células hospedadoras donde una o más enzimas hospedadoras están modificadas. Dichos métodos incluyen, pero no se limitan a, la expresión en exceso, la activación, la supresión o inhibición de dicha enzima hospedadora mediante medios genéticos o químicos, la adición del ADN que codifica dichas enzimas hospedadoras, la adición del ARN silenciador (ARNsi) para suprimir la traducción del ARNm, y la adición de cofactores para la formación de dichos intermedios a partir de D-ornitina.

La incorporación específica de sitio de la PCL generada biosintéticamente en los sitios codificados para TAG se ha llevado a cabo también en cuatro sitios del dominio Fc de la IgG 1 de ratón (véase el Ejemplo 5 y la Figura 17A) y once sitios de la eritropoyetina (EPO) (véase el Ejemplo 6 y la Figura 17B). Se consiguió la incorporación de PCL en ambos conjuntos de proteínas en células de mamíferos HEK293F transfectadas con los genes *pylT*, *pylS*, *pylC* y *pylD* naturales y usando la D-ornitina añadida al medio como precursor. La Figura 17A es el SDS-PAGE que muestra las cuatro proteínas mFc (FL) de longitud completa, donde la PCL se ha incorporado en cuatro sitios del dominio IgG 1 Fc utilizando células HEK293F transfectadas simultáneamente con las construcciones mutantes TAG de mFc, pCMVpylT, pCMVpylS, pCMVpylC y pCMVpylD (Figura 4A) y las células crecen en presencia de D-ornitina. La Figura 17B es el SDS-PAGE que muestra las cuatro proteínas mEPO (FL) de longitud completa, donde la PCL se ha incorporado en once sitios de EPO de ratón utilizando células HEK293F transfectadas simultáneamente con las construcciones

mutantes TAG de mEPO, pCMVpyIT, pCMVpyIS, pCMVpyIC y pCMVpyID y las células crecen en presencia de D-ornitina.

La incorporación específica de sitio de la PCL generada biosintéticamente en los sitios codificados para TAG se ha llevado a cabo también utilizando células de *Escherichia coli* (véase el Ejemplo 7). Se construyó un plásmido, pARA-pylSTBCD, que codifica pylB, pylC, pylD, pylS y pylIT (Figura 6A). Se llevó a cabo la incorporación de PCL-A (o pCL-B) en dos sitios del dominio de la tioesterasa de la ácido graso sintetasa humana (FAS-TE) (véanse los Ejemplos 8 y la Figura 18), un sitio de FKBP-12 (véanse los Ejemplos 9 y la Figura 19) y los sitios 20 del factor de crecimiento de fibroblastos 21 (FGF21) (véanse los ejemplos 10 y la Figura 20) transformando las células de *Escherichia coli* con pARA-pylSTBCD y un segundo plásmido de expresión con el gen para la proteína de interés y añadiendo D-ornitina al medio de crecimiento durante la expresión de la proteína.

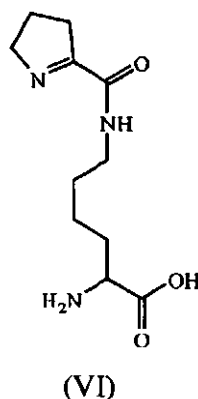
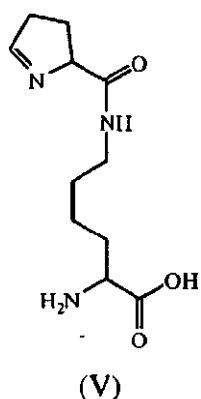
La Figura 18 muestra el SDS-PAGE y los espectros de masas para la incorporación de PCL en dos sitios en la tioesterasa de la ácido graso sintetasa humana (FAS-TE). Se expresó FAS-TE Tyr2454PCL, y las fracciones de proteína insoluble se purificaron mediante Ni-NTA. FAS-TE Leu2222PCL/Leu2223Ile se expresó con y sin D-ornitina en cultivos duplicados. Se muestran eluciones Ni-NTA en los geles. La masa obtenida en cada uno es consistente con la esperada. La Figura 19 muestra el SDS-PAGE y el espectro de masas para la incorporación de PCL en un sitio de FKBP-12. La masa obtenida (12085,6 Da) es consistente con la esperada (12084 Da) para la incorporación de PCL en un sitio único. Asimismo, se muestra en la Figura 19 un cristal de FKBP12-Ile90PCL. La Figura 20 muestra un SDS-PAGE que muestra la incorporación de PCL en sitios múltiples en FGF21.

Se encontró que la incorporación de pirrolisina (Pyl) y PCL era dependiente de la presencia del gen pylB en el sistema de expresión. La Figura 21A muestra el análisis de SDS-PAGE de la incorporación de Pyl o PCL a mTNF-a con el codón de Gln21 (CAA) mutado a un codón de detención TAG (Ejemplo 13). Las bandas 2 y 4 muestran niveles de expresión de la proteína similares en presencia y ausencia de pylB, respectivamente cuando la D-ornitina está presente. La banda 3 (pylB presente) y 5 (pylB ausente) muestra que no se expresa proteína en ausencia de D-ornitina.

Además, mEGF Tyr10TAG se expresó en presencia de los genes pylB, pylC, pylD, pylIT y pylS en *Escherichia coli* utilizando D-ornitina como precursor (Ejemplo 14). Los espectros de MS intactos (Figure 21C, parte inferior) sugirieron una mezcla de proteínas con Pyl y PCL incorporadas. El pico de MS del mEGF que contiene Pyl a 7309 Da es aproximadamente 6 veces más grande que el de la PCL que contiene mEGF a 7295 Da. La incorporación de Pyl y PCL en la posición TAG se verificó mediante los espectros de MS en tándem MS del péptido del extremo N MNSYPGCPSS(PCL)DGYCLNGGVCM (SEC ID NO:32) y MNSYPGCPSS(Pyl)DGYCLNGGVCM (SEC ID N:33) de la proteína mutante mEGF Tyr10TAG. La cuantificación de la abundancia relativa de masa del precursor de diferentes péptidos revela que Pyl es de 5 a 10 veces más abundante que PCL, lo que está de acuerdo aproximadamente con las medidas de masa intacta (Figura 21C parte inferior). Cuando mEGF Tyr10TAG se expresó en presencia de pylB, los espectros de MS intactos (Figura 21C, parte superior) sugirieron una única incorporación de la PCL.

De igual forma, cuando mTNF-a Gln21TAG se expresó en presencia de los genes pylB, pylC, pylD, pylIT y pylS en *Escherichia coli* utilizando los MS en tándem de la D-ornitina, se verificó la incorporación de Pyl y PCL en el péptido NH(Pyl)VEEQLEWL-SQR (SEC ID N:34) y NH(PCL)VEEQLEWLSQR (SEQ ID NO:35) de mTNF Gln21PCL. En contraste con mEGF Tyr10TAG, la cuantificación de la abundancia relativa de masa del precursor de diferentes péptidos identificados con MS en tándem revela que la PCL es 7 veces más abundante que Pyl en la proteína mTNF-a Gln21TAG. El mTNF-a Gln21TAG se expresó en ausencia del gen pylB, la cuantificación de la abundancia relativa de masa del precursor de diferentes péptidos en las medidas de MS en tándem muestra solo la proteína PCL en el intervalo dinámico del experimento. Estos experimentos indican claramente que la incorporación de Pyl depende estrictamente de la presencia de pylB, la presunta metiltransferasa en la biosíntesis de Pyl. Las relaciones relativas de PCL y Pyl incorporadas a las proteínas en presencia del gen pylB, parecen sin embargo variar de proteína a proteína y de experimento a experimento de fermentación y depende por tanto probablemente del tipo de vector de expresión, de las condiciones de crecimiento y/o de otras propiedades del cultivo de células hospedadoras.

Se proporcionan en el presente documento aminoácidos que tienen la estructura de la Fórmula (V) y la Fórmula (VI)



5 donde dichos compuestos de Fórmula (V) o Fórmula (VI) están generados biosintéticamente en una célula que comprende un gen *pylB*, un gen *pylC*, y un gen *pylD*, y la célula está en contacto con un medio de crecimiento que comprende un precursor. Dichos compuestos de Fórmula (V) o Fórmula (VI) se pueden generar biosintéticamente en una célula que comprende un gen *pylC* y un gen *pylD* y la célula está en contacto con un medio de crecimiento que comprende un precursor. En dicha biosíntesis, el precursor puede ser ornitina o arginina. En dicha biosíntesis, el precursor puede ser D-ornitina o D-arginina. En dicha biosíntesis, el precursor puede ser el ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico. En dicha biosíntesis, el precursor puede ser el ácido (2S)-2-amino-6-(2,5-diaminopentanamido)hexanoico.

15 Se proporcionan también en el presente documento células donde se biosintetizan los compuestos de Fórmula (V) y (VI) que comprenden además un gen *pylS* y un gen *pylT* y los compuestos de Fórmula (V) o Fórmula (VI) se incorporan a una proteína en el interior de la célula mediante una aminoacil ARNt sintetasa y un ARNt que reconoce al menos un codón selector de ARNm en la célula. Dicha aminoacil ARNt sintetasa es un producto génico del gen *pylS* y el ARNt es un producto génico del gen *pylT*. En determinadas realizaciones, dichos compuestos de Fórmula (V) o Fórmula (VI) se incorporan a una proteína en el interior de la célula mediante un ARNt ortogonal (ARNt-O) y una aminoacil ARNt sintetasa ortogonal (O-RS), donde los aminoacilatos O-RS del ARNt-O con el compuesto de Fórmula (V) o Fórmula (VI) y el ARNt-O reconocen al menos un codón selector de un ARNm en la célula.

20 El codón selector para la incorporación de la pirrolisina, la PCL y otros análogos de pirrolisina, que incluyen compuestos de Fórmula (V) y Fórmula (VI), está en un codón ámbar (TAG).

25 Las células usadas para la biosíntesis y/o la incorporación de la pirrolisina, la PCL y otros análogos de pirrolisina, que incluyen compuestos de Fórmula (V) y Fórmula (VI), son tanto células procariontas como células eucariotas. En determinadas realizaciones, las células procariontas incluyen, pero no se limitan a, *Escherichia coli*, *Mycobacterium smegmatis*, células de *Lactococcus lactis* y *Bacillus subtilis*. En otras realizaciones, las células eucariotas incluyen, pero no se limitan a, células de mamíferos, células de levaduras, células fúngicas, células vegetales o células de insectos. Dichas células de mamífero incluyen, pero no se limitan a, células de riñón embrionario humano (HEK293F), células de carcinoma epiteloide humano (HeLa y GH3), células de riñón de mono (COS), células C6 de glioma de rata, células de riñón de cría de hámster (BHK-21) y células de ovario de hámster chino (CHO). En determinadas realizaciones, las células de levadura incluyen, pero no se limitan a, células de *Saccharomyces cerevisiae* y *Pichia pastoris*. En otras realizaciones, las células de insecto incluyen, pero no se limitan a, células de *Spodoptera frugiperda* sf9 y sf21, células de *Trichoplusia ni* (BTI TN-5B1-4 o High-Five™) y células de *Mamestra brassicae*.

35 En otro aspecto proporcionado en el presente documento, la pirrolisina, la PCL y otros análogos de pirrolisina, que incluyen compuestos de Fórmula (V) y Fórmula (VI) se biosintetizan en células alimentadoras en contacto con un medio de crecimiento que comprende un precursor, y las células alimentadoras contienen un gen *pylB*, un gen *pylC* y un gen *pylD*. En otras realizaciones, la pirrolisina, la PCL y análogos de pirrolisina, que incluyen compuestos de Fórmula (V) y Fórmula (VI), se biosintetizan en células alimentadoras en contacto con un medio de crecimiento que comprende un precursor, y las células alimentadoras contienen un gen *pylC* y un gen *pylD*. Dicha pirrolisina biosintetizada, la PCL y otros análogos de pirrolisina, que incluyen compuestos de Fórmula (V) y Fórmula (VI), secretados a partir de las células alimentadoras en el medio de crecimiento se capturan por una segunda célula y posteriormente se incorporan a una proteína sintetizada en la segunda célula. Dichas segundas células contienen un gen *pylS* y un gen *pylT*. La pirrolisina, la PCL y otros análogos de pirrolisina, que incluyen compuestos de Fórmula (V) y Fórmula (VI), se incorporan a la proteína de la segunda célula mediante una aminoacil ARNt sintetasa y un ARNt que reconoce al menos un codón selector de ARNm en la segunda célula. Dicha aminoacil ARNt sintetasa es un producto génico del gen *pylS* y el ARNt es un producto génico del gen *pylT*. En determinadas realizaciones, dichos compuestos de Fórmula (V) o Fórmula (VI) se incorporan a la proteína de la segunda célula mediante un ARNt ortogonal (ARNt-O) y una aminoacil ARNt sintetasa ortogonal (O-RS), donde los aminoacilatos O-RS del ARNt-O con el compuesto de Fórmula (V) o Fórmula (VI) y el ARNt-O reconocen al menos un codón selector de un ARNm en la segunda célula.

En determinadas realizaciones de dicha biosíntesis que utiliza células alimentadoras, el precursor es ornitina o arginina. En determinadas realizaciones de dicha biosíntesis, el precursor es D-ornitina o D-arginina. En determinadas realizaciones, el precursor es el ácido (2S)-2-amino-6-(2,5-diaminopentanamido)hexanoico. En determinadas realizaciones, el precursor es el ácido (2S)-2-amino-6-((R)2,5-diaminopentanamido)hexanoico.

5 El codón selector para la incorporación de la pirrolisina, la PCL y otros análogos de pirrolisina, que incluyen compuestos de Fórmula (V) y Fórmula (VI) biosintetizados utilizando células alimentadoras, está en un codón ámbar (TAG).

10 En determinadas realizaciones, la segunda célula es el mismo tipo de célula que la célula alimentadora, mientras que en otras realizaciones, la segunda célula es un tipo de célula diferente a la célula alimentadora. Las células usadas para dicha biosíntesis y/o la incorporación de la pirrolisina, la PCL y otros análogos de pirrolisina, que incluyen compuestos de Fórmula (V) y Fórmula (VI), son tanto células procariontas como células eucariotas. En determinadas realizaciones, las células procariontas incluyen, pero no se limitan a, *Escherichia coli*, *Mycobacterium smegmatis*, células de *Lactococcus lactis* y *Bacillus subtilis*. En otras realizaciones, las células eucariotas incluyen, pero no se limitan a, células de mamíferos, células de levaduras, células fúngicas, células vegetales o células de insectos. Dichas células de mamífero incluyen, pero no se limitan a, células de riñón embrionario humano (HEK293F), células de carcinoma epiteloidal humano (HeLa y GH3), células de riñón de mono (COS), células C6 de glioma de rata, células de riñón de cría de hámster (BHK-21) y células de ovario de hámster chino (CHO). En determinadas realizaciones, las células de levadura incluyen, pero no se limitan a, células de *Saccharomyces cerevisiae* y *Pichia pastoris*. En otras realizaciones, las células de insecto incluyen, pero no se limitan a, células de *Spodoptera frugiperda* sf9 y sf21, células de *Trichoplusia ni* (BTI TN-5B1-4 o High-Five™) y células de *Mamestra brassicae*.

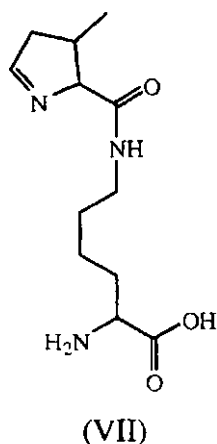
25 En otro aspecto proporcionado en el presente documento, la pirrolisina, la PCL y otros análogos de pirrolisina, que incluyen compuestos de Fórmula (V) y Fórmula (VI) se biosintetizan en células alimentadoras y dicha pirrolisina y/o PCL biosintetizadas, que incluyen compuestos de Fórmula (V) y Fórmula (VI) se purifican a continuación a partir del cultivo de células alimentadoras y se añaden al medio de crecimiento del segundo cultivo que contiene una segunda célula. Dicha pirrolisina purificada, la PCL y/u otros análogos de pirrolisina se incorporan a continuación a una proteína sintetizada en la segunda célula. En determinadas realizaciones, la pirrolisina, la PCL y otros análogos de pirrolisina, que incluyen compuestos de Fórmula (V) y Fórmula (VI) se biosintetizan en células alimentadoras en contacto con un medio de crecimiento que comprende un precursor, y las células alimentadoras contienen un gen *pylB*, un gen *pylC* y un gen *pylD*. En otras realizaciones, la pirrolisina, la PCL y otros análogos de pirrolisina, que incluyen compuestos de Fórmula (V) y Fórmula (VI), se biosintetizan en células alimentadoras en contacto con un medio de crecimiento que comprende un precursor, y las células alimentadoras contienen un gen *pylC* y un gen *pylD*. Las segundas células usadas en este aspecto contienen un gen *pylS* y un gen *pylT*. La pirrolisina, la PCL y otros análogos de pirrolisina, que incluyen compuestos de Fórmula (V) y Fórmula (VI), se incorporan a la proteína de la segunda célula mediante una aminoacil ARNt sintetasa y un ARNt que reconoce al menos un codón selector de ARNm en la segunda célula. Dicha aminoacil ARNt sintetasa es un producto génico del gen *pylS* y el ARNt es un producto génico del gen *pylT*. En determinadas realizaciones, dichos compuestos de Fórmula (V) o Fórmula (VI) se incorporan a la proteína de la segunda célula mediante un ARNt ortogonal (ARNt-O) y una aminoacil ARNt sintetasa ortogonal (O-RS), donde los aminoacilatos O-RS del ARNt-O con el compuesto de Fórmula (V) o Fórmula (VI) y el ARNt-O reconocen al menos un codón selector de un ARNm en la segunda célula.

45 En determinadas realizaciones de dicha biosíntesis que utiliza células alimentadoras, el precursor es ornitina o arginina. En determinadas realizaciones de dicha biosíntesis, el precursor es D-ornitina o D-arginina. En determinadas realizaciones, el precursor es el ácido (2S)-2-amino-6-(2,5-diaminopentanamido)hexanoico. En determinadas realizaciones, el precursor es el ácido (2S)-2-amino-6-((R)2,5-diaminopentanamido)hexanoico.

50 El codón selector para la incorporación de la pirrolisina, la PCL y otros análogos de pirrolisina, que incluyen compuestos de Fórmula (V) y Fórmula (VI) biosintetizados utilizando células alimentadoras, está en un codón ámbar (TAG).

55 En determinadas realizaciones, la segunda célula es el mismo tipo de célula que la célula alimentadora, mientras que en otras realizaciones, la segunda célula es un tipo de célula diferente a la célula alimentadora. Las células usadas para dicha biosíntesis y/o la incorporación de la pirrolisina, la PCL y otros análogos de pirrolisina, que incluyen compuestos de Fórmula (V) y Fórmula (VI), son tanto células procariontas como células eucariotas. En determinadas realizaciones, las células procariontas incluyen, pero no se limitan a, *Escherichia coli*, *Mycobacterium smegmatis*, células de *Lactococcus lactis* y *Bacillus subtilis*. En otras realizaciones, las células eucariotas incluyen, pero no se limitan a, células de mamíferos, células de levaduras, células fúngicas, células vegetales o células de insectos. Dichas células de mamífero incluyen, pero no se limitan a, células de riñón embrionario humano (HEK293F), células de carcinoma epiteloidal humano (HeLa y GH3), células de riñón de mono (COS), células C6 de glioma de rata, células de riñón de cría de hámster (BHK-21) y células de ovario de hámster chino (CHO). En determinadas realizaciones, las células de levadura incluyen, pero no se limitan a, células de *Saccharomyces cerevisiae* y *Pichia pastoris*. En otras realizaciones, las células de insecto incluyen, pero no se limitan a, células de *Spodoptera frugiperda* sf9 y sf21, células de *Trichoplusia ni* (BTI TN-5B1-4 o High-Five™) y células de *Mamestra brassicae*.

Se proporcionan también en el presente documento aminoácidos que tienen la estructura de la Fórmula (VII);



5 donde el compuesto de Fórmula (VII), o sus isómeros o tautómeros, se generan biosintéticamente en una célula que contiene un gen pylB, un gen pylC, y un gen pylD, y la célula está en contacto con un medio de crecimiento que contiene D-ornitina o D-arginina o ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico o el ácido 2,5-diamino-3-metilpentanoico or el ácido (2S)-2-amino-6-(2,5-diaminopentanamido)hexanoico. Mientras que en otras realizaciones, dichos compuestos de Fórmula (VII) se generan biosintéticamente en una célula que contiene un gen pylC y un gen pylD y la célula está en contacto con un medio de crecimiento que contiene ácido 2,5-diamino-3-metilpentanoico. En determinadas realizaciones, la célula está en contacto con un medio de crecimiento que comprende ácido D-2,5-diamino-3-metilpentanoico. En determinadas realizaciones, el ácido 2,5-diamino-3-metilpentanoico es ácido (2R,3S)-2,5-diamino-3-metilpentanoico. En determinadas realizaciones, el ácido 2,5-diamino-3-metilpentanoico es ácido (2R,3R)-2,5-diamino-3-metilpentanoico.

15 Las células donde se genera biosintéticamente el compuesto de Fórmula (VII) comprenden además un gen pylS y un gen pylT y el compuesto de Fórmula (VII) se incorpora a una proteína de la célula mediante una aminoacil ARNt sintetasa y un ARNt que reconoce al menos un codón selector de un ARNm en la célula. Dicha aminoacil ARNt sintetasa es un producto génico del gen pylS y el ARNt es un producto génico del gen pylT. En determinadas realizaciones, dicho compuesto de Fórmula (VII) se incorporan a una proteína de la célula mediante un ARNt ortogonal (ARNt-O) y una aminoacil ARNt sintetasa ortogonal (O-RS), donde los aminoacilatos O-RS del ARNt-O con el compuesto de Fórmula (VII) y el ARNt-O reconocen al menos un codón selector de un ARNm en la célula.

25 El codón selector para la incorporación de un compuesto de Fórmula (VII) es un codón ámbar (TAG).

30 Las células usadas para la biosíntesis y/o la incorporación de un compuesto de Fórmula (VII) son tanto células procariotas como células eucariotas. En determinadas realizaciones, las células procariotas incluyen, pero no se limitan a, Escherichia coli, Mycobacterium smegmatis, células de Lactococcus lactis y Bacillus subtilis. En otras realizaciones, las células eucariotas incluyen, pero no se limitan a, células de mamíferos, células de levaduras, células fúngicas, células vegetales o células de insectos. Dichas células de mamífero incluyen, pero no se limitan a, células de riñón embrionario humano (HEK293F), células de carcinoma epiteloides humano (HeLa y GH3), células de riñón de mono (COS), células C6 de glioma de rata, células de riñón de cría de hámster (BHK-21) y células de ovario de hámster chino (CHO). En determinadas realizaciones, las células de levadura incluyen, pero no se limitan a, células de Saccharomyces cerevisiae y Pichia pastoris. En otras realizaciones, las células de insecto incluyen, pero no se limitan a, células de Spodoptera frugiperda sf9 y sf21, células de Trichoplusia ni (BTI TN-5B1-4 o High-Five™) y células de Mamestra brassicae.

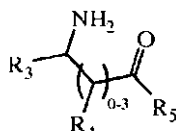
40 En determinadas realizaciones, una o más pirrolisinas, PCL y/u otros análogos de pirrolisina se incorporan a las proteínas, polipéptidos y/o péptidos usando los métodos proporcionados en el presente documento, y dicha pirrolisina y/o PCL se derivatizan usando los métodos proporcionados en el presente documento.

Derivatización de PCL y pirrolisina

45 En determinadas realizaciones, n es un número entero de 1 a 4000, En determinadas realizaciones, n es un número entero de 1 a 3000, En determinadas realizaciones, n es un número entero de 1 a 2000, En determinadas realizaciones, n es un número entero de 1 a 1000, En determinadas realizaciones, n es un número entero de 1 a 700, En determinadas realizaciones, n es un número entero de 1 a 800, En determinadas realizaciones, n es un número entero de 1 a 600, En determinadas realizaciones, n es un número entero de 1 a 500, En determinadas realizaciones, n es un número entero de 1 a 400, En determinadas realizaciones, n es un número entero de 1 a 300, En determinadas realizaciones, n es un número entero de 1 a 200, En determinadas realizaciones, n es un número entero de 1 a 100, En

determinadas realizaciones, n es un número entero de 1 a 90, En determinadas realizaciones, n es un número entero de 1 a 80, En determinadas realizaciones, n es un número entero de 1 a 70, En determinadas realizaciones, n es un número entero de 1 a 60, En determinadas realizaciones, n es un número entero de 1 a 50, En determinadas realizaciones, n es un número entero de 1 a 40, En determinadas realizaciones, n es un número entero de 1 a 30, En determinadas realizaciones, n es un número entero de 1 a 20, En determinadas realizaciones, n es un número entero de 1 a 10, En determinadas realizaciones, n es un número entero de 1 a 5.

Asimismo, se proporcionan en el presente documento métodos para el marcado específico de proteínas, polipéptidos y/o péptidos de Fórmula (I), donde el método implica una premezcla de dichas proteínas, polipéptidos y/o péptidos que contienen una o más pirrolisinas y/o restos de PCL con un reactivo que tiene la estructura de la Fórmula (III):



(III)

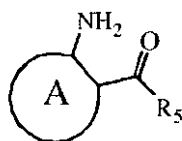
donde:

R₃, R₅ y cada R₄ se seleccionan independientemente entre H, -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidroxil-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹;

L se selecciona entre un enlace, alquileo C₁₋₈, alquileo C₁₋₈ halo-sustituido, alquileo C₁₋₈ hidroxil-sustituido, alquilenilo C₂₋₈, alquilenilo C₁₋₈ halo-sustituido, alquilenilo C₂₋₈ hidroxil-sustituido, un polialquilenglicol, -O(CR¹¹R¹²)_k-, -S(CR¹¹R¹²)_k-, -S(O)_k(CR¹¹R¹²)_k-, -O(CR¹¹R¹²)_k-NR¹¹C(O)-, -O(CR¹¹R¹²)_kC(O)NR¹¹-, -C(O)-, -C(O)(CR¹¹R¹²)_k-, -C(S)-, -C(S)(CR¹¹R¹²)_k-, C(O)NR¹¹-, -NR¹¹C(O)-, -NR¹¹(CR¹¹R¹²)_k-, -CONR¹¹(CR¹¹R¹²)_k-, -N(R¹¹)CO(CR¹¹R¹²)_k-, -C(O)NR¹¹(CR¹¹R¹²)_k-, -NR¹¹C(O)(CR¹¹R¹²)_k-, en donde cada R¹¹ y R¹² son independientemente H, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, o alquilo C₁₋₈ hidroxil-sustituido, y k es un número entero de 1 a 12;

X¹ se ha seleccionado de una marca, un colorante, un polímero, un polímero soluble en agua, un polialquilenglicol, un poli(etilenglicol), un derivado de poli(etilenglicol), un azúcar, un lípido, un fotorreticulador, un compuesto citotóxico, un fármaco, una marca de afinidad, una marca de fotoafinidad, un compuesto reactivo; una resina, un péptido, una segunda proteína o polipéptido o análogo de polipéptido, un anticuerpo o fragmento de anticuerpo, un quelante metálico, un cofactor, un ácido graso, un hidrato de carbono, un polinucleótido, un ADN, un ARN, una sonda de la PCR, un polinucleótido antisentido, un ribooligonucleótido, un desoxirribooligonucleótido, ADN modificado con fosforotioato, ADN y ARN modificados, ácido nucleico peptídico, un sacárido, un disacárido, un oligosacárido, un polisacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial, una nanopartícula, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un grupo funcional novedoso, un grupo que interactúa de forma covalente o no covalente con otras moléculas, un resto fotoestimulado, un resto excitable mediante radiación actínica, un ligando un resto fotoisomerizable, biotina, un análogo de biotina, un resto que incorpora un átomo pesado, un grupo escindible químicamente, un grupo fotoescindible, una cadena lateral alargada, un azúcar unido a carbono, un agente redox activo, un aminotioácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo cromóforo, un grupo quimioluminiscente, un resto fluorescente, un grupo denso en electrones, un grupo magnético, un grupo intercalante, un grupo quelante, un cromóforo, un agente de transferencia de energía, un agente biológicamente activo, una marca detectable, una molécula pequeña, un ácido nucleico inhibidor, un ARNip, un radionucleótido, un agente de captura de neutrones, un derivado de biotina, punto(s) cuántico(s), un nanotransmisor, un radiotransmisor, una abzima, una enzima, un activador complejo activado, un virus, una toxina, un adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, un epítipo de linfocito T, un fosfolípido, una molécula de tipo LPS, hemocianina de lapa californiana (KLH), un hapteno inmunógeno, un aglicano, un alérgeno, una angioestatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimotopo, un receptor, una micela inversa, detergentes, potenciadores de la respuesta inmune, colorantes fluorescentes, reactivos FRET, sondas de radioimagen, otra sonda de espectroscopía, profármacos, toxinas para inmunoterapia, un soporte sólido, -CH₂CH₂-(OCH₂CH₂O)_p-OX² -O-(CH₂CH₂O)_pCH₂CH₂-X² y cualquier combinación de los mismos, donde p es de 1 a 10.000 y X² es H, un alquilo C₁₋₈, un grupo protector o un grupo funcional terminal.

Se proporcionan además en el presente documento métodos para el marcado específico de sitios de proteínas, polipéptidos y/o péptidos de Fórmula (I), donde el método implica premezclar y hacer reaccionar en condiciones adecuadas, dichas proteínas, polipéptidos y/o péptidos que contienen una o más pirrolisinas y/o restos de PCL con un reactivo que tiene la estructura de la Fórmula (IV):



(IV)

donde:

- 5 R₅ se selecciona entre -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidroxi-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹; A es un cicloalquilo C₃₋₈, heterocicloalquilo C₃₋₈, un arilo monocíclico de 5-6 miembros, un heteroarilo monocíclico de 5-6 miembros, un anillo bicíclico condensado de 9-10 miembros o un anillo tricíclico condensado de 13-14 miembros, en donde A está opcionalmente sustituido con 1 a 5 sustituyentes seleccionados independientemente entre -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidroxi-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹;
- 10 L se selecciona entre un enlace, alquilenilo C₁₋₈, alquilenilo C₁₋₈ halo-sustituido, alquilenilo C₁₋₈ hidroxi-sustituido, alquenileno C₂₋₈, alquenileno C₂₋₈ halo-sustituido, alquenileno C₂₋₈ hidroxi-sustituido, un polialquilenglicol, un polietilenglicol),
 $-O(CR^{11}R^{12})_k-$, $-S(CR^{11}R^{12})_k-$, $-S(O)_k(CR^{11}R^{12})_k-$,
 $-O(CR^{11}R^{12})_k-NR^{11}C(O)-$, $O(CR^{11}R^{12})_kC(O)NR^{11}-$, $-C(O)-$, $-C(O)(CR^{11}R^{12})_k-$, $-C(S)-$, $-C(S)(CR^{11}R^{12})_k-$, $C(O)NR^{11}-$,
 $-NR^{11}C(O)-$, $-NR^{11}(CR^{11}R^{12})_k-$, $-CONR^{11}(CR^{11}R^{12})_k-$, $-N(R^{11})CO(CR^{11}R^{12})-$, $-C(O)NR^{11}(CR^{11}R^{12})_k-$,
 $-NR^{11}C(O)(CR^{11}R^{12})_k-$, en donde cada R¹¹ y R¹² son independientemente H, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, o alquilo C₁₋₈ hidroxi-sustituido, y k es un número entero de 1 a 12;
- 15 X¹ se ha seleccionado de una marca, un colorante, un polímero, un polímero soluble en agua, un polialquilenglicol, un poli(etilenglicol), un derivado de poli(etilenglicol), un azúcar, un lípido, un fotorreticulador, un compuesto citotóxico, un fármaco, una marca de afinidad, una marca de fotoafinidad, un compuesto reactivo; una resina, un péptido, una segunda proteína o polipéptido o análogo de polipéptido, un anticuerpo o fragmento de anticuerpo, un quelante metálico, un cofactor, un ácido graso, un hidrato de carbono, un polinucleótido, un ADN, un ARN, una sonda de la PCR, un polinucleótido antisentido, un ribooligonucleótido, un desoxirribooligonucleótido, ADN modificado con fosforioato, ADN y ARN modificados, ácido nucleico peptídico, un sacárido, un disacárido, un oligosacárido, un polisacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial, una nanopartícula, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un grupo funcional novedoso, un grupo que interactúa de forma covalente o no covalente con otras moléculas, un resto fotoestimulado, un resto excitable mediante radiación actínica, un ligando un resto fotoisomerizable, biotina, un análogo de biotina, un resto que incorpora un átomo pesado, un grupo escindible químicamente, un grupo fotoescindible, una cadena lateral alargada, un azúcar unido a carbono, un agente redox activo, un aminotioácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo cromóforo, un grupo quimioluminiscente, un resto fluorescente, un grupo denso en electrones, un grupo magnético, un grupo intercalante, un grupo quelante, un cromóforo, un agente de transferencia de energía, un agente biológicamente activo, una marca detectable, una molécula pequeña, un ácido nucleico inhibidor, un ARNip, un radionucleótido, un agente de captura de neutrones, un derivado de biotina, punto(s) cuántico(s), un nanotransmisor, un radiotransmisor, una abzima, una enzima, un activador complejo activado, un virus, una toxina, un adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, un epítipo de linfocito T, un fosfolípido, una molécula de tipo LPS, hemocianina de lapa californiana (KLH), un hapteno inmunógeno, un aligiano, un alérgeno, una angioestatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimotopo, un receptor, una micela inversa, detergentes, potenciadores de la respuesta inmune, colorantes fluorescentes, reactivos FRET, sondas de radioimagen, otra sonda de espectroscopía, profármacos, toxinas para inmunoterapia, un soporte sólido,
 $-CH_2CH_2-(OCH_2CH_2O)_p-OX^2rO-(CH_2CH_2O)_pCH_2CH_2-X^2$, y cualquier combinación de los mismos, donde p es de 1 a 10.000 y X² es H, un alquilo C₁₋₈, un grupo protector o un grupo funcional terminal.
- 20
25
30
35
40
45

En determinadas realizaciones, k es un número entero de 1 a 11, En determinadas realizaciones, k es un número entero de 1 a 10, En determinadas realizaciones, k es un número entero de 1 a 9, En determinadas realizaciones, k es un número entero de 1 a 8, En determinadas realizaciones, k es un número entero de 1 a 7, En determinadas realizaciones, k es un número entero de 1 a 6, En determinadas realizaciones, k es un número entero de 1 a 5, En determinadas realizaciones, k es un número entero de 1 a 4, En determinadas realizaciones, k es un número entero de 1 a 3, En determinadas realizaciones, k es un número entero de 1 a 2,

50

En determinadas realizaciones, p es un número entero de 1 a 8000, En determinadas realizaciones, p es un número entero de 1 a 7000, En determinadas realizaciones, p es un número entero de 1 a 6000, En determinadas realizaciones, p es un número entero de 1 a 5000, En determinadas realizaciones, p es un número entero de 1 a 4000, En determinadas realizaciones, p es un número entero de 1 a 3000, En determinadas realizaciones, p es un número entero de 1 a 2000, En determinadas realizaciones, p es un número entero de 1 a 1000, En determinadas

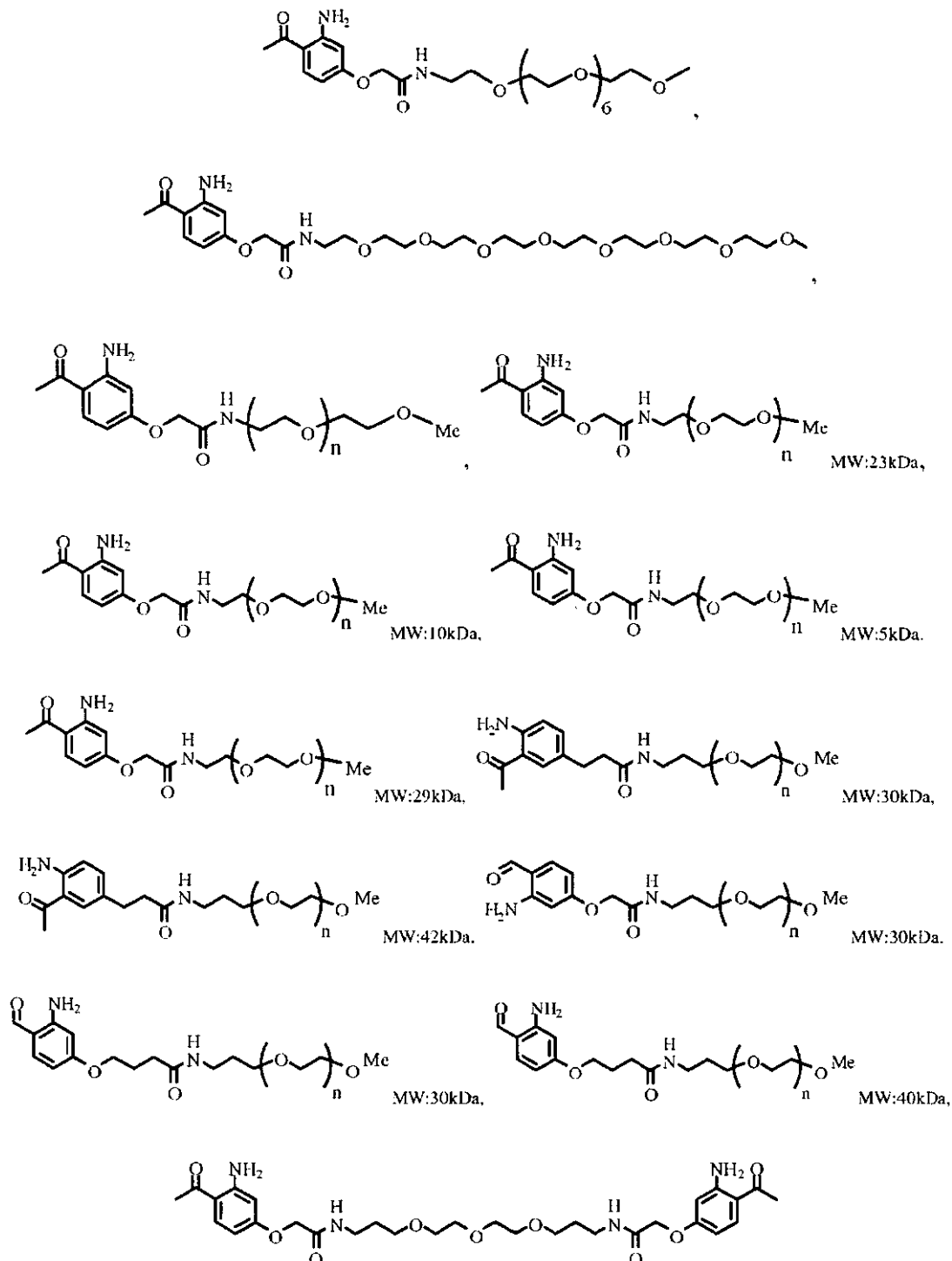
55

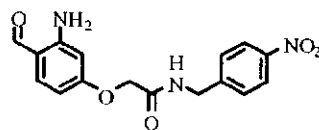
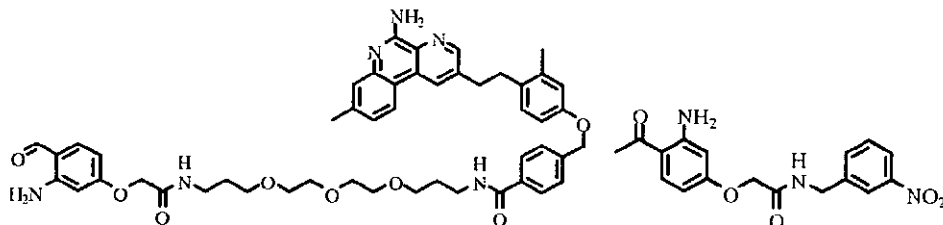
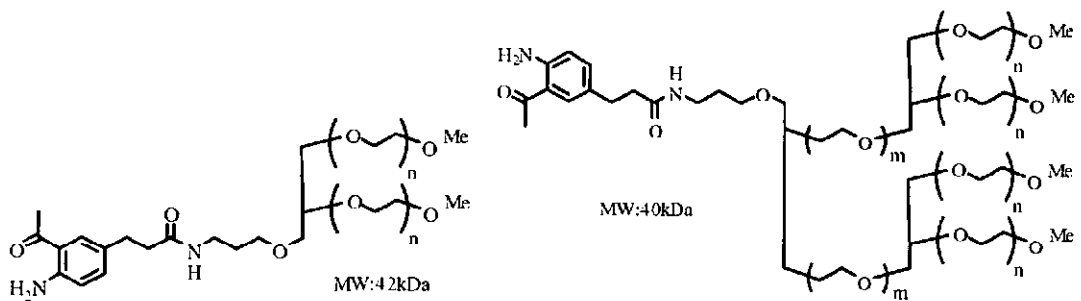
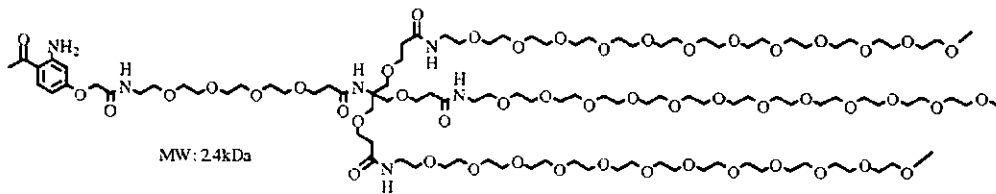
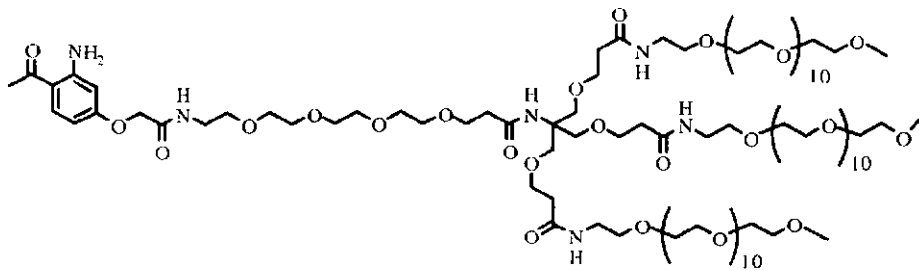
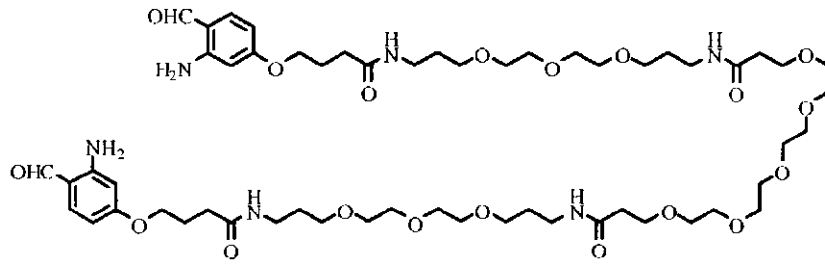
realizaciones, p es un número entero de 1 a 500, En determinadas realizaciones, p es un número entero de 1 a 400, En determinadas realizaciones, p es un número entero de 1 a 300, En determinadas realizaciones, p es un número entero de 1 a 200, En determinadas realizaciones, p es un número entero de 1 a 100, En determinadas realizaciones, p es un número entero de 1 a 90, En determinadas realizaciones, p es un número entero de 1 a 80, En determinadas realizaciones, p es un número entero de 1 a 70, En determinadas realizaciones, p es un número entero de 1 a 60, En determinadas realizaciones, p es un número entero de 1 a 50, En determinadas realizaciones, p es un número entero de 1 a 40, En determinadas realizaciones, p es un número entero de 1 a 30, En determinadas realizaciones, p es un número entero de 1 a 20, En determinadas realizaciones, p es un número entero de 1 a 10, En determinadas realizaciones, p es un número entero de 1 a 5.

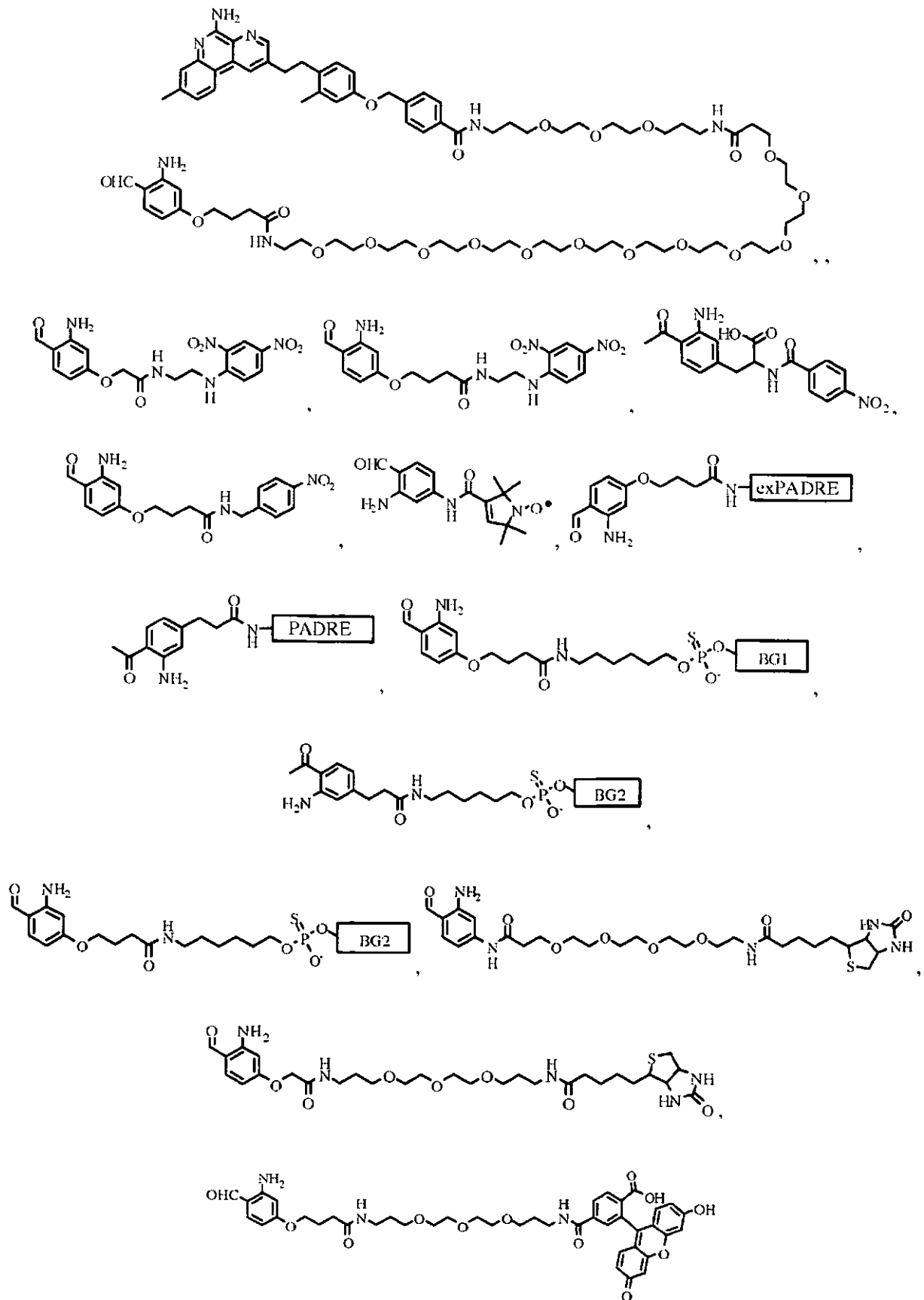
5

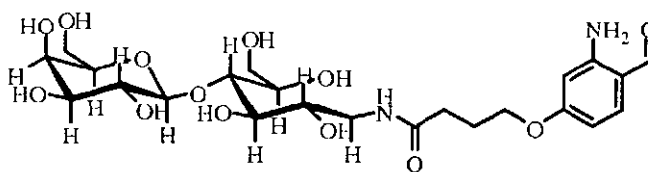
10

Y el ejemplo no limitante de los compuestos de Fórmula (IV) incluye:

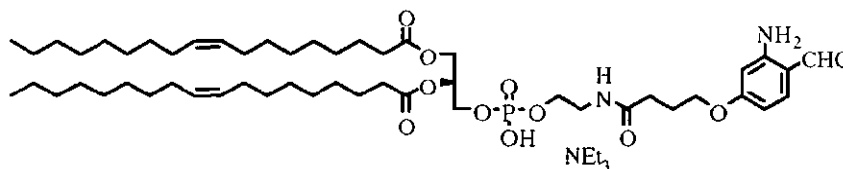








y



5

donde los compuestos que tienen uno o más restos de polietilenglicol (PEG) tienen un peso molecular promedio en el intervalo de 1000 Da a 50 kDa, y n es de 20 a 1200 y donde exPADRE es AlaGlySerArgSerGly(DAla)LysCha-ValAlaAlaTrpThrLeuLysAla(D-Ala)Gly-OH, PADRE es Gly(DAla)LysChaValAlaAlaTrpThrLeuLysAla(D-Ala)Gly-OH, BG-1 es 5'*t*c*.A*T*G*A*C*G*T*T*C*C*T*G*A*C*G*T*T-3' y BG2 es 5'*T*C*G*T*C*G*T*T*T*T*C*G*G*C*G*C*G*C*C*G-3', y donde * denota un enlace fosfoitoato.

10

En determinadas realizaciones, las proteínas, polipéptidos y/o péptidos que contienen una o más pirrolisinas y/o PCL derivatizadas usando los métodos proporcionados en el presente documento que tienen la estructura de la Fórmula (II):

15



donde:

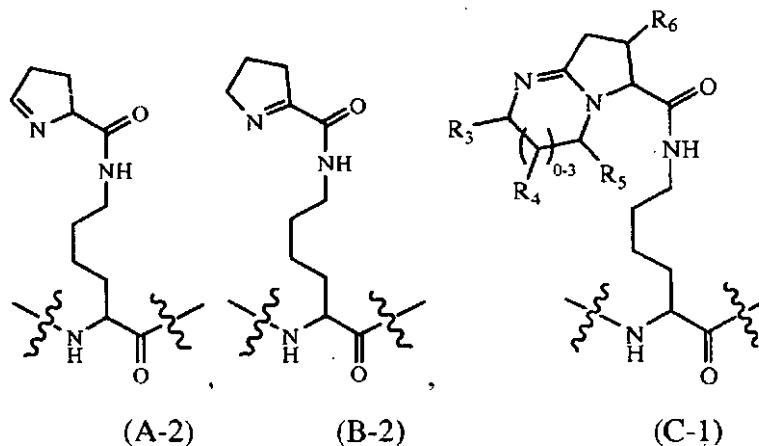
20

R₁ es H o un grupo de modificación del amino terminal; R₂ es OH o un grupo de modificación del carboxi terminal; n es un número entero de 1 a 5000;

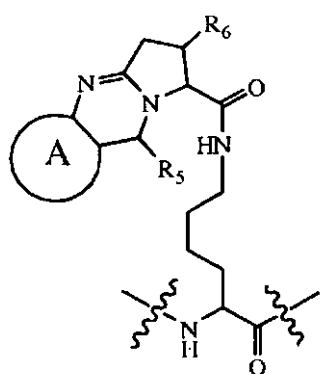
25

cada BB se selecciona de forma independiente entre un resto de aminoácido, un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (A-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (B-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (C-1) un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (D-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (E-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (F-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (G-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (H-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (I-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (J-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (K-1) y un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (L-1), o isómeros de los mismos;

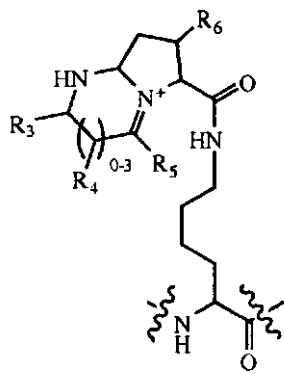
30



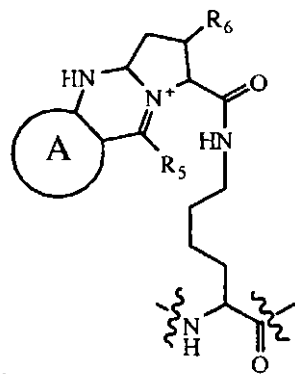
35



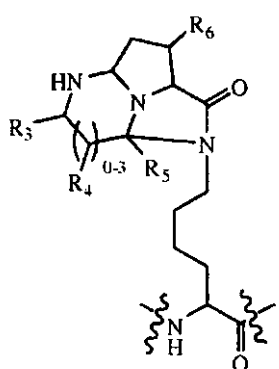
(D-1)



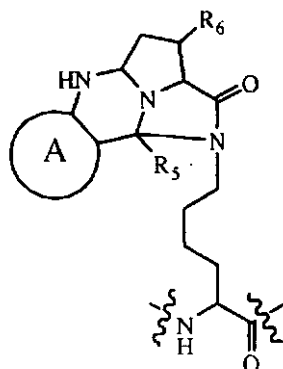
(E-1)



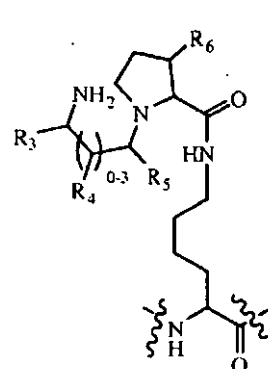
(F-1)



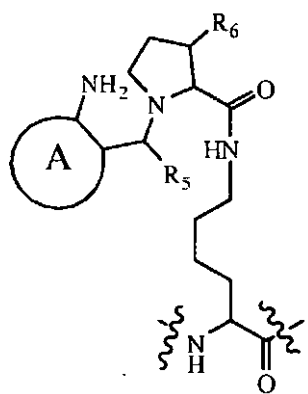
(G-1)



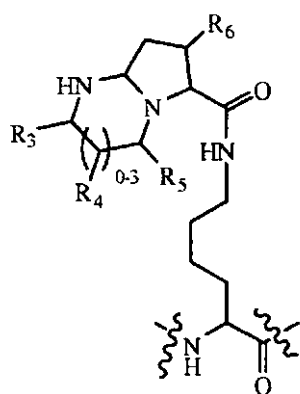
(H-1)



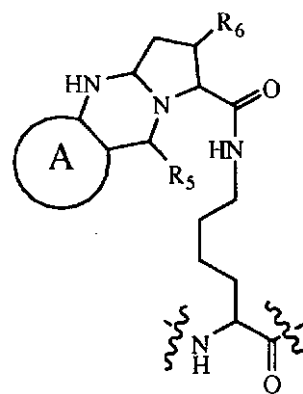
(I-1)



(J-1)



(K-1)



(L-1)

donde:

R₃, R₅ y cada R₄ se seleccionan independientemente entre H, -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidroxi-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹; R₆ es H o alquilo C₁;

A es un cicloalquilo C₃₋₈, heterocicloalquilo C₃₋₈, un arilo monocíclico de 5-6 miembros, un heteroarilo monocíclico de 5-6 miembros, un anillo bicíclico condensado de 9-10 miembros o un anillo tricíclico condensado de 3-14 miembros, en donde A está opcionalmente sustituido con 1 a 5 sustituyentes seleccionados independientemente entre -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidroxi-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹;

L se selecciona entre un enlace, alquilenilo C₁₋₈, alquilenilo C₁₋₈ halo-sustituido, alqueno C₁₋₈ hidroxi-sustituido, alquenileno C₂₋₈, alquenileno C₂₋₈ halo-sustituido, alquenileno C₂₋₈ hidroxi-sustituido, un polialquilenglicol, un polietilenglicol), -O(CR¹¹R¹²)_k-, -S(CR¹¹R¹²)_k-, -S(O)_k(CR¹¹R¹²)_k-, -O(CR¹¹R¹²)_k-NR¹¹C(O)-, -O(CR¹¹R¹²)_kC(O)NR¹¹-, -C(O)-, -C(O)(CR¹¹R¹²)_k-, -C(S)-, -C(S)(CR¹¹R¹²)_k-, -C(O)NR¹¹-, -NR¹¹C(O)-,

$-\text{NR}^{11}(\text{CR}^{11}\text{R}^{12})_k$, $-\text{CONR}^{11}(\text{CR}^{11}\text{R}^{12})_k$, $-\text{N}(\text{R}^{11})\text{CO}(\text{CR}^{11}\text{R}^{12})_k$, $-\text{C}(\text{O})\text{NR}^{11}(\text{CR}^{11}\text{R}^{12})_k$, $-\text{NR}^{11}\text{C}(\text{O})(\text{CR}^{11}\text{R}^{12})_k$, en donde cada R^{11} y R^{12} son independientemente H, alquilo C_{1-8} , alquilo C_{1-8} halo-sustituido, o alquilo C_{1-8} hidroxil-sustituido, y k es un número entero de 1 a 12, y

X^1 se ha seleccionado de una marca, un colorante, un polímero, un polímero soluble en agua, un polialquilenglicol, un poli(etilenglicol), un derivado de poli(etilenglicol), un azúcar, un lípido, un fotorreticulador, un compuesto citotóxico, un fármaco, una marca de afinidad, una marca de fotoafinidad, un compuesto reactivo; una resina, un péptido, una segunda proteína o polipéptido o análogo de polipéptido, un anticuerpo o fragmento de anticuerpo, un quelante metálico, un cofactor, un ácido graso, un hidrato de carbono, un polinucleótido, un ADN, un ARN, una sonda de la PCR, un polinucleótido antisentido, un ribooligonucleótido, un desoxirribooligonucleótido, ADN modificado con fosforotioato, ADN y ARN modificados, ácido nucleico peptídico, un sacárido, un disacárido, un oligosacárido, un polisacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial, una nanopartícula, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un grupo funcional novedoso, un grupo que interactúa de forma covalente o no covalente con otras moléculas, un resto fotoestimulado, un resto excitable mediante radiación actínica, un ligando un resto fotoisomerizable, biotina, un análogo de biotina, un resto que incorpora un átomo pesado, un grupo escindible químicamente, un grupo fotoescindible, una cadena lateral alargada, un azúcar unido a carbono, un agente redox activo, un aminotioácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo cromóforo, un grupo quimioluminiscente, un resto fluorescente, un grupo denso en electrones, un grupo magnético, un grupo intercalante, un grupo quelante, un cromóforo, un agente de transferencia de energía, un agente biológicamente activo, una marca detectable, una molécula pequeña, un ácido nucleico inhibidor, un ARNip, un radionucleótido, un agente de captura de neutrones, un derivado de biotina, punto(s) cuántico(s), un nanotransmisor, un radiotransmisor, una abzima, una enzima, un activador complejo activado, un virus, una toxina, un adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, un epítipo de linfocito T, un fosfolípido, una molécula de tipo LPS, hemocianina de lapa californiana (KLH), un hapteno inmunógeno, un aglicano, un alérgeno, una angioestatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimotopo, un receptor, una micela inversa, detergentes, potenciadores de la respuesta inmune, colorantes fluorescentes, reactivos FRET, sondas de radioimagen, otra sonda de espectroscopía, profármacos, toxinas para inmunoterapia, un soporte sólido, $-\text{CH}_2\text{CH}_2-(\text{OCH}_2\text{CH}_2\text{O})_p-\text{QX}^2$ $-\text{O}-(\text{CH}_2\text{CH}_2\text{O})_p\text{CH}_2\text{CH}_2-\text{X}^2$, y cualquier combinación de los mismos, donde p es de 1 a 10.000, y X^2 es H, un alquilo C_{1-8} , un grupo protector o un grupo funcional terminal, y donde al menos un BB es un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (C-1) o la Fórmula (D-1) o la Fórmula (E-1) o la Fórmula (F-1) o la Fórmula (G-1) o la Fórmula (H-1) o la Fórmula (I-1) o la Fórmula (J-1) o la Fórmula (K-1) o la Fórmula (L-1), o isómeros de los mismos.

En determinadas realizaciones, k es un número entero de 1 a 11, En determinadas realizaciones, k es un número entero de 1 a 10, En determinadas realizaciones, k es un número entero de 1 a 9, En determinadas realizaciones, k es un número entero de 1 a 8, En determinadas realizaciones, k es un número entero de 1 a 7, En determinadas realizaciones, k es un número entero de 1 a 6, En determinadas realizaciones, k es un número entero de 1 a 5, En determinadas realizaciones, k es un número entero de 1 a 4, En determinadas realizaciones, k es un número entero de 1 a 3, En determinadas realizaciones, k es un número entero de 1 a 2,

En determinadas realizaciones, p es un número entero de 1 a 8000, En determinadas realizaciones, p es un número entero de 1 a 7000, En determinadas realizaciones, p es un número entero de 1 a 6000, En determinadas realizaciones, p es un número entero de 1 a 5000, En determinadas realizaciones, p es un número entero de 1 a 4000, En determinadas realizaciones, p es un número entero de 1 a 3000, En determinadas realizaciones, p es un número entero de 1 a 2000, En determinadas realizaciones, p es un número entero de 1 a 1000, En determinadas realizaciones, p es un número entero de 1 a 500, En determinadas realizaciones, p es un número entero de 1 a 400, En determinadas realizaciones, p es un número entero de 1 a 300, En determinadas realizaciones, p es un número entero de 1 a 200, En determinadas realizaciones, p es un número entero de 1 a 100, En determinadas realizaciones, p es un número entero de 1 a 90, En determinadas realizaciones, p es un número entero de 1 a 80, En determinadas realizaciones, p es un número entero de 1 a 70, En determinadas realizaciones, p es un número entero de 1 a 60, En determinadas realizaciones, p es un número entero de 1 a 50, En determinadas realizaciones, p es un número entero de 1 a 40, En determinadas realizaciones, p es un número entero de 1 a 30, En determinadas realizaciones, p es un número entero de 1 a 20, En determinadas realizaciones, p es un número entero de 1 a 10, En determinadas realizaciones, p es un número entero de 1 a 5,

En determinadas realizaciones, n es un número entero de 1 a 4000, En determinadas realizaciones, n es un número entero de 1 a 3000, En determinadas realizaciones, n es un número entero de 1 a 2000, En determinadas realizaciones, n es un número entero de 1 a 1000, En determinadas realizaciones, n es un número entero de 1 a 700, En determinadas realizaciones, n es un número entero de 1 a 600, En determinadas realizaciones, n es un número entero de 1 a 500, En determinadas realizaciones, n es un número entero de 1 a 400, En determinadas realizaciones, n es un número entero de 1 a 300, En determinadas realizaciones, n es un número entero de 1 a 200, En determinadas realizaciones, n es un número entero de 1 a 100, En determinadas realizaciones, n es un número entero de 1 a 90, En determinadas realizaciones, n es un número entero de 1 a 80, En determinadas realizaciones, n es un número entero de 1 a 70, En determinadas realizaciones, n es un

número entero de 1 a 60, En determinadas realizaciones, n es un número entero de 1 a 50, En determinadas realizaciones, n es un número entero de 1 a 40, En determinadas realizaciones, n es un número entero de 1 a 30, En determinadas realizaciones, n es un número entero de 1 a 20, En determinadas realizaciones, n es un número entero de 1 a 10, En determinadas realizaciones, n es un número entero de 1 a 5,

5 En determinadas realizaciones, los compuestos de Fórmula (III) utilizados en los métodos de derivatización proporcionados en el presente documento incluyen, pero no se limitan a, aminoazúcares. En determinadas realizaciones, dichos aminoazúcares incluyen, pero no se limitan a, D-manosamina y D-galactosamina.

10 En determinadas realizaciones, Los compuestos de Fórmula (IV) usados en los métodos de derivatización proporcionados en el presente documento incluyen, pero no se limitan a, 2-amino-benzaldehído (2-ABA), derivados de 2-amino-benzaldehído (2-ABA), derivados de 2-amino-benzaldehído (2-ABA) proporcionados en el presente documento, 2-amino-acetofenona (2-AAP), derivados de 2-amino-acetofenona (2-AAP), derivados de 2-amino-acetofenona (2-AAP) proporcionados en el presente documento, 2-amino-5-nitro-benzofenona (2-ANBP),
15 derivados de 2-amino-5-nitro-benzofenona (2-ANBP) y derivados de 2-amino-5-nitro-benzofenona (2-ANBP) proporcionados en el presente documento.

Figura 22 muestra un esquema de reacción para la derivatización química de PCL-A con 2-amino-benzaldehído (2-ABA), donde uno o más PCL-A ha sido incorporado específicamente al sitio en la proteína. La reacción de ciclación del semialdehído con 2-ABA es una reacción de tipo Friedlaender que da como resultado la formación de un resto de tipo quinazolina (22-A o 22-B), que se puede reducir adicionalmente con un agente reductor adecuado para formar restos de tipo tetrahidroquinazolina (22-D) o anilinas sustituidas (22-E). Alternativamente, la reacción adicional del resto de tipo quinazolina 22-B da como resultado la formación del resto de anillo fusionado (22-C). Los agentes reductores adecuados para la reducción de los restos de tipo quinazolina incluyen, pero no se limitan a, cianoborohidruro de sodio y borohidruro de sodio.

Además, la Figura 23 muestra las diversas estructuras de los conjugados de proteínas formados tras la reacción de la PCL-A con cualquiera de 2-ABA, 2-AAP o 2-ANBP. La masa esperada aumenta debido a que se muestra también la unión de dichos grupos. Se caracterizaron las estructuras obtenidas utilizando dichos grupos de espectroscopía de RMN (véase el Ejemplo 45). Asimismo, se encontró que la reducción de NaCNBH₃ estabiliza los conjugados de proteínas basados en la PCL y evita la disociación del enlace de PCL-ABA y PCL-AAP incluso a elevadas temperaturas (véase el Ejemplo 44).

La reacción de 2-ABA con el ácido A¹-pirrolina-5-carboxílico (ácido L-1-pirrolina-5-carboxílico) (véase Strecker, H. J., (1971), *Methods in Enzymology* 27B, 254-257; Vogel, H. J., y Davis, B. D. (1953), "Glutamic gamma-semialdehyde and delta-pyrroline-5-carboxylic acid, intermediates in the biosynthesis of proline," *J. Am. Chem. Soc.* 74, 109-112; y Schoepf, C., y Oechler, F., (1936), *Ann. Chem.* 523, 1), y otras pirrolinas (véanse Schoepf, C., y Oechler, F., (1936), *Ann. Chem.* 523, 1; y Schoepf, C., y Steuer, H., (1947), *Ann. Chem.* 558, 124) se han usado como ensayo colorimétrico en estudios del papel del ácido A¹-pirrolina-5-carboxílico en el metabolismo y la biosíntesis de la prolina (véase Vogel, H. J., y Davis, B. D. (1953), "Glutamic gamma-semialdehyde and A¹-pyrroline-5-carboxylic acid, intermediates in the biosynthesis of proline," *J. Am. Chem. Soc.* 74, 109-112; Strecker, H. J., (1960), "The interconversion of glutamic acid and proline," *J. Biol. Chem.* 235, 2045-2050; Mezl, V. A., y Knox, W. E., (1976), "Properties and analysis of a stable derivative of pyrroline-5-carboxylic acid for use in metabolic studies," *Analytical Biochemistry* 74, 430-40; Wu, G. Y., y Seifter, S., (1975), "A new method for the preparation of delta 1-pyrroline 5-carboxylic acid and proline," *Analytical Biochemistry* 67, 413-21; Williams, I., y Frank, L., (1975), "Improved chemical synthesis and enzymatic assay of A¹-pyrroline-5-carboxylic acid," *Anal. Biochem.* 64, 85-97, y Strecker, H. J., (1957), "The interconversion of glutamic acid and proline," *J. Biol. Chem.* 225, 825). Aunque el esquema de reacción muestra la formación del semialdehído-gamma-glutámico intermedio, la reacción puede proceder directamente de la pirrolina-carboxi-lisina sin la formación de dicho intermedio.

Los estudios de espectrometría de masas de alta resolución de la proteína modelo hRBP4 verificaron la derivatización química del resto de la PCL con 2-ABA, donde la PCL era incorporada específicamente al sitio en la proteína hRBP4 (Figuras 25-26, véase el Ejemplo 11). La Figura 24 muestra el análisis por espectrometría de masas de hRBP4 Phe122PCL derivatizada con 2-ABA, donde la proteína derivatizada con 2-ABA da como resultado un pico máximo a 23269,2 Da y se detecta una cantidad mínima de proteína no modificada a 23166,8 Da. El aumento de masa de 102,4 Da en la proteína derivatizada es consistente con el aumento esperado de 103 Da para la unión del resto 2-ABA tal como se muestra en la Figura 23. Demostrando de esta manera que una proteína con incorporación de PCL específica de sitio se modifica de forma específicamente del sitio en el sitio de la PCL. El análisis MS/MS (Figura 25) obtenido después del análisis LC-MS de la digestión trípica de la proteína hRBP4 Phe122PCL derivatizada con 2-ABA identificó el péptido YWGVASF*LQK, donde F* tiene una masa consistente con la de una PCL modificada con 2-ABA. La Figura 25A es el modelo de fragmentación del péptido YWGVASF*LQK, donde F* tiene una masa consistente con la de una PCL modificada con 2-ABA. La Figura 25B es el TIC (cromatograma iónico total) y EIC (cromatograma iónico extraído) de iones 2+ de YWGVASF*LQK (F* = PCL y aducto de PCL-2-ABA), donde la comparación de las EIC para las especies derivatizada y sin derivatizar (no detectable) indica el agotamiento de la reacción. La Figura 25C es el análisis por espectrometría de masas de hRBP4 Phe122PCL derivatizado con 2-ABA que muestra los precursores 3+ y 2+ de YWGVASF*LQK a m/z 459,92 (3+) y 689,37 (2+) respectivamente (F* = aducto

de PCL-2-ABA), demostrando por tanto que las reacciones observadas con 2-ABA se producen específicamente en el sitio con el resto de la PCL incorporado en el sitio TAG deseado en el resto 122.

En la Figura 26 se muestra la evaluación de la dependencia del pH de la derivatización. En la Figura 26A se muestra el espectro de masas de hRBP4 con la PCL incorporada en la posición 122 (hRBP4 Phe122PCL), donde la hRBP4 Phe122PCL no se hace reaccionar con 2-ABA, mientras que la Figura 26B y la Figura 26C son los espectros de masa de hRBP4 Phe122PCL tras la reacción 10 mM 2-ABA en tampón acetato, pH 5,0, y con 10 mM 2-ABA en tampón fosfato, pH 7,4, respectivamente. La reacción del resto de la PCL de la proteína hRBP4 con 2-ABA continúa rápidamente hasta que está completa en un 95 % a temperatura ambiente en tampón acuoso ajustado a pH 5, y en una extensión ligeramente inferior, a aproximadamente un 87 % a pH 7,4. La reacción con 2-ABA es selectiva para el resto de PCL como se observa en las Figuras 27D-F, donde no se observó reacción con 2-ABA para la hRBP4 marcada con O-metil-fenilalanina (OMePhe), un aminoácido no natural inerte, en la posición 62, y con un resto de fenilalanina natural (Phe) en la posición 122.

Evaluación de la eficacia de reacción como función de la relación de concentración de reactivo a proteína, y la reactividad con reactivos de tipo 2-ABA se muestra en la Figura 27. El espectro de masas de hRBP4 con PCL incorporada en la posición 122 (hRBP4 Phe122PCL) tras la reacción con 0,1 mM de 2-amino-benzaldehído (2-ABA) se muestra en la Figura 27A, mientras que los espectros de masas para hRBP4 Phe122PCL tras la reacción con 0,1 mM de 2-amino-acetofenona (2-AAP) y 0,1 mM de 2-amino-5-nitro-benzofenona (2-ANBP) se muestran en la Figura 27B y la Figura 27C, respectivamente. todas las reacciones se llevaron a cabo en acetato de sodio 200 mM, pH 5,0. Los conjugados de proteína se detectan a la masa correcta y con el aumento de masa esperada como se proporcióna en la Figura 23. Además, se consiguieron rendimientos de más de un 88 % de conversión para 2-ABA y 2-AAP, aunque debido a la baja solubilidad se obtuvo un rendimiento de aproximadamente el 5 % para 2-ANBP. Sin embargo, esto demuestra que los análogos de 2-ABA reaccionan eficazmente en el sitio de incorporación de la PCL a las relaciones de concentración de reactivo a proteína de 6 a 1. La extensión de las reacciones es solo ligeramente inferior que las llevadas a cabo a las relaciones de reactivo a proteína de 600 a 1 (Figura 26).

Cuando los agentes de derivatización se añadieron a un intervalo de concentración final entre 0,1 y 10 mM, que correspondían a entre 6 a 600 veces el exceso molar sobre la proteína, no hubo mejora significativa en la extensión de la derivatización. Además, a relaciones molares más grandes que 4700, restos adicionales de proteínas, presumiblemente lisinas, se derivatizaron mediante 2-ABA (Figura 28). En la Figura 28, un exceso grande de 2-ABA sobre la proteína en 10x PBS da como resultado la conjugación en sitios adicionales para la hRBP4 incorporada a PCL (A, ~17 mM, 4700 veces en exceso de 2-ABA sobre la proteína) y para hRBP4 incorporada a OMePhe (B, ~6,5 mM, 15400 veces de exceso) lo que ilustra que la conjugación en estos sitios adicionales está mediada por restos diferentes de la PCL.

Para ilustrar adicionalmente la utilidad de los métodos proporcionados en el presente documento, la Figura 29 muestra los espectros de masas antes y después de la derivatización de la PCL incorporada en FAS-TE con 2-amino-acetofenona (2-AAP) a pH 5,0 y pH 7,4 (véase, Ejemplo 12). La Figura 29A muestra el espectro de masas de FAS-TE Tyr2454PCL sin reaccionar, mientras que la Figura 29B muestra el espectro de masas de la mezcla de reacción a pH 5,0. Aquí se produce un 100 % de la intensidad del pico observable a 33318,8 Da, que es 116,8 Da más grande que la del material sin reaccionar, tal como se esperaba para el aumento de masa de 117 Da anticipado para FAS-TE Tyr2454PCL modificado con 2- AAP. De igual forma, a pH 7,4 la reacción avanza hasta un agotamiento del 95 % (Figura 29C (sin reaccionar) y Figura 29D (con reacción)).

La Figura 30 es un esquema de reacción general para la modificación específica de sitio de proteínas mediante la derivatización química de pirrolisina y/o PCL con 2-amino-benzaldehído o análogos de 2-amino-benzaldehído. En el presente documento se proporcionan determinadas realizaciones de dichos análogos. La reacción del anillo pirrolina en la pirrolisina y PCL con 2-ABA es similar a la reacción del ácido A¹-pirrolina-5-carboxílico con 2-ABA. De igual forma, las reacciones de pirrolisina y PCL con 2-AAP y 2- ANBP dan como resultado proteínas modificadas con los restos respectivos sustituidos de la Figura 23.

En determinadas realizaciones, la funcionalidad 2-ABA se usa para unir específicamente al sitio diversos grupos a proteínas, polipéptidos y/o péptidos que tienen pirrolisina y/o PCL incorporados en los anteriores. Dichos grupos incluyen, pero sin limitación, una marca, un colorante, un polímero, un polímero soluble en agua, un polialquilenglicol, un poli(etilenglicol), un derivado de poli(etilenglicol), un azúcar, un lípido, un fotorreticulador, un compuesto citotóxico, un fármaco, una marca de afinidad, una marca de fotoafinidad, un compuesto reactivo; una resina, un péptido, una segunda proteína o polipéptido o análogo de polipéptido, un anticuerpo o fragmento de anticuerpo, un quelante metálico, un cofactor, un ácido graso, un hidrato de carbono, un polinucleótido, un ADN, un ARN, una sonda de la PCR, un polinucleótido antisentido, un ribooligonucleótido, un desoxirribooligonucleótido, ADN modificado con fosforotioato, ADN y ARN modificados, ácido nucleico peptídico, un sacárido, un disacárido, un oligosacárido, un polisacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial, una nanopartícula, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un grupo funcional novedoso, un grupo que interactúa de forma covalente o no covalente con otras moléculas, un resto fotoestimulado, un resto excitable mediante radiación actínica, un ligando un resto fotoisomerizable, biotina, un análogo de biotina, un resto que incorpora un átomo pesado, un grupo escindible químicamente, un grupo fotoescindible, una cadena lateral alargada, un azúcar

unido a carbono, un agente redox activo, un aminotioácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo cromóforo, un grupo quimioluminiscente, un resto fluorescente, un grupo denso en electrones, un grupo magnético, un grupo intercalante, un grupo quelante, un cromóforo, un agente de transferencia de energía, un agente biológicamente activo, una marca detectable, una molécula pequeña, un ácido nucleico inhibidor, un ARNip, un radionucleótido, un agente de captura de neutrones, un derivado de biotina, punto(s) cuántico(s), un nanotransmisor, un radiotransmisor, una abzima, una enzima, un activador complejo activado, un virus, una toxina, un adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, un epítipo de linfocito T, un fosfolípido, una molécula de tipo LPS, hemocianina de lapa californiana (KLH), un hapteno inmunógeno, un aglicano, un alérgeno, una angioestatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimotopo, un receptor, una micela inversa, detergentes, potenciadores de la respuesta inmune, colorantes fluorescentes, reactivos FRET, sondas de radioimagen, otra sonda de espectroscopía, profármacos, toxinas para inmunoterapia, un soporte sólido, $-\text{CH}_2\text{CH}_2-(\text{OCH}_2\text{CH}_2\text{O})_p-\text{OX}^2$, $-\text{O}-(\text{CH}_2\text{CH}_2\text{O})_p\text{CH}_2\text{CH}_2\text{X}^2$, y cualquier combinación de los mismos, donde p es de 1 a 10.000 y X^2 es H, un alquilo C_{1-8} , un grupo protector o un grupo funcional terminal.

En determinadas realizaciones, la funcionalidad 2-AAP se usa para unir de manera específica del sitio diversos grupos a proteínas, polipéptidos y/o péptidos que tienen pirrolisina y/o PCL incorporados en los anteriores. Dichos grupos incluyen, pero sin limitación, una marca, un colorante, un polímero, un polímero soluble en agua, un polialquilenglicol, un poli(etilenglicol), un derivado de poli(etilenglicol), un azúcar, un lípido, un fotorreticulador, un compuesto citotóxico, un fármaco, una marca de afinidad, una marca de fotoafinidad, un compuesto reactivo; una resina, un péptido, una segunda proteína o polipéptido o análogo de polipéptido, un anticuerpo o fragmento de anticuerpo, un quelante metálico, un cofactor, un ácido graso, un hidrato de carbono, un polinucleótido, un ADN, un ARN, una sonda de la PCR, un polinucleótido antisentido, un ribooligonucleótido, un desoxirribooligonucleótido, ADN modificado con fosforotioato, ADN y ARN modificados, ácido nucleico peptídico, un sacárido, un disacárido, un oligosacárido, un polisacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial, una nanopartícula, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un grupo funcional novedoso, un grupo que interactúa de forma covalente o no covalente con otras moléculas, un resto fotoestimulado, un resto excitable mediante radiación actínica, un ligando un resto fotoisomerizable, biotina, un análogo de biotina, un resto que incorpora un átomo pesado, un grupo escindible químicamente, un grupo fotoescindible, una cadena lateral alargada, un azúcar unido a carbono, un agente redox activo, un aminotioácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo cromóforo, un grupo quimioluminiscente, un resto fluorescente, un grupo denso en electrones, un grupo magnético, un grupo intercalante, un grupo quelante, un cromóforo, un agente de transferencia de energía, un agente biológicamente activo, una marca detectable, una molécula pequeña, un ácido nucleico inhibidor, un ARNip, un radionucleótido, un agente de captura de neutrones, un derivado de biotina, punto(s) cuántico(s), un nanotransmisor, un radiotransmisor, una abzima, una enzima, un activador complejo activado, un virus, una toxina, un adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, un epítipo de linfocito T, un fosfolípido, una molécula de tipo LPS, hemocianina de lapa californiana (KLH), un hapteno inmunógeno, un aglicano, un alérgeno, una angioestatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimotopo, un receptor, una micela inversa, detergentes, potenciadores de la respuesta inmune, colorantes fluorescentes, reactivos FRET, sondas de radioimagen, otra sonda de espectroscopía, profármacos, toxinas para inmunoterapia, un soporte sólido, $-\text{CH}_2\text{CH}_2-(\text{OCH}_2\text{CH}_2\text{O})_p-\text{OX}^2$, $-\text{O}-(\text{CH}_2\text{CH}_2\text{O})_p\text{CH}_2\text{CH}_2\text{X}^2$, y cualquier combinación de los mismos, donde p es de 1 a 10.000 y X^2 es H, un alquilo C_{1-8} , un grupo protector o un grupo funcional terminal.

En determinadas realizaciones, la funcionalidad 2-ANPA se usa para unir específicamente al sitio diversos grupos a proteínas, polipéptidos y/o péptidos que tienen pirrolisina y/o PCL incorporados en los anteriores. Dichos grupos incluyen, pero sin limitación, una marca, un colorante, un polímero, un polímero soluble en agua, un polialquilenglicol, un poli(etilenglicol), un derivado de poli(etilenglicol), un azúcar, un lípido, un fotorreticulador, un compuesto citotóxico, un fármaco, una marca de afinidad, una marca de fotoafinidad, un compuesto reactivo; una resina, un péptido, una segunda proteína o polipéptido o análogo de polipéptido, un anticuerpo o fragmento de anticuerpo, un quelante metálico, un cofactor, un ácido graso, un hidrato de carbono, un polinucleótido, un ADN, un ARN, una sonda de la PCR, un polinucleótido antisentido, un ribooligonucleótido, un desoxirribooligonucleótido, ADN modificado con fosforotioato, ADN y ARN modificados, ácido nucleico peptídico, un sacárido, un disacárido, un oligosacárido, un polisacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial, una nanopartícula, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un grupo funcional novedoso, un grupo que interactúa de forma covalente o no covalente con otras moléculas, un resto fotoestimulado, un resto excitable mediante radiación actínica, un ligando un resto fotoisomerizable, biotina, un análogo de biotina, un resto que incorpora un átomo pesado, un grupo escindible químicamente, un grupo fotoescindible, una cadena lateral alargada, un azúcar unido a carbono, un agente redox activo, un aminotioácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo cromóforo, un grupo quimioluminiscente, un resto fluorescente, un grupo denso en electrones, un grupo magnético, un grupo intercalante, un grupo quelante, un cromóforo, un agente de transferencia de energía, un agente biológicamente activo, una marca detectable, una molécula pequeña, un ácido nucleico inhibidor, un ARNip, un radionucleótido, un agente de captura de neutrones, un derivado de biotina, punto(s) cuántico(s), un nanotransmisor, un radiotransmisor, una abzima, una enzima, un activador complejo activado, un virus, una toxina, un adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, un epítipo de linfocito T, un fosfolípido, una molécula de tipo LPS, hemocianina de lapa californiana

(KLH), un hapteno inmunógeno, un aglicano, un alérgeno, una angioestatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimotopo, un receptor, una micela inversa, detergentes, potenciadores de la respuesta inmune, colorantes fluorescentes, reactivos FRET, sondas de radioimagen, otra sonda de espectroscopía, profármacos, toxinas para inmunoterapia, un soporte sólido, $-\text{CH}_2\text{CH}_2-(\text{OCH}_2\text{CH}_2\text{O})_p-\text{OX}^2$, $-\text{O}-(\text{CH}_2-\text{CH}_2\text{O})_p\text{CH}_1\text{CH}_2-\text{X}^2$, y cualquier combinación de los mismos, donde p es de 1 a 10.000 y X^2 es H, un alquilo C_{1-8} , un grupo protector o un grupo funcional terminal.

Cualquiera de los puntos de unión R_7 a R_{11} que se muestran en la Figura 30 se usan para unir cualquiera de estos grupos anteriormente mencionados a 2-ABA, 2-AAP y 2-ANPA. En otras realizaciones, el anillo de benceno se sustituye por cualquier otra estructura de anillo proporcionada en el presente documento, incluyendo, pero sin limitarse a: naftaleno. En otras realizaciones, el anillo de benceno se sustituye por azúcares.

La concentración del agente de derivatización utilizado en los métodos proporcionados en el presente documento para la derivatización de pirrolisina y PCL, incorporadas a las proteínas, polipéptidos y/o péptidos, está en el intervalo de aproximadamente 0,005 mM a aproximadamente 50 mM. En determinadas realizaciones, la concentración del agente de derivatización utilizado en los métodos proporcionados en el presente documento para la derivatización de pirrolisina o PCL está en el intervalo de aproximadamente 0,005 a aproximadamente 25 mM. En determinadas realizaciones, la concentración del agente de derivatización utilizado en los métodos proporcionados en el presente documento para la derivatización de pirrolisina o PCL está en el intervalo de aproximadamente 0,005 a aproximadamente 10 mM. En determinadas realizaciones, la concentración del agente de derivatización utilizado en los métodos proporcionados en el presente documento para la derivatización de pirrolisina o PCL está en el intervalo de aproximadamente 0,005 a aproximadamente 5 mM. En determinadas realizaciones, la concentración del agente de derivatización utilizado en los métodos proporcionados en el presente documento para la derivatización de pirrolisina o PCL está en el intervalo de aproximadamente 0,005 a aproximadamente 2 mM. En determinadas realizaciones, la concentración del agente de derivatización utilizado en los métodos proporcionados en el presente documento para la derivatización de pirrolisina o PCL está en el intervalo de aproximadamente 0,005 a aproximadamente 1 mM.

En los métodos de derivatización proporcionados en el presente documento, la relación molar del agente de derivatización con respecto a las proteínas, polipéptidos o péptidos que tienen uno o más pirrolisina o PCL incorporadas en las anteriores, está en el intervalo de aproximadamente 0,05 a aproximadamente 10000. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 0,05 a aproximadamente 8000. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 0,05 a aproximadamente 6000. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 0,05 a aproximadamente 4000. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 0,05 a aproximadamente 2000. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 0,05 a aproximadamente 1000. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 0,05 a aproximadamente 800. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 0,05 a aproximadamente 600. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 0,05 a aproximadamente 400. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 0,05 a aproximadamente 200. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 0,05 a aproximadamente 100. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 0,05 a aproximadamente 50. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 0,05 a aproximadamente 25. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 0,05 a aproximadamente 10. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 0,05 a aproximadamente 1.

En otras realizaciones adicionales, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 1,5 a aproximadamente 10000. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 1,5 a aproximadamente 8000. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 1,5 a aproximadamente 6000. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 1,5 a aproximadamente 4000. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 1,5 a aproximadamente 2000. En

determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 1,5 a aproximadamente 1000. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 1,5 a aproximadamente 800. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 1,5 a aproximadamente 600. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 1,5 a aproximadamente 400. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 1,5 a aproximadamente 200. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 1,5 a aproximadamente 100. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 1,5 a aproximadamente 50. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 1,5 a aproximadamente 25. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 1,5 a aproximadamente 10. En determinadas realizaciones, la relación molar del agente de derivatización con respecto a dichas proteínas, polipéptidos o péptidos está en el intervalo de aproximadamente 6 a aproximadamente 600.

El pH de la disolución tampón utilizada en la derivatización de proteínas, polipéptidos y/o péptidos usando los métodos y composiciones que se proporcionan en el presente documento es de aproximadamente pH 2 a aproximadamente pH 10. En determinadas realizaciones, el pH es de aproximadamente pH 4 a aproximadamente pH 8. En determinadas realizaciones, el pH es de aproximadamente pH 4 a aproximadamente pH 7,5. En determinadas realizaciones, el pH es de aproximadamente pH 4 a aproximadamente pH 7. En determinadas realizaciones, el pH es de aproximadamente pH 4 a aproximadamente pH 6,5. En determinadas realizaciones, el pH es de aproximadamente pH 4 a aproximadamente pH 6. En determinadas realizaciones, el pH es de aproximadamente pH 4 a aproximadamente pH 5,5. En determinadas realizaciones, el pH es de aproximadamente pH 4 a aproximadamente pH 5. En determinadas realizaciones, el pH es de aproximadamente pH 4 a aproximadamente pH 4,5. En determinadas realizaciones, el pH es de aproximadamente pH 6 a aproximadamente pH 8. En determinadas realizaciones, el pH es de aproximadamente pH 6 a aproximadamente pH 7,5. En determinadas realizaciones, el pH es de aproximadamente pH 6 a aproximadamente pH 7. En determinadas realizaciones, el pH es de aproximadamente pH 6 a aproximadamente pH 6,5. En determinadas realizaciones, el pH es de aproximadamente pH 7 a aproximadamente pH 8. En determinadas realizaciones, el pH es de aproximadamente pH 7 a aproximadamente pH 7,5.

Los métodos de derivatización proporcionados en el presente documento son útiles para la derivatización de la pirrolisina o la PCL incorporada específicamente al sitio en las proteínas. En determinadas realizaciones, dichos métodos de derivatización se usan para la derivatización de un resto de PCL incorporado a un sitio único en una proteína, polipéptido y/o péptido. En otras realizaciones, dichos métodos de derivatización se usan para la derivatización de un resto de PCL incorporado a múltiples sitios en una proteína, polipéptido y/o péptido. En determinadas realizaciones, la proteína, polipéptidos y/o péptidos derivatizados usando los métodos proporcionados en el presente documento contienen 1, 2, 3, 4, 5, 6, 7, 6, 9, 10, 11, 12, 13, 14, 15 o más pirrolisinas o análogos de pirrolisina.

En otro aspecto proporcionado en el presente documento, las proteínas, polipéptidos y péptidos que tienen una o más pirrolisinas o PCL incorporadas en la anterior se derivatizan mediante al menos una modificación simultánea de la traducción o posterior a la traducción. Los ejemplos no limitantes de dichas modificaciones simultáneas de la traducción o posteriores a la traducción incluyen, pero no se limitan a, glicosilación, acetilación, acilación, metilación, nitración, sulfatación, modificación de lípidos, palmitoilación, adición de palmitato, fosforilación, modificación del enlace glicolípido, y similares. Incluidos también con este aspecto están los procedimientos para producir, purificar, caracterizar y usar dichas proteínas, polipéptidos y péptidos que contienen al menos una modificación simultánea de la traducción o una modificación posterior a la traducción.

Agentes bioterapéuticos con modificaciones específicas de sitio

Polímeros macromoleculares acoplados a PCL y pirrolisina incorporados en proteínas, polipéptidos y/o péptidos

En determinadas realizaciones, las composiciones, métodos, técnicas y estrategias descritas en el presente documento, se usan para añadir polímeros macromoleculares a restos de pirrolisina o PCL incorporados en proteínas, polipéptidos y/o péptidos. Se pueden acoplar una amplia variedad de polímeros macromoleculares acoplados a restos de pirrolisina y PCL incorporados en las proteínas, polipéptidos y/o péptidos descritos en el presente documento. Dichas modificaciones se usan para modular las propiedades biológicas de dichas proteínas, polipéptidos y/o péptidos, y/o proporcionar nuevas propiedades biológicas a dichas proteínas, polipéptidos y/o péptidos. En determinadas realizaciones, los polímeros macromoleculares se acoplan a las proteínas, polipéptidos y/o péptidos en el presente documento mediante acoplamiento directo con el(los) resto(s) de pirrolisina o PCL incorporados en dichas proteínas, polipéptidos y/o péptidos. En otras realizaciones, los polímeros macromoleculares se acoplan a las proteínas, polipéptidos y/o péptidos proporcionados en el presente documento mediante enlazadores bi, tri, tetra, y

5 polifuncionales acoplados al(a los) resto(s) de pirrolisina o PCL. En otras realizaciones, los polímeros macromoleculares se acoplan a las proteínas, polipéptidos y/o péptidos proporcionados en el presente documento mediante enlazadores bifuncionales acoplados al (a los) resto(s) de pirrolisina o PCL. En determinadas realizaciones, dichos enlazadores bi, tri, tetra y polifuncionales son enlazadores monofuncionales donde todos los extremos terminales son sustituyentes que son específicos para la reacción con restos de pirrolisina o PCL. En determinadas realizaciones, dichos enlazadores bi, tri, tetra y polifuncionales son enlazadores heterobifuncionales donde uno o más extremos terminales son un sustituyente que es específico para la reacción con resto(s) de pirrolisina o PCL, aunque los otros extremos terminales son otros sustituyentes funcionales que no reaccionan con resto(s) de pirrolisina o PCL. Se proporcionan en el presente documento determinados sustituyentes que son grupos reactivos específicos de pirrolisina o PCL, y los otros sustituyentes y los enlaces resultantes, que se usan en dichos enlazadores bi, tri, tetra y polifuncionales incluyen, pero no se limitan a, los relacionados en la Tabla 1.

Tabla 1

Electrófilos	Nucleófilos	Enlace covalente
Ésteres activados	Aminas/anilinas	Carboxamidas
Azidas de acilo	Aminas/anilinas	Carboxamidas
Haluros de acilo	Aminas/anilinas	Carboxamidas
Haluros de acilo	Alcoholes/fenoles	Ésteres
Nitrilos de acilo	Alcoholes/fenoles	Ésteres
Nitrilos de acilo	Aminas/anilinas	Carboxamidas
Aldehídos	Aminas/anilinas	Iminas
Aldehídos o cetonas	Hidrazinas	hidrazonas
Aldehídos o cetonas	Hidroxilaminas	Oximas
Haluros de alquilo	Aminas/anilinas	Alquilaminas
Haluros de alquilo	Ácidos carboxílicos	Ésteres
Haluros de alquilo	Tioles	Tioéteres
Haluros de alquilo	Alcoholes/fenoles	Éteres
Sulfonatos de alquilo	Tioles	Tioéteres
Sulfonatos de alquilo	Ácidos carboxílicos	Ésteres
Sulfonatos de alquilo	Alcoholes/fenoles	Éteres
Anhídridos	Alcoholes/fenoles	Ésteres
Anhídridos	Aminas/anilinas	Carboxamidas
Haluros de arilo	Tioles	Tiofenoles
Haluros de arilo	Aminas	Arilaminas
Azindinas	Tioles	Tioéteres
Boronatos	Glicoles	Ésteres de boronato
Ácidos carboxílicos	Aminas/anilinas	Carboxamidas
Ácidos carboxílicos	Alcoholes	Ésteres
Hidrazidas	Ácidos carboxílicos	Hidrazinas
Carbodiimidias	Ácidos carboxílicos	N-acilureas o anhídridos
Diazoalcanos	Ácidos carboxílicos	Ésteres
Epóxidos	Tioles	Tioéteres
Haloacetamidas	Tioles	Tioéteres
Halotriazinas	Aminas/anilinas	Aminotriazinas
Halotriazinas	Alcoholes/fenoles	Éteres de triazinilo
Imidoésteres	Aminas/anilinas	Amidinas
Isocianatos	Aminas/anilinas	Ureas
Isocianatos	Alcoholes/fenoles	Iretanos
Isotiocianatos	Aminas/anilinas	Tioureas
Maleimidias	Tioles	Tioéteres
Fosforamiditas	Alcoholes	Ésteres de fosfato
Haluros de sililo	Alcoholes	Éteres de sililo
Ésteres de sulfonato	Aminas/anilinas	Alquilaminas
Ésteres de sulfonato	Tioles	Tioéteres
Ésteres de sulfonato	Ácidos carboxílicos	Ésteres
Ésteres de sulfonato	Alcoholes	Éteres
Haluros de sulfonilo	Aminas/anilinas	Sulfonamidas
Haluros de sulfonilo	Alcoholes/fenoles	Ésteres de sulfonato

15 La unión covalente de polímeros macromoleculares a una molécula biológicamente activa, tal como las proteínas, polipéptidos y/o péptidos proporcionados en el presente documento, representa una solución para aumentar la solubilidad en agua (tal como en un entorno fisiológico), la biodisponibilidad, el aumento de la semivida en suero, el aumento de la semivida terapéutica, la modulación de la inmunogenicidad, la modulación de la actividad biológica, o extender el tiempo en circulación de dichas moléculas biológicamente activas. Las características importantes de

dichos polímeros macromoleculares incluyen la biocompatibilidad, la ausencia de toxicidad, y la ausencia de inmunogenicidad, y para los usos terapéuticos de las proteínas, polipéptidos y/o péptidos proporcionados en el presente documento que se acoplan a polímeros macromoleculares mediante un resto de pirrolisina o PCL, dichos polímeros macromoleculares son farmacéuticamente aceptables.

5 Determinados polímeros macromoleculares acoplados a resto(s) de pirrolisina o PCL incorporados en proteínas, polipéptidos y/o péptidos descritos en el presente documento son polímeros solubles en agua. En determinadas realizaciones, dichos polímeros solubles en agua están acoplados mediante resto(s) de pirrolisina o PCL incorporados a proteínas, polipéptidos y/o péptidos proporcionados en el presente documento, aunque en otras realizaciones, dichos polímeros solubles en agua están acoplados a las proteínas, polipéptidos y/o péptidos proporcionados en el presente documento mediante cualquier grupo funcional o sustituyente acoplado a resto(s) de pirrolisina o PCL incorporados en dichas proteínas, polipéptidos y/o péptidos. En algunas realizaciones, las proteínas, polipéptidos y/o péptidos proporcionados en el presente documento incluyen uno o más restos de PCL acoplados a polímeros solubles en agua y uno o más aminoácidos que se producen naturalmente unidos a polímeros solubles en agua.

15 Las formas estructurales de los polímeros macromoleculares acoplados a resto(s) de pirrolisina o PCL incorporados en proteínas, polipéptidos y/o péptidos incluyen, pero no se limitan a, polímeros lineales, ahorquillados o ramificados. En determinadas realizaciones, las estructuras de dichos polímeros solubles en agua ramificados o ahorquillados tienen entre 2 y aproximadamente 300 extremos terminales. En determinadas realizaciones, cada extremo terminal de dichos derivados de polímeros multifuncionales, incluyendo, pero sin limitación, los polímeros lineales que tienen dos extremos terminales, incluye un grupo funcional. En determinadas realizaciones dichos grupos funcionales son los mismos, mientras que en otras realizaciones dichos grupos funcionales son diferentes. Los ejemplos no limitantes de dichos grupos funcionales terminales incluyen, pero no se limitan a, N-succinimidil carbonatos, aminas, hidrazidas, propionatos de succinimidilo y butanoatos de succinimidilo, succinatos de succinimidilo, ésteres de succinimidilo, carbonatos de benzotriazol, éteres de glicidilo, oxicarbonilimidazoles, carbonatos de p-nitrofenilo, aldehídos, maleimidas, ortopiridil-disulfuros, acriloles y vinilsulfonas. En otras realizaciones, los grupos funcionales incluyen los relacionados en la Tabla 1.

30 La unión covalente de polímeros solubles en agua (denominados también en el presente documento como polímeros hidrófilo) a una molécula biológicamente activa, tal como las proteínas, polipéptidos y/o péptidos proporcionados en el presente documento, representa una solución para aumentar la solubilidad en agua (tal como en un entorno fisiológico), la biodisponibilidad, el aumento de la semivida en suero, el aumento de la semivida terapéutica, modular la inmunogenicidad, la modulación de la actividad biológica, o extender el tiempo en circulación de dichas moléculas biológicamente activas. Las características importantes de dichos polímeros solubles en agua incluyen la biocompatibilidad, la ausencia de toxicidad, y la ausencia de inmunogenicidad, y para los usos terapéuticos de las proteínas, polipéptidos y/o péptidos proporcionados en el presente documento que se acoplan a polímeros solubles en agua mediante restos de pirrolisina o PCL, dichos polímeros macromoleculares son farmacéuticamente aceptables.

40 Los polímeros hidrófilos acoplados a proteínas, polipéptidos y/o péptidos mediante restos de pirrolisina y/o PCL incluyen, pero no se limitan a, éteres de polialquilo y sus análogos con alcoxi protegido, polivinilpirrolidonas, éteres de polivinilalquilo, polioxazolinas, polialquioxazolinas, polihidroxiálquioxazolinas, poliácridamidas, polialquilacrilamidas, polihidroxiálquilacrilamidas, acrilatos de polihidroxiálquilo, ácidos polisiálicos y sus análogos, secuencias peptídicas hidrófilas; polisacáridos y sus derivados, celulosa y sus derivados, quitina y sus derivados, ácido hialurónico y sus derivados, almidones, alginatos, sulfato de condroitina, albúmina, pululano, carboximetil pululano, poliaminoácidos y sus derivados, copolímeros de anhídrido maleico, alcoholes polivinílicos y sus copolímeros, alcoholes polivinílicos y sus terpolímeros, y sus combinaciones.

50 Dichos éteres de polialquilo y sus análogos con alcoxi protegido incluyen, pero no se limitan a, polioxietilenglicol (conocido también como poli(etilenglicol) o PEG), polioxietilén/propilenglicol, y sus análogos con metoxi o etoxi protegido. Dichas polihidroxiálquilacrilamidas incluyen, pero no se limitan a, polihidroxiopropilmetacrilamida y sus derivados. Dichos polisacáridos y sus derivados incluyen, pero no se limitan a, dextrano y derivados de dextrano (tales como, a modo de ejemplo solamente, carboximetildextrano, sulfatos de dextrano y aminodextrano). Dichas celulosas y sus derivados incluyen, pero no se limitan a, carboximetilcelulosa e hidroxialquilcelulosas. Dicha quitina y sus derivados incluyen, pero no se limitan a, quitosán, succinil quitosán, carboximetilquitina y carboximetilquitosán. Dichos poliaminoácidos y sus derivados, incluyen, pero no se limitan a, ácidos poliglutámicos, polilisinas, ácidos poliaspárticos y poliaspartamidas. Dichos copolímeros de anhídrido maleico incluyen, pero no se limitan a, copolímero de estireno anhídrido maleico y copolímero de éter de diviniletilo anhídrido maleico.

60 En algunas realizaciones, el polímero soluble en agua acoplado directa o indirectamente a proteínas, polipéptidos y/o péptidos proporcionados en el presente documento mediante acoplamiento directo con el(los) resto(s) de pirrolisina o PCL incorporados a dichas proteínas, polipéptidos y/o péptidos es un poli(etilenglicol) (PEG). Se considera que el poli(etilenglicol) es biocompatible, donde PEG es capaz de coexistir con tejidos u organismos vivos sin producir daño. PEG es también sustancialmente no inmunógeno y por tanto no muestra tendencia a producir una respuesta inmune en el cuerpo. PEG es un polímero hidrófilo que se ha usado extensamente en compuestos farmacéuticos, para implantar artificiales, y en otras aplicaciones donde la biocompatibilidad, la ausencia de toxicidad, y la ausencia de inmunogenicidad son de importancia. Los conjugados de PEG no muestran tendencia a producir una respuesta

inmune sustancial o producir coagulación u otros efectos indeseables. Por tanto, cuando se unen a una molécula que tiene alguna función deseable en el cuerpo, tal como un agente biológicamente activo, PEG muestra tendencia a enmascarar el agente y puede reducir o eliminar cualquier respuesta inmune de tal manera que un organismo puede tolerar la presencia del agente.

5 En algunas realizaciones, el poli(etilenglicol) acoplado directa o indirectamente a proteínas, polipéptidos y/o péptidos proporcionados en el presente documento mediante acoplamiento directo con el(los) resto(s) de pirrolisina o PCL incorporados a dichas proteínas, polipéptidos y/o péptidos es un polímero de cadena lineal, mientras que en otras realizaciones dichos polímeros de PEG son polímeros ramificados. El peso molecular o la distribución de pesos
10 moleculares de dichos polímeros PEG de cadena lineal y ramificada está comprendido entre aproximadamente 100 Da y aproximadamente 100.000 Da o más. En algunas realizaciones, el peso molecular o la distribución de pesos moleculares de dichos polímeros PEG está comprendida entre aproximadamente 100 y aproximadamente 50.000 Da. En algunas realizaciones, el peso molecular o la distribución de pesos moleculares de dichos polímeros PEG está comprendida entre aproximadamente 100 y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular o la distribución de pesos moleculares de dichos polímeros PEG está comprendida entre aproximadamente 1.000 y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular o la distribución de pesos moleculares de dichos polímeros PEG está comprendida entre aproximadamente 5.000 y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular o la distribución de pesos moleculares de dichos polímeros PEG está comprendida entre aproximadamente 10.000 y aproximadamente 40.000 Da. En determinadas realizaciones, el peso molecular de dichos polímeros PEG es 100.000 Da, 95.000 Da, 90.000 Da, 85.000 Da, 80.000 Da, 75.000 Da, 70.000 Da, 65.000 Da, 60.000 Da, 55.000 Da, 50.000 Da, 45.000 Da, 40.000 Da, 35.000 Da, 30.000 Da, 25.000 Da, 20.000 Da, 15.000 Da, 10.000 Da, 9.000 Da, 8.000 Da, 7.000 Da, 6.000 Da, 5.000 Da, 4.000 Da, 3.000 Da, 2.000 Da, 1.000 Da, 900 Da, 800 Da, 700 Da, 600 Da, 500 Da, 400 Da, 300 Da, 200 Da, o 100 Da.

25 Las formas estructurales de los polímeros PEG acoplados a resto(s) de pirrolisina o PCL incorporados a proteínas, polipéptidos y o péptidos incluyen, pero no se limitan a, polímeros lineales, ahorquillados o ramificados. En determinadas realizaciones, las estructuras de dichos polímeros ramificados o ahorquillados tienen entre 2 y aproximadamente 300 extremos terminales. En determinadas realizaciones, cada extremo terminal de dichos derivados de polímeros multifuncionales, incluyendo, pero sin limitación, los polímeros lineales que tienen dos
30 extremos terminales, incluye un grupo funcional. En determinadas realizaciones dichos grupos funcionales son los mismos, mientras que en otras realizaciones al menos uno de dichos grupos funcionales es diferente. En otras realizaciones más dichos grupos funcionales son diferentes.

35 En determinadas realizaciones antes del acoplamiento con las proteínas, polipéptidos y/o péptidos, uno o más extremos terminales del polímero de PEG incluyen un grupo funcional. En determinadas realizaciones, el PEG es un polímero lineal con un extremo terminal que tiene un grupo funcional, donde en otras realizaciones el PEG es un polímero lineal con un extremo terminal que tiene un grupo funcional, formando por tanto un polímero PEG bifuncional. En otras realizaciones, el PEG es un polímero ahorquillado con un extremo terminal que tiene un grupo funcional, donde en otras realizaciones el PEG es un polímero ahorquillado con dos o más extremos terminales que tienen un
40 grupo funcional. En otras realizaciones adicionales, el PEG es un polímero ahorquillado con cada extremo terminal que tiene un grupo funcional, formando por tanto un polímero PEG multifuncional. En otras realizaciones, el PEG es un polímero ramificado con un extremo terminal que tiene un grupo funcional, donde en otras realizaciones el PEG es un polímero ramificado con dos o más extremos terminales que tienen un grupo funcional. En otras realizaciones adicionales, el PEG es un polímero ramificado con cada extremo terminal que tiene un grupo funcional, formando por tanto un polímero PEG multifuncional. En determinadas realizaciones de dichos polímeros PEG anteriormente mencionados, los grupos funcionales son los mismos, mientras que en otras realizaciones de dichos grupos funcionales al menos un grupo funcional es diferente. En otras realizaciones adicionales de dichos polímeros PEG anteriormente mencionados, los grupos funcionales son diferentes. Sin embargo, al menos un extremo terminal de los polímeros PEG está disponible para la reacción con al menos un resto de pirrolisina o PCL incorporado en una
50 proteína, polipéptido y/o péptido.

Los ejemplos no limitantes de grupos funcionales terminales incluyen, pero no se limitan a, N-succinimidil carbonatos, aminas, hidrazidas, propionatos de succinimidilo, butanoatos de succinimidilo, succinatos de succinimidilo, ésteres de succinimidilo, carbonatos de benzotriazol, éteres de glicidilo, oxicarbonilimidazoles, carbonatos de p-nitrofenilo, aldehídos, maleimidas, ortopiridil-disulfuros, acríloles, vinilsulfonas, carbonatos activados (incluyendo, pero sin limitarse a, éster de p-nitrofenilo), ésteres activados (incluyendo, pero sin limitarse a, N-hidroxisuccinimida, éster de p-nitrofenilo), oximas, carbonilos, dicarbonilos, hidroxilaminas, hidroxilo, metoxi, benzaldehídos, acetofenonas, 2-amino-benzaldehídos, 2-aminoacetofenonas y 2-amino-5-nitro-benzofenonas.

60 La Figura 31 ilustra una realización de un polímero PEG funcionalizado, 2-amino-acetofenonas-PEG₈ (2-AAP- PEG₈; TU3205-044), acoplado a proteínas mediante un resto de PCL incorporada en dichas proteínas. Tal como se muestra, el restos de 2- amino-acetofenona de 2-AAP-PEG₈ forma un resto de tipo quinazolina como un separador entre el PEG₈ y la proteína. Este resto de tipo quinazolina puede reaccionar además para formar la estructura de anillo fusionado. En el caso del acoplamiento de 2-AAP-PEG₈, dicha derivatización añade 556 Da a la masa de la proteína.
65 Otros PEG funcionalizados utilizados en los métodos proporcionados en el presente documento incluyen, pero no se limitan a, los proporcionados en el Ejemplo 20.

La Figura 32 es el espectro de masas de hRBP4 con PCL incorporada en la posición 122 (hRBP4 Phe122PCL) antes y después de la derivatización con 2-AAP-PEG8 a pH 7,5 (Figura 32A) y pH 5,0 (Figura 32B). Estas reacciones de acoplamiento se llevaron a cabo a pH 7,5 y pH 5, donde se produce una conversión del 89 al 100 % dando como resultado un aumento de la masa esperada de 556 Da con respecto a la proteína sin reaccionar (Figura 32C). hRBP4 natural no reacciona con un exceso de 450 o 2300 veces de 2-AAP-PEG8 (Figura 32D-F). Esto ilustra adicionalmente la especificidad de la reacción de acoplamiento entre PCL y 2AAP-PEG8, ya que no se observa acoplamiento cuando PCL está ausente.

Para ilustrar adicionalmente la utilidad de este método de marcado, el dominio de la tioesterasa de la ácido graso sintetasa humana (FAS- TE) con PCL incorporada en la posición 2454 (FAS-TE Tyr2454PCL) producida en *Escherichia coli* se derivatizó con 2-AAP-PEG8 (Figura 33, véase el Ejemplo 14). La Figura 33A muestra el espectro de masas de la proteína sin reaccionar (masa 33202 Da), mientras que la Figura 33B muestra el espectro de masas de FAS-TE Tyr2454PCL que ha reaccionado hasta la finalización con 2-AAP-PEG8 (TU3205-044) proporcionando la masa esperada de 33756 Da. Se proporciona una ilustración la Figura 33C y la Figura 33D que muestra los espectros de masas de FAS-TE con PCL incorporado en la posición 2454 (FAS-TE Tyr2454PCL) producida en *Escherichia coli* derivatizada con 2,4 kDa 2-AAP-PEG (TU3205-048) a temperatura ambiente (Figura 33C) y a 4°C (Figura 33D). En las condiciones utilizadas solo se produjo una conversión aproximadamente del 25 % para obtener la proteína con la masa esperada de 35585 Da.

Además, La Figura 34 muestra la PEGilación de FAS-TE Tyr2454PCL con 2,4 kDa 2-AAP-PEG y 23 kDa 2-AAP-PEG a las relaciones molares que se muestran. El producto pegilado 2-AAP-PEG (2-AAP-PEG8) de 0,5 kDa no se pudo resolver utilizando SDS-PAGE, sin embargo, se verificó utilizando espectroscopía de masas.

Los métodos proporcionados en el presente documento se han usado para incorporar PCL en veinte posiciones en el factor de crecimiento de fibroblastos 21 (FGF-21) y se han usado posteriormente para PEGILAR los restos de PCL correspondientes (véase el Ejemplo 15). La Figura 35 muestra los espectros de masas obtenidos antes y durante la derivatización FGF21 Lys84PCL con 2-AAP-PEG8. La Figura 35A es el espectro de masas de FGF21 Lys84PCL sin reaccionar (21235,6 Da), mientras que la Figura 35B es el espectro de masas de FGF21 Lys84PCL que ha reaccionado con 2-AAP-PEG8. La reacción de PEGilación continuó hasta el agotamiento y dio como resultado una proteína con 21792,4 Da. El aumento de 556,8 Da está de acuerdo con el aumento de 556 Da esperado para la derivatización. La Figura 36 muestra los resultados de SDS-PAGE obtenidos tras la derivatización de siete de los mutantes FGF21 PCL con 2-AAP-PEG de 23 kDa. La Figura 36A es el gel SDS de las mezclas de reacción de los mutantes FGF21 PCL (entre 0,1 y 0,4 mM) que ha reaccionado con 2-AAP-PEG de 23 kDa (pH 7,4, 4°C, 60 horas) y muestra PEG-FGF21, (FL) FGF21-PCL de longitud completa FGF21-PCL (TR) y truncada. La Figura 36B muestra los resultados de SDS-PAGE de ocho mutantes FGF21 PCL tras la purificación parcial de PEG-FGF21, de longitud completa (FL) y FGF21 (TR) y truncada.

Los métodos proporcionados en el presente documento se han usado para incorporar PCL en once posiciones en eritropoyetina de ratón (EPO) y se han usado posteriormente para PEGILAR los restos de PCL correspondientes de los tres mutantes EPO PCL (véase el Ejemplo 6). La Figura 37 es un gel SDS obtenido tras la PEGilación de mutantes EPO PCL de ratón. Se hizo reaccionar EPO de ratón con PCL incorporada en células HEK293F en tres posiciones diferentes con 2-AAP-mPEG23k (TU3205-052). La EPO PEGilada y la EPO no PEGilada se indican mediante flechas.

En determinadas realizaciones, polímeros PEG bifuncionales o multifuncionales unidos a una proteína, polipéptido y/o péptido mediante resto(s) PCL se derivatizan o sustituyen adicionalmente utilizando los grupos funcionales restantes en los extremos terminales sin reaccionar. En determinadas realizaciones, polímeros PEG bifuncionales o multifuncionales unidos a una proteína, polipéptido y/o péptido mediante la reacción de un resto benzaldehído, benzofenona o acetofenona con resto(s) PCL se derivatizan o sustituyen adicionalmente utilizando los grupos funcionales restantes en los extremos terminales sin reaccionar.

Los métodos y composiciones descritos en el presente documento proporcionan un método muy eficaz para la modificación selectiva de proteínas, polipéptidos y péptidos con derivados PEG, que implica la incorporación selectiva del aminoácido PCL o pirrolisina a las proteínas en respuesta a un codón selector y la posterior modificación de los restos de aminoácidos correspondientes. Por tanto, se proporcionan en el presente documento proteínas, polipéptidos y/o péptidos que tienen restos(s) de PCL o pirrolisina incorporados en los anteriores, y se proporcionan también en el presente documento las mencionadas proteínas, polipéptidos y/o péptidos unidos a polímeros solubles en agua, tales como poli(etilenglicoles) (PEG), mediante dichos restos(s) de PCL o pirrolisina.

La extensión de la PEGilación a una proteína, polipéptido y/o péptido (que es el número de polímeros PEG unidos a una proteína, polipéptido y/o péptido) descrito en el presente documento es ajustable para proporcionar características farmacológicas, farmacocinéticas o farmacodinámicas alteradas. En determinadas realizaciones, dichas alteraciones son aumentos en dichas características, mientras que en otras realizaciones, dichas alteraciones son disminuciones en dichas características. En determinadas realizaciones, dicha característica es la semivida in vivo. En algunas realizaciones, la semivida de una proteína, polipéptido y/o péptido PEGilado usando los métodos y composiciones proporcionados en el presente documento está aumentada al menos aproximadamente 10, 20, 30, 40, 50, 60, 70, 80, 90 por ciento, dos veces, cinco veces, 10 veces, 50 veces, 100 veces o al menos aproximadamente 200 veces con

respecto a una proteína, polipéptido y/o péptido sin modificar.

5 En determinadas realizaciones, las proteínas, polipéptidos y/o péptidos PEGilados usando los métodos y composiciones proporcionados en el presente documento se purifican usando métodos que incluyen, pero sin limitación, cromatografía hidrófoba, cromatografía de afinidad; cromatografía de intercambio aniónico, cromatografía de intercambio catiónico, resina de amonio cuaternario Q, DEAE SEPHAROSE, cromatografía sobre gel de sílice; HPLC en fase inversa, filtración en gel (utilizando, incluyendo pero sin limitarse a, SEPHADEX G-75, Sephacryl S-100, SUPERDEX 75, 100 y 200), cromatografía de interacción hidrófoba, cromatografía de exclusión por tamaño, cromatografía de quelatos metálicos; ultrafiltración/diafiltración, precipitación con etanol, cromatofocalización, cromatografía por desplazamiento, procedimientos electroforéticos (que incluyen, pero no se limitan a focalización isoelectrica preparativa), solubilidad diferencial (que incluye, pero no se limita a precipitación con sulfato de amonio), y extracción.

15 En otros aspectos, los polímeros PEG usados en los métodos proporcionados en el presente documento tienen enlaces débiles o degradables en la estructura. Por medio solo de ejemplo, dichos polímeros PEG tienen enlaces éster en la estructura polimérica que están sujetos a hidrólisis, y esta hidrólisis da como resultado la escisión del enlace para liberar la proteína, el polipéptido o el péptido al cual se une PEG.

20 En otros aspectos, los métodos y composiciones descritos en el presente documento se usan para producir preparaciones sustancialmente homogéneas de conjugados de polímero-proteína. "Sustancialmente homogéneas" tal como se usa en el presente documento significa que se observa que las moléculas conjugadas de polímero:proteína son mayores que la mitad de la proteína total. El conjugado de polímero:proteína tiene actividad biológica y las presentes preparaciones de polipéptido PEGilado "sustancialmente homogéneo" proporcionadas en el presente documento son aquellas que son suficientemente homogéneas para mostrar las ventajas de una preparación homogénea, tales como, a modo de ejemplo solamente, facilitar la aplicación clínica en la predictibilidad de la farmacocinética entre lotes.

30 En otros aspectos, los métodos y composiciones descritos en el presente documento se usan para producir preparaciones sustancialmente homogéneas de conjugados de polímero-proteína. "Sustancialmente homogéneas" tal como se usa en el presente documento significa que se observa que las moléculas conjugadas de polímero:proteína son mayores que la mitad de la proteína total. El conjugado de polímero:proteína tiene actividad biológica y las presentes preparaciones de polipéptido PEGilado "sustancialmente homogéneo" proporcionadas en el presente documento son aquellas que son suficientemente homogéneas para mostrar las ventajas de una preparación homogénea, tales como, a modo de ejemplo solamente, facilitar la aplicación clínica en la predictibilidad de la farmacocinética entre lotes.

40 En otro aspecto, los métodos y composiciones descritos en el presente documento se usan para preparar una mezcla de moléculas de conjugados de polímero-proteína, y la ventaja proporcionada en el presente documento es que la proporción de conjugado de monopolímero:proteína para incluir en la mezcla es seleccionable. Por tanto, en determinadas realizaciones, se prepara una mezcla de diversas proteínas con diversas cantidades de restos de polímeros unidos (es decir, bi, tri, tetra, etc.) y estos conjugados se combinan opcionalmente con un conjugado de monopolímero-proteína preparado utilizando los métodos descritos en el presente documento, proporcionando por tanto una mezcla con una proporción predeterminada de conjugados de monopolímero:proteína.

45 La proporción de moléculas de poli(etilenglicol) a moléculas de proteína variará, así como sus concentraciones en la mezcla de reacción. En determinadas realizaciones, la relación óptima (en términos de eficacia de reacción en que existe un mínimo exceso de proteína o polímero sin reaccionar) se determina mediante el peso molecular del poli(etilenglicol) seleccionado y el número de grupos reactivos disponibles. Cuanto mayor es el peso molecular del polímero, menor es el número de moléculas de polímero que se unen a la proteína. De igual forma, en otras realizaciones se tiene en cuenta la ramificación de un polímero PEG cuando se optimizan estos parámetros. En este caso, cuanto mayor es el peso molecular (o de más ramificaciones) mayor es la relación polímero:proteína.

Acoplamiento directo de aminoazúcares con pirrolisina y/o PCL incorporadas en proteínas. Polipéptidos y/o péptidos

55 En otro aspecto proporcionado en el presente documento, proteínas, los polipéptidos y/o péptidos que tienen uno o más resto(s) de pirrolisinas y/o PCL incorporados en el anterior se derivatizan con aminoazúcares. La Figura 38 muestra un esquema de reacción generalizado donde D-manosamina se acopla a una proteína (ilustrada como R₁) que tiene PCL incorporada en la anterior. En dicha realización el resto de aminoaldehído de D-manosamina reacciona con el resto de PCL aumentando por tanto la masa de la proteína en 161 Da. Otros aminoazúcares usados en dichas realizaciones incluyen, pero no se limitan a, manosamina, galactosamina, y glucosamina. En la Figura 39 se muestra el espectro de masas de hRBP4 con la PCL incorporada en la posición 122 (hRBP4 Phe122PCL) después de la reacción con manosamina, mientras que la Figura 40 muestra el espectro de masas de la ácido graso sintetasa humana (FAS-TE) con PCL incorporada en la posición 2222 (FAS-TE Leu2222PCL/Leu2223Ile) antes (Figura 40A) y después (Figura 40B) de la reacción con manosamina.

65

Glicosilación de proteínas, polipéptidos y/o péptidos mediante acoplamiento directo con pirrolisina y PCL a dichas proteínas, polipéptidos y/o péptidos

Los métodos y composiciones descritos en el presente documento se usan para obtener proteínas, polipéptidos y péptidos con uno o más restos de pirrolisina y/o PCL que soportan restos de sacáridos. Los restos de sacáridos pueden ser tanto naturales (incluyendo pero sin limitarse a, N-acetilglucosamina y D-manosamina) como no naturales (incluyendo pero sin limitarse a, 3-fluorogalactosa). En determinadas realizaciones, los sacáridos se unen al resto de pirrolisina y/o PCL mediante un enlace no natural que incluye, pero sin limitación, mediante la formación de un resto de tipo quinazolina. En determinadas realizaciones, los sacáridos se unen al resto de pirrolisina y/o PCL mediante un enlace no natural que incluye, pero sin limitación, mediante la formación de un resto de tipo quinazolina que se ha reducido adicionalmente con un agente reductor. En determinadas realizaciones, los sacáridos se unen al resto de pirrolisina y/o PCL mediante un enlace no natural que incluye, pero sin limitación, mediante la formación de un resto de tipo quinazolina que ha reaccionado adicionalmente formando por tanto el sistema de anillo fusionado que se proporciona en el presente documento (véanse las Figuras 22, 23 y 30). En determinadas realizaciones, la adición de un sacárido o sacáridos, incluyendo, pero sin limitación, restos de glicosilo, que se añaden a proteínas, polipéptidos y péptidos que tienen uno o más pirrolisinas o PCL incorporadas en las anteriores, se produce in vivo. En otras realizaciones, la adición de un sacárido o sacáridos, incluyendo, pero sin limitación, restos de glicosilo, se añaden a proteínas, polipéptidos y péptidos que tienen uno o más pirrolisinas o PCL incorporadas en las anteriores, se produce in vitro. En otras realizaciones, una vez unidos al resto de pirrolisina y/o PCL, el sacárido se modifica adicionalmente mediante tratamiento con glicosiltransferasas y/u otras enzimas para generar un oligosacárido que se une a la proteína, polipéptidos o péptidos que tienen uno o más pirrolisinas o PCL incorporadas en las anteriores.

La Figura 41 ilustra una realización donde la modificación simultánea a la traducción o posterior a la traducción comprende la unión de un oligosacárido a una proteína, polipéptido o péptido mediante la formación de un resto de tipo quinazolina. En determinadas realizaciones, los sacáridos se unen al resto de pirrolisina y/o PCL mediante un enlace no natural que incluye, pero sin limitación, mediante la formación de un resto de tipo quinazolina que se ha reducido adicionalmente con un agente reductor. En determinadas realizaciones, los sacáridos se unen al resto de pirrolisina y/o PCL mediante un enlace no natural que incluye, pero sin limitación, mediante la formación de un resto de tipo quinazolina que ha reaccionado adicionalmente formando por tanto el sistema de anillo fusionado que se proporciona en el presente documento (Figuras 22, 23 y 30). En dichas realizaciones, a modo de ejemplo solamente, el oligosacárido comprende el núcleo (GlcNAc-Man)₂-Man-GlcNAc-GlcNAc unido a 2-amino-acetofenona (2-AAP) se conjuga a un resto de pirrolisina y/o PCL incorporado a la proteína, polipéptido o péptido. En otras realizaciones el oligosacárido comprende el núcleo (GlcNAc-Man)₂-Man-GlcNAc-GlcNAc unido a 2-amino-benzaldehído (2-ABA) y el conjugado de proteína se forma mediante reacción del 2-ABA con un resto de pirrolisina y/o PCL incorporado a la proteína, polipéptido o péptido. En otras realizaciones más el oligosacárido comprende el núcleo de (GlcNAc-Man)₂-Man-GlcNAc-GlcNAc unido a 2-amino-benzofenona (2-ABP) y el conjugado de proteína se forma mediante reacción del 2-ABP con un resto de pirrolisina y/o PCL incorporado a la proteína, polipéptido o péptido. Otros oligosacáridos usados en dichas realizaciones poseen restos 2-ABA, 2-AAP o 2-ANBP.

Unión covalente orientada de proteínas en heterodímeros, heterotrímeros, heteromultímeros, homodímeros, homotrímeros, homomultímeros, homodímeros no simétricos, homotrímeros no simétricos y homomultímeros no simétricos.

En otro aspecto que utiliza los métodos proporcionados en el presente documento, la incorporación específica del sitio de una o más pirrolisinas y/o PCL en proteínas, polipéptidos y/o péptidos se usa para formar conjugados de proteína-proteína que incluyen, pero sin limitación, heterodímeros, heterotrímeros, heteromultímeros, homodímeros, homotrímeros, homomultímeros, homodímeros no simétricos, homotrímeros no simétricos y homomultímeros no simétricos. Dicha formación de conjugado proteína-proteína se lleva a cabo utilizando la reticulación específica del sitio entre proteínas que tienen resto(s) de pirrolisina y PCL incorporados en el anterior. La Figura 42 ilustra determinadas realizaciones (heterodímeros, heterotrímeros, homotrímeros) de dicha formación del conjugado proteína-proteína, donde las proteínas que tienen resto(s) de PCL incorporados en el anterior se reticulan. En la Figura 42, el enlace del conjugado de proteína formado mediante dicha reticulación es un resto de tipo quinazolina que une las proteínas juntas, sin embargo, en otras realizaciones el enlace es una forma reducida del resto de tipo quinazolina (Figuras 22, 23 y 30). En otras realizaciones, el enlace es el resto anillo fusionado (Figuras 22, 23 y 30).

Los ejemplos no limitantes de dichos conjugados de proteína-proteína incluyen, pero no se limitan a, conjugados de proteína-proteína con una citoquina, un factor de crecimiento, un receptor del factor de crecimiento, un interferón, una interleuquina, una molécula inflamatoria, un producto de oncogén, una hormona peptídica, una molécula de transducción de la señal, un receptor de la hormona esteroidea, un activador de la transcripción, un supresor de la transcripción, eritropoyetina (EPO), factores de crecimiento de fibroblastos, factor de crecimiento de fibroblastos 21 (FGF21), leptina, insulina, hormona del crecimiento humano, Péptido -78 activador neutrófilo epitelial, GROa/MGSA, GROp, GROy, MIP-1a, MIP-1b, MCP-1, factor de crecimiento de hepatocitos, factor de crecimiento de tipo insulina, factor de inhibición de la leucemia, oncostatina M, PD-ECSF, PDGF, pleiotropina, SCF, ligando c-kit, VEGF, G-CSF, IL-1, IL-2, IL-8, IGF-I, IGF-II, FGF (factor de crecimiento de fibroblastos), PDGF, TNF, TGF- α , TGF- β , EGF (factor de crecimiento epidérmico), KGF (factor de crecimiento de queratinocitos), CD40L/CD40, VLA-4/VCAM-1, ICAM-1/LFA-1, hialurina/CD44, Mos, Ras, Raf, Met; p53, Tat, Fos, Myc, Jun, Myb, Rel, receptor de estrógenos, receptor de

progesterona, receptor de testosterona, receptor de aldosterona, receptor de LDL, y/o corticosterona. En otro conjunto de realizaciones, La proteína es homóloga a una terapéutica u otra proteína tal como: una alfa-1 antitripsina, una angiostatina, un factor antihemolítico, un anticuerpo, una apolipoproteína, una apoproteína, un factor natriurético atrial, un polipéptido natriurético atrial, un péptido atrial, una quimioquina C-X-C, T39765, NAP-2, ENA-78, una Gro-a, una Gro-b, una Gro-c, una IP-10, una GCP-2, una NAP-4, un SDF-1, una PF4, una MIG, una Calcitonina, un ligando c-kit, una citoquina, una quimioquina CC, una proteína-1 quimioattractora de monocitos, una proteína-2 quimioattractora de monocitos, una proteína-3 quimioattractora de monocitos, una proteína 1 alfa inflamatoria de monocitos, una proteína 1 beta inflamatoria de monocitos, RANTES, 1309, R83915, R91733, HCC1, T58847, D31065, T64262, un CD40, un ligando de CD40, ligando C-kit, un colágeno, un factor estimulador de colonias (CSF), un factor 5a del complemento, un inhibidor del complemento, un receptor 1 del complemento, una citoquina, un Péptido -78 activador neutrófilo epitelial, una GROa/MGSA, una GROp, una GROy, una MIP-1a, una MIP-16, una MCP-1, un factor de crecimiento epidérmico (EGF), un péptido activador neutrófilo epitelial, una toxina exfoliante, un Factor IX, un Factor VII, un Factor VIII, un Factor X, un factor de crecimiento de fibroblastos (FGF), una fibrinógeno, una fibronectina, una G-CSF, una GM-CSF, una glucocerebrosidasa, una gonadotropina, un factor de crecimiento, un receptor del factor de crecimiento, una proteína Hedgehog, una hemoglobina, una proteína de crecimiento de hepatocitos (HGF), una hirudina, una albúmina sérica humana, un ICAM-1, un receptor de ICAM-1, un LFA-1, un receptor de LFA-1, una insulina, un factor de crecimiento de tipo insulina (IGF), un IGF-I, un IGF-II, un interferón, un IFN-a., un IFN-p, un IFN-y, una interleuquina, una IL-1, una IL-2, una IL-3, una IL-4, una IL-5, una IL-6, una IL-7, una IL-8, una IL-9, una IL-10, una IL-11, una IL-12, un factor de crecimiento de queratinocitos (KGF), una lactoferrina, un factor de inhibición de la leucemia, una luciferasa, una neurturina, un factor inhibidor de neutrófilos (FIN), una oncoestatina M, una proteína osteogénica, un producto de oncogén, una hormona paratiroidea, un PD-ECSF, un PDGF, una hormona peptídica, una hormona del crecimiento humano, una pleiotropina, una proteína A, una proteína G, unas exotoxinas A pirógenas, B, o C, una relaxina, una renina, un SCF, un receptor I del complemento soluble, un I-CAM 1 soluble, unos receptores de interleuquina solubles, un receptor TNF soluble, una somatomedina, una somatostatina, una somatotropina, una estreptoquinasa, unos superantígenos, unas enterotoxinas estafilocócicas, un SEA, un SEB, un SEC 1, un SEC2, un SEC3, un SED, un SEE, un receptor de la hormona esteroidea, una superóxido dismutasa, una toxina de síndrome de choque tóxico, una timosina alfa 1, un activador del plasminógeno tisular, un factor de crecimiento tumoral (TGF), un TGF-a, un TGF-p, un factor de necrosis tumoral, un factor alfa de necrosis tumoral, un factor beta de necrosis tumoral, un receptor del factor de necrosis tumoral (TNFR), una proteína VLA-4, una proteína VCAM-1, un factor A de crecimiento endotelial vascular (VEGF), una uroquinasa, un Mos, un Ras, un Raf, un Met; un p53, un Tat, un Fos, un Myc, un Jun, un Myb, un Rel, un receptor de estrógenos, un receptor de progesterona, un receptor de testosterona, un receptor de aldosterona, un receptor de LDL, y/o corticosterona.

Los reticuladores usados para formar conjugados de proteína-proteína que incluyen, pero sin limitación, heterodímeros, heterotrímeros, heteromultímeros, homodímeros, homotrímeros, homomultímeros, homodímeros no simétricos, homotrímeros no simétricos y homomultímeros no simétricos, poseen restos de benzaldehído, acetofenona y/o benzofenona en sus respectivos extremos terminales. Dichos restos reaccionan con resto(s) de pirrolisina y/o PCL incorporados en proteínas que utilizan los métodos proporcionados en el presente documento, y dichas proteínas incluyen aquellas proporcionadas en el presente documento. Un ejemplo no limitante de un reticulador bifuncional utilizado para formar un homodímero se muestra en la Figura 43A. Este enlazador bifuncional se utilizó para reticular el factor 21 de crecimiento de fibroblastos (FGF-21) con un resto de PCL incorporado en la posición 84 (FGF21 Lys84PCL). Figura 43B es el espectro de masas de la mezcla de reacción, donde el pico del dímero FGF21 Lys48PCL covalente en la masa esperada de 43037,2 Da es la especie dominante. FGF21 Lys84PCL sin reaccionar se detecta a 21235,2 Da, aunque algunos FGF21 monoméricos modificados con un extremo del reticulador unido se detectan a 21820,0 Da. La modificación de las condiciones de reacción permitió la reacción completa de FGF21 Lys84PCL. En la Figura 43, el enlace del conjugado de proteína formado mediante dicha reticulación es un resto de tipo quinazolina, sin embargo, en otras realizaciones el enlace es una forma reducida del resto de tipo quinazolina (Figuras 22, 23 y 30). En otras realizaciones, el enlace es el resto anillo fusionado (Figuras 22, 23 y 30).

Figura 44A es el espectro de masas de la mezcla de reacción, donde el pico del dímero FGF21 covalente en la masa esperada de 43034 Da es la especie dominante. FGF21 Lys84PCL sin reaccionar no se detectó a 21234 Da, aunque algunos FGF21 monoméricos modificados con un extremo del reticulador unido se detectan a 21818 Da. La Figure 43B es el SDS-PAGE de las reacciones a pH 5.0 y pH 7.5, indicando una mejor conversión a pH 5.0. Se ha utilizado la reticulación específica de PCL para preparar los dímeros de FGF21. Sin embargo la metodología es aplicable a cualquiera de las proteínas relacionadas anteriormente y se puede usar para formar dímeros, trímeros, y multímeros. Dicha reticulación se usa para potenciar la actividad de las proteínas biológicamente activas. Además, esta solución se usa para estabilizar complejos para estudios estructurales y/o estabilizar señuelos de receptores. La Figura 45 muestra una realización para un reticulador utilizado para formar trímeros.

60 Marcado específico de sitios

En otro aspecto que utiliza los métodos proporcionados en el presente documento, la incorporación específica del sitio de una o más pirrolisinas y/o PCL en proteínas, polipéptidos y/o péptidos se usa para permitir el marcado específico de sitio de dichas proteínas, polipéptidos y/o péptidos. Dicho marcado es el resultado de la reacción de resto(s) de pirrolisina y/o PCL con una marca derivatizada con un grupo funcional que reacciona selectivamente con los restos(s) de pirrolisina y/o PCL. La marca usada incluye, pero sin limitación, un colorante, un anticuerpo o fragmento de

anticuerpo, un quelante metálico, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un resto excitable mediante radiación actínica, un ligando biotina, un análogo de biotina, un agente redox activo, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo quimioluminiscente, un grupo denso en electrones, un grupo magnético, un grupo intercalante, un cromóforo; un agente de transferencia de energía, punto(s) cuántico(s), colorantes fluorescentes, reactivos FRET, y cualquier combinación de los mismos. La marca es también cualquier X^1 que se define en el presente documento.

En realizaciones de dicho marcado específico de sitios de dichas proteínas, polipéptidos y/o péptidos que contienen uno o más restos de pirrolisina y/o PCL es la reacción de los restos de pirrolisina y/o PCL con una marca derivatizada con restos de benzaldehído, acetofenona o benzofenona. La marca usada incluye, pero sin limitación, un colorante, un anticuerpo o fragmento de anticuerpo, un quelante metálico, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un resto excitable mediante radiación actínica, un ligando biotina, un análogo de biotina, un agente redox activo, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo quimioluminiscente, un grupo denso en electrones, un grupo magnético, un grupo intercalante, un cromóforo; un agente de transferencia de energía, punto(s) cuántico(s), colorantes fluorescentes, reactivos FRET, y cualquier combinación de los mismos. La marca es también cualquier X^1 que se define en el presente documento.

En un aspecto adicional de dicho marcado específico de sitio proporcionado en el presente documento, la reacción del(de los) resto(s) de pirrolisina y/o PCL con una marca derivatizada con un resto que reacciona selectivamente con el(los) restos) de pirrolisina y/o PCL permite un aumento de la sensibilidad de la detección. Específicamente, dichas reacciones forma un resto detectable que tiene propiedades espectrales diferentes de las de los reactivos presentes antes de la reacción, potenciando por tanto la relación señal a ruido minimizando la señal de fondo en relación con la señal del resto detectable. En determinadas realizaciones de dicho marcado específico de sitio proporcionado en el presente documento, la reacción del(de los) resto(s) de pirrolisina o PCL con una marca derivatizada con restos de benzaldehído, acetofenona o benzofenona forma un resto de tipo quinazolina (Figuras 22, 23 y 30) que tiene propiedades espectrales diferentes de las del benzaldehído, acetofenona o benzofenona. En determinadas realizaciones de dicho marcado específico de sitio proporcionado en el presente documento, la reacción del(de los) resto(s) de pirrolisina o PCL con una marca derivatizada con restos de benzaldehído, acetofenona o benzofenona forma un resto de tipo quinazolina reducido (Figuras 22, 23 y 30) que tiene propiedades espectrales diferentes de las del benzaldehído, acetofenona o benzofenona. En determinadas realizaciones de dicho marcado específico de sitio proporcionado en el presente documento, la reacción del(de los) resto(s) de pirrolisina o PCL con una marca derivatizada con restos de benzaldehído, acetofenona o benzofenona forma un resto de anillo fusionado (Figuras 22, 23 y 30) que tiene propiedades espectrales diferentes de las del benzaldehído, acetofenona o benzofenona.

En determinadas realizaciones, dichas proteínas marcadas, polipéptidos y/o péptidos se usan para diagnóstico y diagnóstico por imágenes, mientras que en otras realizaciones dichas proteínas, polipéptidos y/o péptidos marcados se usan en selección de alto rendimiento. En otras realizaciones, dichas proteínas marcadas, polipéptidos y/o péptidos marcados se usan para evaluar propiedades de transporte, mientras que en otras realizaciones dichas proteínas, polipéptidos y/o péptidos marcados se usan en estudios de localización.

La Figura 46 ilustra determinadas realizaciones de dicho marcado específico de sitio. La Figura 46A ilustra el marcado con un resto fluorescente. La Figura 46B ilustra la reacción de un resto de PCL con un resto no fluorescente que tras la reacción se vuelve fluorescente. La Figura 46C ilustra la biotilación de una proteína. La Figura 46D ilustra el marcado con un resto radioactivo. La Figura 46E ilustra el marcado con 1-(2-amino-5-yodofenil)etanona y la Figura 46F ilustra el marcado con 1-(2-amino-5-bromofenil)etanona; se pueden usar ambas para obtener información por etapas del procedimiento de determinación de estructuras cristalinas de proteínas mediante rayos X. 1-(2-amino-5-yodofenil)etanona se puede usar también para marcar proteínas con un resto radioactivo. En la Figura 46, el enlace del conjugado de proteína formado mediante dicho marcado es un resto de tipo quinazolina, sin embargo, en otras realizaciones el enlace es una forma reducida del resto de tipo quinazolina (Figuras 22, 23 y 30). En otras realizaciones, el enlace es el resto anillo fusionado (Figuras 22, 23 y 30).

En realizaciones de dicho marcado específico de sitios de dichas proteínas, polipéptidos y péptidos que contienen uno o más resto(s) de pirrolisina y/o PCL, los propios resto(s) de pirrolisina o PCL pueden marcarse utilizando un precursor marcado. La Figura 46G y la Figura 46H ilustran, a modo de ejemplo solamente, el marcado con isótopos radioactivos o estables procedentes de precursores marcados de forma diferente.

Otro aspecto de la presente invención proporciona la producción de proteínas que son homólogas para cualquier proteína disponible, pero que comprenden uno o más resto(s) de pirrolisina o PCL. Por ejemplo, en determinadas realizaciones, las proteínas terapéuticas se preparan de tal manera que comprenden una o más pirrolisinas o PCL y son homólogas a una o más proteínas terapéuticas. Por ejemplo, en un aspecto, La proteína es homóloga a una terapéutica u otra proteína tal como: una citoquina, un factor de crecimiento, un receptor del factor de crecimiento, un interferón, una interleuquina, una molécula inflamatoria, un producto de oncogén, una hormona peptídica, una molécula de transducción de la señal, un receptor de la hormona esteroidea, un activador de la transcripción, un supresor de la transcripción, eritropoyetina (EPO), factores de crecimiento de fibroblastos, factor de crecimiento de fibroblastos 21 (FGF21), leptina, insulina, hormona del crecimiento humano, Péptido -78 activador neutrófilo epitelial, GROa/MGSA, GROp, GROy, MIP-1a, MIP- 16, MCP-1, factor de crecimiento de hepatocitos, factor de crecimiento de

tipo insulina, factor de inhibición de la leucemia, oncostatina M, PD-ECSF, PDGF, pleiotropina, SCF, ligando c-kit, VEGF, G-CSF, IL-1, IL-2, IL-8, IGF-I, IGF-II, FGF (factor de crecimiento de fibroblastos), PDGF, TNF, TGF-a, TGF-p, EGF (factor de crecimiento epidérmico), KGF (factor de crecimiento de queratinocitos), CD40L/CD40, VLA-4/VCAM-1, ICAM-1/LFA-1, hialurina/CD44, Mos, Ras, Raf, Met; p53, Tat, Fos, Myc, Jun, Myb, Rel, receptor de estrógenos, receptor de progesterona, receptor de testosterona, receptor de aldosterona, receptor de LDL, y/o corticosterona. En otro conjunto de realizaciones, La proteína es homóloga a una terapéutica u otra proteína tal como: una alfa-1 antitripsina, una angiostatina, un factor antihemolítico, un anticuerpo, una apolipoproteína, una apoproteína, un factor natriurético atrial, un polipéptido natriurético atrial, un péptido atrial, una quimioquina C-X-C, T39765, NAP-2, ENA-78, una Gro-a, una Gro-p, una Gro-y, una IP-10, una GCP-2, una NAP-4, un SDF-1, una PF4, una MIG, una Calcitonina, un ligando c-kit, una citoquina, una quimioquina CC, una proteína-1 quimioattractora de monocitos, una proteína-2 quimioattractora de monocitos, una proteína-3 quimioattractora de monocitos, una proteína 1 alfa inflamatoria de monocitos, una proteína 1 beta inflamatoria de monocitos, RANTES, 1309, R83915, R91733, HCC1, T58847, D31065, T64262, un CD40, un ligando de CD40, ligando C-kit, un colágeno, un factor estimulador de colonias (CSF), un factor 5a del complemento, un inhibidor del complemento, un receptor 1 del complemento, una citoquina, un Péptido -78 activador neutrófilo epitelial, una GROa/MGSA, una GROp, una GROy, una MIP-1a, una MIP-16, una MCP-1, un factor de crecimiento epidérmico (EGF), un péptido activador neutrófilo epitelial, interferón alfa (INF-a), o interferón beta (INF-p), una toxina exfoliante, un Factor IX, un Factor VII, un Factor VIII, un Factor X, un factor de crecimiento de fibroblastos (FGF), una fibrinógeno, una fibronectina, una G-CSF, una GM-CSF, una glucocerebrosidasa, una gonadotropina, un factor de crecimiento, un receptor del factor de crecimiento, una proteína Hedgehog, una hemoglobina, un factor de crecimiento de hepatocitos (HGF), una hirudina, una albúmina sérica humana, un ICAM-1, un receptor de ICAM-1, un LFA-1, un receptor de LFA-1, una insulina, un factor de crecimiento de tipo insulina (IGF), un IGF-I, un IGF-II, un interferón, un IFN-a, un IFN-p, un IFN-y, una interleuquina, una IL-1, una IL-2, una IL-3, una IL-4, una IL-5, una IL-6, una IL-7, una IL-8, una IL-9, una IL-10, una IL-11, una IL-12, un factor de crecimiento de queratinocitos (KGF), una lactoferrina, un factor de inhibición de la leucemia, una luciferasa, una neurturina, un factor inhibidor de neutrófilos (FIN), una oncoestatina M, una proteína osteogénica, un producto de oncogén, una hormona paratiroidea, un PD-ECSF, un PDGF, una hormona peptídica, una hormona del crecimiento humano, una pleiotropina, una proteína A, una proteína G, unas exotoxinas A pirógenas, B, o C, una relaxina, una renina, un SCF, un receptor I del complemento soluble, un I-CAM 1 soluble, unos receptores de interleuquina solubles, un receptor TNF soluble, una somatomedina, una somatostatina, una somatotropina, una estreptoquinasa, unos superantígenos, unas enterotoxinas estafilocócicas, un SEA, un SEB, un SEC 1, un SEC2, un SEC3, un SED, un SEE, un receptor de la hormona esteroidea, una superóxido dismutasa, una toxina de síndrome de choque tóxico, una timosina alfa 1, un activador del plasminógeno tisular, un factor de crecimiento tumoral (TGF), un TGF-a, un TGF-p, un factor de necrosis tumoral, un factor alfa de necrosis tumoral, un factor beta de necrosis tumoral, un receptor del factor de necrosis tumoral (TNFR), una proteína VLA-4, una proteína VCAM-1, un factor A de crecimiento endotelial vascular (VEGF), una uroquinasa, un Mos, un Ras, un Raf, un Met; un p53, un Tat, un Fos, un Myc, un Jun, un Myb, un Rel, un receptor de estrógenos, un receptor de progesterona, un receptor de testosterona, un receptor de aldosterona, un receptor de LDL, y/o corticosterona.

En un aspecto, las composiciones en el presente documento comprenden una proteína, incluyendo, cualquiera de las proteínas señaladas anteriormente, que comprenden uno o más resto(s) de pirrolisina o PCL y un excipiente farmacéuticamente aceptable.

En determinadas realizaciones, la proteína es una proteína terapéutica, mientras que en otras realizaciones, la proteína terapéutica es eritropoyetina (EPO), factor de crecimiento de fibroblastos 21 (FGF21), interferón alfa (INF-a), o interferón beta (INF-p). En determinadas realizaciones, la proteína terapéutica es un anticuerpo o fragmento de anticuerpo, incluyendo, pero sin limitarse a los Fab. En determinadas realizaciones, la proteína terapéutica se produce como una proteína de fusión donde la molécula de fusión asociada está modificada. En determinadas realizaciones, la proteína terapéutica es una fusión con un dominio Fc.

se puede inferir la homología con la proteína o el polipéptido llevando a cabo una alineación de secuencias, por ejemplo, utilizando BLASTN o BLASTP, por ejemplo, configurando los parámetros por defecto. Por ejemplo, en una realización, la proteína tiene al menos aproximadamente 50 %, al menos aproximadamente 75 %, al menos aproximadamente 80 %, al menos aproximadamente 90 % o al menos aproximadamente 95 % idéntica a una proteína terapéutica conocida (por ejemplo, una proteína presente en el Genbank u otras bases de datos disponibles).

Para demostrar el marcado específico del sitio, PCL se incorporó en mEGF utilizando los métodos proporcionados en el presente documento (véase el Ejemplo 12), y el resto de PCL se acopló a una biotina mediante un poliéter funcionalizado con un resto ABA (véase X3626-140, Ejemplo 40). La Figura 47A muestra el análisis por espectrometría de masas ESI de mEGF-Tyr10PCL conjugado con biotina (véase el Ejemplo 24)). La Figura 47B muestra la transferencia Western del conjugado de mEGF-Tyr10PCL-ABA-biotina utilizando una peroxidasa de rábano picante (HRP) conjugada con un anticuerpo de cabra dirigido contra biotina. mEGF-Tyr10PCL desacoplada y fluoresceína-conjugada con mEGF-Tyr10PCL sirvieron como controles negativos.

Para demostrar adicionalmente el marcado específico del sitio, PCL se incorporó en mEGF utilizando los métodos proporcionados en el presente documento (véase el Ejemplo 12), y el resto de PCL se acopló a una fluoresceína mediante un poliéter funcionalizado con un resto ABA (véase X3757-48, Ejemplo 41). La Figura 47C muestra el

análisis por espectrometría de masas ESI de mEGF-Tyr10PCL conjugado con biotina (véase el Ejemplo 25)). Además, PCL se incorporó en mEGF utilizando los métodos proporcionados en el presente documento (véase el Ejemplo 12), y el resto de PCL se acopló a un disacárido funcionalizado con un resto ABA (véase 3793-050, Ejemplo 42). La Figura 47D muestra el análisis por espectrometría de masas ESI de mEGF-Tyr10PCL conjugado el disacárido (véase el Ejemplo 26). Se entiende que la química del acoplamiento usada para la unión del disacárido puede aplicarse fácilmente al acoplamiento de otros azúcares y oligo y polisacáridos complejos.

Acoplamiento de moduladores inmunes a pirrolisina y/o análogos de pirrolisina incorporados en proteínas, polipéptidos y/o péptidos e inmunogenicidad potenciada mediante dicho acoplamiento

En otro aspecto proporcionado en el presente documento, las proteínas, polipéptidos y/o péptidos que tienen uno o más pirrolisinas y/o PCL incorporados en el anterior se derivatizan con uno o más inmunomoduladores. Utilizando los métodos proporcionados en el presente documento los restos inmunoestimuladores pueden acoplarse fácilmente a las proteínas, polipéptidos y/o péptidos mediante resto(s) de pirrolisina y/o PCL. Dichos restos inmunoestimuladores incluyen, pero no se limitan a, uno o más grupos nitro, lípidos, fosfolípidos, moléculas de tipo LPS, adyuvantes, moléculas de tipo adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, epítomos de linfocito T, hemocianina de lapa californiana (KLH), haptenos inmunógenos, halógenos, grupos arilo, grupos heteroarilo, grupos cicloalquilo o grupos heterocicloalquilo. En determinadas realizaciones, los restos inmunomoduladores unidos a pirrolisina y/o PCL se usan para estimular una respuesta inmune contra el autoantígeno de una manera similar a la p-nitrofenilalanina incorporada en proteínas (Grunewals J, Tsao ML, Perera R, Dong L, Niessen F, Wen BG, Kubitz DM, Smider VV, Ruf W, Nasoff M, Lemer RA, Schultz PG, Terminación inmunológica de la autotolerancia, Proc Natl Acad Sci U S A. 12 Ago 2008;105(32):11276-80). En determinadas realizaciones, un autoantígeno con un inmunomodulador unido a pirrolisina y/o PCL se usa como una vacuna contra el cáncer. En otras realizaciones, el antígeno con un inmunomodulador unido a pirrolisina y/o PCL se usa como una vacuna para enfermedades infecciosas. En otras realizaciones, el antígeno con un inmunomodulador unido a pirrolisina y/o PCL se usa como una vacuna para enfermedades que implican la formación de amiloides, incluyendo pero sin limitarse a, enfermedad de Alzheimer.

En una realización se genera una biblioteca de restos inmunomoduladores creando una biblioteca de compuestos de 2-aminobenzaldehído, el acoplamiento de esta biblioteca con pirrolisina o PCL incorporadas en un antígeno de interés y seleccionando a continuación el producto acoplado para la generación de elevados títulos de anticuerpos contra el antígeno modificado y no modificado. En dicha realización, el resto de 2-aminobenzaldehído está sustituido con diversos sustituyentes que incluyen, pero sin limitación, uno o más grupos nitro, lípidos, fosfolípidos, moléculas de tipo LPS, adyuvantes, moléculas de tipo adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, epítomos de linfocito T, hemocianina de lapa californiana (KLH), haptenos inmunógenos, halógenos, grupos arilo, grupos heteroarilo, grupos cicloalquilo o grupos heterocicloalquilo.

En una realización se genera una biblioteca de restos inmunoestimuladores creando una biblioteca de compuestos de 2-aminoacetofenona, el acoplamiento de esta biblioteca con pirrolisina o PCL incorporadas en un antígeno de interés y seleccionando a continuación el producto acoplado para la generación de elevados títulos de anticuerpos contra el antígeno modificado y no modificado. En dicha realización, el resto de 2-aminoacetofenona está sustituido con diversos sustituyentes que incluyen, pero sin limitación, uno o más grupos nitro, lípidos, fosfolípidos, moléculas de tipo LPS, adyuvantes, moléculas de tipo adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, epítomos de linfocito T, hemocianina de lapa californiana (KLH), haptenos inmunógenos, halógenos, grupos arilo, grupos heteroarilo, grupos cicloalquilo o grupos heterocicloalquilo.

En otra realización se genera una biblioteca de restos inmunoestimuladores creando una biblioteca de compuestos de 2-amino-5-nitro-benzofenona, el acoplamiento de esta biblioteca con pirrolisina o PCL incorporadas en un antígeno de interés y seleccionando a continuación el producto acoplado para la generación de elevados títulos de anticuerpos contra el antígeno modificado y no modificado. En dicha realización, el resto de 2-amino-5-nitro-benzofenona está sustituido con diversos sustituyentes que incluyen, pero sin limitación, uno o más grupos nitro, lípidos, fosfolípidos, moléculas de tipo LPS, adyuvantes, moléculas de tipo adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, epítomos de linfocito T, hemocianina de lapa californiana (KLH), haptenos inmunógenos, halógenos, grupos arilo, grupos heteroarilo, grupos cicloalquilo o grupos heterocicloalquilo.

Para demostrar este aspecto se incorporó PCL en mTNF-a y mEGF utilizando los métodos proporcionados en el presente documento (véase los Ejemplos 11 y 12), y el resto de PCL se acopló a un hapteno que contenía uno o más grupos nitrofenilo (véanse los Ejemplos 27 y 28). La Figura 48A y la Figura 48B demuestran la unión de un mononitrofenil hapteno (véase 3793-001, Ejemplo 38-8) en el sitio de incorporación de la PCL en mTNF-Gln21PCL (Figura 48A) y mEGF-Tyr10PCL (Figura 48B). La Figura 48C y la Figura 48D muestra la unión de un dinitrofenilhapteno (TU3627-088, Ejemplo 38-7) en el sitio de incorporación de la PCL en mTNF-a (Figura 48C) y mEGF (Figura 48D).

Las Figuras 49A demuestran la unión de un agonista de TLR7 (véase X3678-114; Ejemplo 38-3) en el sitio de incorporación de la PCL en mEGF-Tyr10PCL (véase el Ejemplo 29). Las Figuras 49B demuestran la unión de un fosfolípido (véase TU3627-092; Ejemplo 43-1) en el sitio de incorporación de la PCL en mEGF-Tyr10PCL (véase el

Ejemplo 31).

Las Figuras 50-51 demuestran la conjugación de los péptidos PADRE: PX2-PADRE (PM = 1585) (véase 3465-143; Ejemplo 38-11), y BHA-exPADRE (PM = 2060) (véase 3647-104; Ejemplo 38-10) para cualquier mTNF-Gln21PCL o mEGF-Tyr10PCL. La Figura 50A y la Figura 50B muestran el análisis por espectrometría de masas MALDI-TOF de la reacción de conjugación de mTNF-Gln21PCL con PX2-PADRE a dos diferentes valores de pH (véase el Ejemplo 30), mientras que la Figura 50C muestra el análisis por espectrometría de masas ESI de la reacción de conjugación de mTNF-Gln21PCL con BHA-exPADRE (véase el Ejemplo 30). La Figura 51 muestra el acoplamiento de BHA-exPADRE con mEGF-Tyr10PCL (véase el Ejemplo 30). La secuencia del péptido PADRE (SEC ID N°:28) es Gly(DAla)LysXValAlaAlaTrpThrLeuLysAla(D-Ala)Gly-OH, donde X es ciclohexil-alanina. La secuencia del péptido PADRE (SEC ID N°:29) es Gly(DAla)LysXValAlaAlaTrpThrLeuLysAla(D-Ala)Gly-OH, donde X es ciclohexil-alanina. En determinadas realizaciones, los epitopos de péptidos inmunogénicos conocidos se acoplarán a antígenos a fin de potenciar la inmunogenicidad de dicho antígeno. En determinadas realizaciones, los péptidos derivados de un antígeno se acoplarán a una proteína portadora inmunógena. En determinadas realizaciones, la proteína portadora inmunógena es KLH. Se entiende que la química del acoplamiento usada para la unión de los péptidos PADRE puede aplicarse fácilmente al acoplamiento de otros péptidos a PCL y pirrolisina.

Immuno-PCR: Acoplamiento de ADN y otros oligonucleótidos a pirrolisina y/o análogos de pirrolisina incorporados en proteínas, Polipéptidos y/o péptidos

En otro aspecto proporcionado en el presente documento, las proteínas, polipéptidos y/o péptidos que tienen una o más pirrolisinas y/o PCL incorporadas en el anterior se derivatizan con ADN usando los métodos proporcionados en el presente documento. El ADN está modificado para tener el 2-amino-benzaldehído terminal, restos de 2-amino-acetofenona o 2-amino-5-nitro-benzofenona que reaccionan con pirrolisina y/o PCL incorporadas en la proteína, polipéptido y/o péptido, existe el ADN acoplado a la proteína, polipéptido y/o péptido mediante un enlace de tipo quinazolina, redujo el enlace de quinazolina o el enlace del anillo fusionado (véanse las Figuras 22, 23 y 30).

para demostrar la unión del ADN a PCL, se incorporó la PCL en mTNF-a y mEGF utilizando los métodos proporcionados en el presente documento (véase los Ejemplos 11 y 12), y el resto de PCL se acopló a un oligonucleótido CpG que es también un resto inmunoestimulador (véanse los Ejemplos 32). BHA-BG1 (véase 3647-057; Ejemplo 38-12) y BHA-BG2 (véase 3597-167; Ejemplo 38-14) donde el reactivo CpG usado para unir CpG a mTNF-Gln21PCL y mEGF-Tyr10PCL en el sitio de la incorporación de PCL. El acoplamiento de BHA-BG1 (7,4 kDa) y BHA-BG2 (7,4 kDa) a mTNF-Gln21PCL (19,3 kDa) se confirmó mediante el ensayo del cambio de gel (Figura 52A) que fue el acoplamiento de BHA-BG2 (7,4 kDa) a mEGF-Tyr10PCL (7,2 kDa) (Figura 52B). La secuencia de BG 1 (SEC ID N°:30) es 5'*T*C*A*T*G*A*C*G*T*T*C*T*G*A*C*G*T*T-3', donde * denota un enlace fosfotioato. La secuencia de BG2 (SEC ID N°:31) es 5'*T*C*A*T*G*A*C*G*T*T*C*T*G*A*C*G*T*T-3', donde * denota un enlace fosfotioato. En determinadas realizaciones, el oligonucleótido que se va a acoplar será un ácido desoxirribonucleico, un ácido ribonucleico, un ácido nucleico de péptido u otro oligonucleótido modificado. Inicialmente acoplado a pirrolisina o PCL como un oligonucleótido monocatenario, ADN bicatenario, ARN, se pueden preparar fácilmente polímeros de PNA o híbridos mediante hibridación con una hebra complementaria. Se entiende que la química del acoplamiento usada para la unión de oligonucleótidos CpG puede aplicarse fácilmente al acoplamiento de otros oligonucleótidos.

En determinadas realizaciones, las proteínas, polipéptidos y/o péptidos son anticuerpos y dichos anticuerpos que tienen ADN unido se usan para llevar a cabo el análisis de inmuno-PCR. Se llevó a cabo la inmuno-PCR uniendo el ADN a un anticuerpo utilizando los métodos descritos en el presente documento, y a continuación uniendo el ADN unido al anticuerpo con un antígeno diana. Después de la unión, se amplificó el ADN utilizando las técnicas de amplificación del ADN y el producto de amplificación resultante se analizó utilizando técnicas electroforéticas, métodos de microplacas o PCR en tiempo real.

Unión específica a sitio y orientada

En otro aspecto proporcionado en el presente documento, la incorporación específica del sitio de una o más pirrolisinas y/o PCL en proteínas, polipéptidos y/o péptidos se usa para controlar la orientación de dichas proteínas, polipéptidos y/o péptidos tras la unión a la superficie de un soporte. Dicha unión es el resultado de la reacción de los restos de pirrolisina o PCL con una superficie derivatizada con restos de benzaldehído, restos de acetofenona, restos de benzofenona o sus combinaciones, formando por tanto un resto de quinazolina que se une y que orienta la proteína sobre la superficie. Incorporando la pirrolisina y/o la PCL en posiciones específicas en la proteína, se controla la orientación de la proteína sobre la superficie. Dicho control de la orientación de la unión de la proteína se usa como componente de un kit de herramienta de diseño mediante ingeniería genética de una proteína para la evaluación de las propiedades de la proteína. La Figura 53 ilustra una realización de dicha unión orientada específicamente al sitio, donde una superficie se derivatiza con 20 restos de amino-acetofenonas que reaccionan con restos de PCL incorporados en una proteína uniendo por tanto la proteína a la superficie mediante un enlace de tipo quinazolina. En la Figura 53, el enlace del conjugado de proteína formado es un resto de tipo quinazolina, sin embargo, en otras realizaciones, es una forma reducida del resto de tipo quinazolina (véanse las Figuras 22, 23 y 30). En otras realizaciones, el enlace es el resto del anillo fusionado (véanse las Figuras 22, 23 y 30).

Los ejemplos no limitantes de soportes incluyen, pero no se limitan a, matrices sólidas y semisólidas, tales como aerogeles e hidrogeles, resinas, perlas, biochips (que incluyen biochips revestidos con películas finas), chip microfluídico, un chip de silicio, placas multipocillos (denominadas también placas de microvaloración o microplacas), membranas, células, metales conductores y no conductores, soportes de vidrio (incluyendo portas de microscopio) y magnéticos. Otros ejemplos no limitantes de soportes sólidos usados en los métodos y composiciones descritos en el presente documento incluyen geles de sílice, membranas poliméricas, partículas, películas de plástico derivatizadas, vidrio derivatizado, sílice derivatizada, perlas de vidrio, algodón, perlas de plástico, geles de alúmina, polisacáridos tales como Sefarosa, poli(acrilato), poliestireno, poli(acrilamida), poliol, agarosa, agar, celulosa, dextrano, almidón, FICOLL, heparina, glucógeno, amilopectina, manano, inulina, nitrocelulosa, diazocelulosa, cloruro de polivinilo, polipropileno, polietileno (incluyendo poli(etilenglicol)), nailon, perla de látex, perla magnética, perla paramagnética, perla superparamagnética, almidón y similares. En determinadas realizaciones, los soportes usados en los métodos y composiciones descritos en el presente documento son soportes usados para el análisis superficial tales como dispositivos de ondas acústicas superficiales o dispositivos que utilizan análisis de ondas evanescente, tales como análisis de resonancia de plasmón superficial.

Las superficies de los soportes sólidos usados para la unión de proteínas, polipéptidos y/o péptidos que tienen pirrolisina y/o análogos de pirrolisina incorporados en el anterior tienen grupos funcionales reactivos, incluyendo, pero sin limitación, hidroxilo, carboxilo, amino, tiol, aldehído, halógeno, nitro, ciano, amido, urea, carbonato, carbamato, isocianato, sulfona, sulfonato, sulfonamida y sulfóxido. Dichos grupos funcionales se usan para unir covalentemente los restos de 2-amino-benzaldehídos, restos de 2-amino-acetofenona y/o restos de 2-amino-5-nitro-benzofenona que reaccionan con la pirrolisina o análogos de pirrolisina para formar el resto de quinazolina. En determinadas realizaciones, dichos restos de 2-amino-benzaldehído, restos de 2-amino-acetofenona y restos de 2-amino-5-nitro-benzofenona son parte de un enlazador polimérico acoplado al soporte sólido mediante reacción con un grupo funcional reactivo con soporte sólido, tales como, pero sin limitación, hidroxilo, carboxilo, amino, tiol, aldehído, halógeno, nitro, ciano, amido, urea, carbonato, carbamato, isocianato, sulfona, sulfonato, sulfonamida y sulfóxido.

En otras realizaciones, las superficies de los soportes sólidos tienen estreptavidina o avidina unida a los anteriores, y se usan para la unión de proteínas, polipéptidos y/o péptidos que tienen pirrolisina y/o análogos de pirrolisina incorporados al anterior, donde la pirrolisina y/o los análogos de pirrolisina se usan para unir específicamente en el sitio la biotina a las proteínas, polipéptidos y/o péptidos,

Otros soportes usados en los métodos y composiciones descritos en el presente documento incluyen, resinas usadas en la síntesis de péptidos tales como, a modo de ejemplo solamente, poliestireno, resina-PAM, resina POLYHIPE™, resina de poliamida, resina de poliestireno injertada con poli(etilenglicol), resina de polidimetil-acrilamida y perlas PEGA.

En determinadas realizaciones, las proteínas, polipéptidos y/o péptidos que tienen pirrolisina y/o PCL incorporadas en los anteriores se depositan sobre un soporte sólido en un formato de matriz. En determinadas realizaciones, dicha deposición se lleva a cabo mediante contacto superficial directo entre la superficie del soporte y un mecanismo de liberación, tal como una horquilla o un capilar, o mediante tecnologías de chorro de tinta que utilizan propulsión piezoeléctrica y otras formas de propulsión para transferir líquidos a partir de boquillas en miniatura a superficies sólidas. En el caso de la impresión por contacto, los sistemas de control robóticos y cabezales de impresión multiplexados que permiten la fabricación de micromatrices automatizadas. Para la deposición sin contacto mediante tecnologías de propulsión piezoeléctrica, los sistemas robóticos permiten también la fabricación automática de micromatrices usando dispositivos tanto continuos como de goteo controlado.

Ejemplos

Se entiende que las realizaciones descritas en el presente documento son solo a fines ilustrativos.

Ejemplo 1: Incorporación específica del sitio de PCL generado biosintéticamente a proteínas usando células de mamífero.

Este ejemplo proporciona una descripción de las construcciones génicas que, cuando se transfectan a células de mamífero, permiten la incorporación de PCL a sitios codificados por TAG en una proteína diana. Aunque es posible la incorporación simultánea en varios sitios, todos los ejemplos del presente documento ilustran la incorporación de PCL en un solo sitio por molécula de proteína, solamente. Este ejemplo también describe los procedimientos generales para incorporar PCL a proteínas mediante el uso de células de mamífero.

Construcciones:

pyIT de longitud completa y las regiones codificantes de pyIS, pyIT, pyIB, pyIC y pyID se amplificaron a partir de ADN genómico de *Methanosarcina mazei* Gol mediante PCR con cebadores diseñados basados en las secuencias de nucleótidos de pyIS, pyIT, pyIB, pyIC y pyID de *Methanosarcina mazei* Gol (NC_003901, véase más adelante. El codón de inicio de los genes se cambió a ATG. PyIT se clonó en dirección 5' del promotor CMV en el vector pCMV. La

transcripción está bajo el control de un promotor U6 para formar el vector pCMVU6. pylS, pylB, pylC y pylD se clonaron de manera similar en pCMVU6. La transcripción de pylS, pylB, prIC y pylD está bajo el control del promotor CMV (Figura 4A).

- 5 Las regiones de codificación de los genes diana, proteína de unión al retinol humano (hRBP4), eritropoyetina humana y de ratón (EPO), y mIgG1 Fc (mFc) se clonaron en el vector pRS bajo el control de un promotor CMV con una etiqueta His en el extremo C (Figura 4B). Los restos para su mutación a un codón TAG se seleccionaron en función de su exposición a disolvente en modelos estructurales existentes tal como se indica en la Tabla 2. Los codones TAG se introdujeron en los genes diana mediante métodos de PCR.

10

Tabla 2:

Mutantes TAG de sitio único introducidos en proteínas expresadas en células HEK293F.				
n°	EPO humana	EPO de ratón	hRBP4	Fc de IgG1 de ratón
1	R31	R30	Y51	K333
2	A57	A56	F62	K336
3	N63	N60	W83	T394
4	K79	K78	Y116	L426
5	Q92	Q91	W117	
6	R103	Q102	F122	
7	W115	P114	Y140	
8	K143	K142	Y191	
9	A152	T151	Y199	
10	R158	R157		
11	R 193	R192		

La Figura 4B muestra las construcciones de TAG mutante de hRBP4, mEPO, hEPO, y mIgG1. Los sitios de las mutaciones se han relacionado en la Tabla 2 y se indican en la Figura 4B mediante flechas. Para cada construcción, una etiqueta His está unida en el extremo C como herramienta de purificación y detección. Solamente las proteínas de longitud completa pueden tener la marca His. pylT, pylS, pylB, pylC, y pylD se clonaron en el pCMVU6 bajo el control de un CMV (Figura 4A). Se utilizaron en los estudios de transfección simultánea.

15

Cultivo celular y transfección:

20

Las células HEK293F se hicieron crecer en suspensión en el medio de expresión 293 Freestyle a 37°C con CO₂ al 5%. Un día antes de la transfección, las células se dividieron a 0,7x10⁶ células/ml. Se preparó ADN plásmido usando el kit de preparación de plásmidos Qiagen Maxi. Las transfecciones se llevaron a cabo con el método PEI. PEI se mezcló con el ADN plásmido en una relación 2 a 1 en Opti-MEM y el complejo de ADN se añadió a las células HEK293F a 1 ug de ADN plásmido/ml cultivo celular y las células se cultivaron durante cuatro días en presencia de Cyc o D-ornitina 5 mM. Cuando las células se transfectaron simultáneamente con varios plásmidos, la relación entre PEI y la cantidad total de ADN plásmido fue siempre de 2 a 1.

25

Purificación y análisis de proteínas:

30

Cuatro días después de la transfección, los cultivos celulares se centrifugaron a 2000 g durante 20 min y el medio se recogió para la purificación de las proteínas etiquetadas con His. Los medios se introdujeron en columnas Ni-NTA equilibradas previamente con TrisHCl 20 mM (7,5), NaCl 150 mM que contenía imidazol 10 mM. Las columnas se lavaron con el mismo tampón y se eluyeron con tampón de elución (TrisHCl 20 mM 7,5, NaCl 150 mM, imidazol 300 mM). Las proteínas eluidas se ensayaron con el método de Bradford y SDS-PAGE. En algunos casos, los medios y las proteínas purificadas se analizaron mediante transferencia Western con un anticuerpo dirigido contra la etiqueta his o contra proteínas.

35

Las secuencias de los genes pyl codificadas a partir de ADN genómico de Methanosarcina maize se proporcionan a continuación:

40

Secuencia de pyIS: (SEC ID N°:1)

atggataaaaaaccactaaacactctgatatctgcaaccgggctctggatgtccaggaccggaacaattcataaaataaacaccac
 gaagtctctgaagcaaaatctatattgaaatggcatgcccagaccacftgttgaacaactccaggagcagcaggactgcaag
 agcgtcaggcaccacaatacaggaaagacctgcaaacgctgcagggttcggatgaggatcctaataagttcctcacaaggca
 aacgaagaccagacaagcgtaaaagtcaaggctgtttctgccctaccagaacgaaaaaggcaatgccaaaatccgttgcgaga
 gccccgaaacctcttgagaatacagaagcggcacaggctcaaccttctggatctaaatttcacctgcgataccggttccaccaa
 gagtcagtttctgccggcatctgttcaacatcaatatacaagcatttctacaggagcaactgcacccgactggtaaaaggaatac
 gaacccaffacatccatgtctgccctgttcaggcaagtccccgcacttacgaagagccagactgacaggcttgaagtcctgtt
 aaacccaaaagatgagattccctgaattccggcaagcctttcaggagccttgagtcgaattgctctcgcagaaaaaagacct
 gcagcagatctacgcggaagaaagggagaattatctgggggaaactcgcagcgtgaaattaccagggtcttctggacaggggtttct
 ggaaataaaatccccgatctgatccctcttgagtatacgaaggatggcattgataatgataccgaactttcaaacagatctca
 ggttgacaagaacttctgcctgagacccatgcttgcctcaaacctttacaactacctgcgcaagccttgacagggccctgcctgatc
 caataaaaattttgaaataggccatgctacagaaaagagtcggacggcaagaacacctcgaagagttaccatgctgaacttct
 gccagatgggagcgggagtcacacgggaaaatctgaaagcataattacggacttctgaaccacctgggaattgattcaagatc
 gtaggcgattcctgcatggtctatggggatacccttcatgtaatgcacggagacctggaacttccctgcagtagtcggaccatac
 cgcttgaccgggaatgggtattgataaacctggatagggggcagggttcgggctcgaacgccttctaagggttaaacacgactta
 aaaaatcaagagagctgcaagggtccgagtcttactataacgggatttctaccaacctgtaa

5 Secuencia de pyIB: (SEC ID N°:2)

atgatccagaaaaatggcaaccgaagaacttgacagggtcggggagaaaattattgaaagggtttaaattgctgatgatgacctcagg
 gctcttcttcttgaattcgaagaagagctggaaaaacttactatgtagctagaaggtcagaaactattttcggcaacagggtg

 tttctaacgttttatttctcaacttattgtaaaaaccagtgccttttctactataactgtaaaaacgaaattaaccgctaccgct
 gaccggggaagagggttaaaagagatgtgcaaaagccctgaaaagggtcagaggcttcacatgatcgaacctgacaatgggagagggatcc
 ctattactatgatgacctgaccgctcgttgaactgtcagacagtaaaagaaactcgggcttccaataatgatttctccgggag
 ttatggatgacagcaccctctgaaagccagggaagaaggagcaaatcttcttccctttatcaggagacttatgaccgcaacttta
 tggaaaactaagggtaggctcagcttcgaaggaaagggttaaatgcccgagggttgcaaaagaacaggggtactgtatagaagac
 ggcatcttaccggcgtaggaaatgatatcgaatcaactcttatccctgaaggggatgaaagcaacaatcctgatatggtaagg
 gtaatgaclttctgctcagggaaggaaactccgctgaaagggttcagcagatgltcaaaagctttcggagctgaaaatcatagcgattct
 caggctcatgttctgaatgctgataccggcttcttgcacttgaaggcatagacggcatggtgcaccgtttaaagccggagca
 aatattgaacctccatctcccaagattcacgctggaagggggtgccaatfacgaccgcccagatggaagagaggacaggggacg
 ttacaagcgttgcaaaaggctgaagggtatgggaatggaacctgcgccgagcgtgagttgagagagctcgggggtgctaa

Secuencia de py/C: (SEC ID N°:3)

atgagagagtcctgggggtgctagcctgaaaacaatatgccttataggcgggaaagtgcagggttcgaggctgcatacctatctaa
 gaaagccgggaatgaaagtgcctgtaatagacaaaaaccgcaggcgccttataagggaattatgcggatgagttccagtgtttaacat
 aacggaagagccggaaaaactgcgcgatataaaaaatggtgatccatactgccggtaaatgaaaacctggaatgtatagaatt
 tctgaattctataaaagaaaaattctcctgcccgtactttcgattttgaagcttacaggatcagcaggataagagaaatcaaaag
 aatacttcgcatccataggaaacccgacccctcaggacaaaccgtcgggaaccaccttattttgtaaagccctcctgcgaaagcagc
 agtggtggagcgagaataatccatgacaggaaagagcttaaaagccttgagcccgggatgctcatagaagaatacgttgaagg
 gaagtgtctcacttgaggtcatagggatgaaataaattttgctgtgtaaaggaaacctgtacatatcgtgacacctatgactg
 ccatatggtgaccctctccctctagaccctccttcagggaactatcctactccctgcagcaaacctgcccttaaaaggaatlatgg
 acgtggaaagcgttccggcccctgggggtlaaaagtattgagatagatgccgttcccggagccagactccgactgcggtctatt
 attctccgggatcaaccicatagaactcctgtccgggctttaatggaggcatagaagagatcaaaactctccctgaagacaggta
 ctgcattacgaacatctcatgctgcagaaaaatggagtacttaccctgtgggagaaacaggtcctgtccatgggaaatgattacggc
 aattattatgaagaacctggaatagagatttctgtgcaaggagagaacctgtattcaccctggttttctggggcagagacaagg
 aagaagctgaagctagaaaaaacaagggttctgattctaaaaagccgttcggagctgctgcataa

5 Secuencia de py/D: (SEC ID N°:4)

atggcactttaaccccagaagacctggaatattaacaacagcttcaagaagctgattctactgtccgcagagttacagggtt
 atataaaaggatctgtaaaagtcttacggcacaactccatgctgtgaaaaagtaggtatcgtgcctgtgacctcagggaacgggat
 cataggagctttccgaatccctgaatgcaattgccgggtatttcgggttgacagttttactgatagcctgacgtcagcggat
 tatgaggcagtaagaacggagcccgatcacttatggcagatgataaaccttccctgcccaacctgaaaaatggaaaaatc
 gccataaccagccgtgtacaggcataatfatgctgaaatagcttcaagatacctgaaagccgattcaaaagaagtgttgcctg
 ggtcttgggaagggttgattccgggagcagccatctcgtacagaaaggcttcaagggttacggatgatgctgacagaacctt
 ctgaaaaaaagcgttccagcctcggaaftatacttcaatccctcagccccgaagcgcagaggcaaaaggaaagtttccattttt
 cgaagcaacccctgtgcagacacgattccggaatccgtaatttcggaaaactgtgtgatttctacccctgggataccctgtgcaatc
 tcaaaaggagctgcaaaaaaagtgtggaagtgaactglaatggaaccactggggataggatagcatcaatgctgtattctgtactt
 aa

Secuencia de py/T: (SEC ID N°:5)

GGAAACCTGATCATGTAGATCGAATGGACTCTAAATCCGTTTCAGCCGGGTTAG
 ATCCCGGGGTTTCGCCA

Ejemplo 2: Incorporación específica del sitio de PCL generado biosintéticamente a hRBP4 usando células de mamífero.

Este ejemplo demuestra la incorporación específica del sitio de PCL a sitios codificados mediante TAG en la proteína humana modelo RBP4 (proteína 4 de unión a retinol humano). hRBP4 es una proteína secretada y su estructura se ha caracterizado, y los restos de hRBP4 expuestos al disolvente están bien definidos.

hRBP4 se clonó en el vector pRS con una etiqueta Flag-His en el extremo C (Figura 4A). hRBP4 se expresó bien en células HEK293F usando transfección transitoria. El codón TAG se introdujo en nueve posiciones en construcciones independientes de hRBP4 tal como se indica en la Tabla 2 y en las siguientes secuencias. Destacar que los codones TAG individuales se introdujeron en el codón de los restos de aminoácidos subrayados.

(A). Secuencia de aminoácidos de hRBP4: (SEC ID N°:6)

MKTFILLLWVLLLVVIFLLPGATAQPERDCRVSSFRVKENFDKARFSGTWY51A
 MAKKDPEGLF62LQDNIVA EFSVDETQMSATAKGRVLLNNW93DVCADMV
 GTFTDTEPAKFKMKY116W117GVASF122LQKGNDDHWIVD TDYDTY140AV
 QYSCRLNLDGTCADSYSFVFSRDPNGLPPEAQKIVRQRQEELCLARQY191RLI
 VHNGY199CDGRSERNLLDYKDDDDKHHHHH

5 (B). Secuencia de nucleótidos de los mutantes hRBP4 y TAG. hRBP4 natural: (SEC ID N°:7)

ATGAAAACATTCATACTCCTGCTcTGGGTACTGCTGCTCTGGGTTatcttctgcttc
 ccggtgccactgctcagcctgagcgcgactgccgagtgagcagctccgagtcaggagaacttcgacaaggctcgcttct
 ctgggacctggTACgccatggccaagaaggaccccgaggcctcTTTctgcaggacaacatcgtcgggagttctcc
 gtggacgagaccggccagatgagcggcacagccaagggccgagtcctgttttgaataacTGGgacgtgtgcgcagac
 atggtgggcaccttcacagacaccgaggaccctgccaaagtcaagatgaagTACTGGggcgtagcctccTTTctcc
 agaaaggaaatgatgacctggtcgtcgcacacagactacgacacgTATgccgtgcagtactcctgccgctcctgaac
 ctgcatggcacctgtgctgacagctactcctcgtgttttccgggacccaacggcctgccccagaagcgcagaagattgt
 aaggcagcggcaggaggagctgtgcctggccaggcagTACaggctgatcgtccacaacggTACTgcatggcaga
 tcagaaagaaaccttttgactataaagacgatgacgataagcatcaccatcaccatcacTAA

10 hRBP4 mutante 1: (SEC ID N°:8)

ATGAAAACATTCATACTCCTGCTcTGGGTACTIGCTGCTCTGGGTTatcttctgcttc
 ccggtgccactgctcagcctgagcgcgactgccgagtgagcagctccgagtcaggagaacttcgacaaggctcgcttct
 ctgggacctggTAGgccatggccaagaaggaccccgaggcctcTTTctgcaggacaacatcgtcgggagttctcc
 gtggacgagaccggccagatgagcggcacagccaagggccgagtcctgttttgaataacTGGgacgtgtgcgcagac
 atggtgggcaccttcacagacaccgaggaccctgccaaagtcaagatgaagTACTGGggcgtagcctccTTTctcc
 agaaaggaaatgatgacctggtcgtcgcacacagactacgacacgTATgccgtgcagtactcctgccgctcctgaac
 ctgcatggcacctgtgctgacagctactcctcgtgttttccgggacccaacggcctgccccagaagcgcagaagattgt
 aaggcagcggcaggaggagctgtgcctggccaggcagTACaggctgatcgtccacaacggTACTgcatggcaga
 tcagaaagaaaccttttgactataaagacgatgacgataagcatcaccatcaccatcacTAA

hRBP4 mutante 2: (SEC ID N°:9)

ATGAAAACATTCATACTCCTGCTcTGGGTACTGCTGCTCTGGGTTatcttctgcttc
 ccggtgccactgctcagcctgagcgcgactgccgagtgagcagctccgagtcaggagaacttcgacaaggctcgcttct
 ctgggacctggTACgccatggccaagaaggaccccgaggcctcTAGctgcaggacaacatcgtcgggagttctcc
 gtggacgagaccggccagatgagcggcacagccaagggccgagtcctgttttgaataacTGGgacgtgtgcgcagac
 atggtgggcaccttcacagacaccgaggaccctgccaaagtcaagatgaagTACTGGggcgtagcctccTTTctcc
 agaaaggaaatgatgacctggtcgtcgcacacagactacgacacgTATgccgtgcagtactcctgccgctcctgaac
 ctgcatggcacctgtgctgacagctactcctcgtgttttccgggacccaacggcctgccccagaagcgcagaagattgt
 aaggcagcggcaggaggagctgtgcctggccaggcagTACaggctgatcgtccacaacggTACTgcatggcaga
 tcagaaagaaaccttttgactataaagacgatgacgataagcatcaccatcaccatcacTAA

15

hRBP4 mutante 3: (SEC ID N°:10)

ATGAAAACATTCATACTCCTGCTcTGGGTACTGCTGCTCTGGGTTatcttctgcttc
 ccggtgccactgctcagcctgagcgcgactgccgagtgagcagctccgagtcaaggagaacttcgacaaggctcgttct
 ctgggacctggTACgcatggccaagaaggaccccgagggcctcTTTctgcaggacaacatcgtcgcggagtctcc
 gtggacgagaccggccaagatgagcggccacagccaaggggccgagtccttttgaataacTAGgacgtgtgcgcagac
 atggtgggcaccttcacagacaccgaggaccctgccaagtcaagatgaagTACTGGggcgtagcctccTTTctcc
 agaaaggaaatgatgaccactggatcgtcgcacagactacgacacgTATgccgtgcagtactcctgccgctcctgaac
 ctcgatggcacctgtgctgacagctactccttcgtgtttccgggaccccaacggcctgccccagaagcgcagaagattgt
 aaggcagcggcaggaggagctgtgcctggccaggcagTACaggctgatcgtccacaacgggTACTgcatggcaga
 tcagaaagaaaccttttgactataaagacgatgacgataagcatcaccatcaccatcacTAA

5 hRBP4 mutante 4: (SEC ID N°: 11)

ATGAAAACATTCATACTCCTGCTcTGGGTACTGCTGCTCTGGGTTatcttctgcttc
 ccggtgccactgctcagcctgagcgcgactgccgagtgagcagctccgagtcaaggagaacttcgacaaggctcgttct
 ctgggacctggTACgcatggccaagaaggaccccgagggcctcTTTctgcaggacaacatcgtcgcggagtctcc
 gtggacgagaccggccaagatgagcggccacagccaaggggccgagtccttttgaataacTGGgacgtgtgcgcagac
 atggtgggcaccttcacagacaccgaggaccctgccaagtcaagatgaagTAGTGGggcgtagcctccTTTctcc
 agaaaggaaatgatgaccactggatcgtcgcacagactacgacacgTATgccgtgcagtactcctgccgctcctgaac
 ctcgatggcacctgtgctgacagctactccttcgtgtttccgggaccccaacggcctgccccagaagcgcagaagattgt
 aaggcagcggcaggaggagctgtgcctggccaggcagTACaggctgatcgtccacaacgggTACTgcatggcaga
 tcagaaagaaaccttttgactataaagacgatgacgataagcatcaccatcaccatcacTAA

10 hRBP4 mutante 5: (SEC ID N°:12)

ATGAAAACATTCATACTCCTGCTcTGGGTACTGCTGCTCTGGGTTatcttctgcttc
 ccggtgccactgctcagcctgagcgcgactgccgagtgagcagctccgagtcaaggagaacttcgacaaggctcgttct
 ctgggacctggTACgcatggccaagaaggaccccgagggcctcTTTctgcaggacaacatcgtcgcggagtctcc
 gtggacgagaccggccaagatgagcggccacagccaaggggccgagtccttttgaataacTGGgacgtgtgcgcagac
 atggtgggcaccttcacagacaccgaggaccctgccaagtcaagatgaagTACTAGggcgtagcctccTTTctcc
 agaaaggaaatgatgaccactggatcgtcgcacagactacgacacgTATgccgtgcagtactcctgccgctcctgaac
 ctcgatggcacctgtgctgacagctactccttcgtgtttccgggaccccaacggcctgccccagaagcgcagaagattgt
 aaggcagcggcaggaggagctgtgcctggccaggcagTACaggctgatcgtccacaacgggTACTgcatggcaga
 tcagaaagaaaccttttgactataaagacgatgacgataagcatcaccatcaccatcacTAA

hRBP4 mutante 6: (SEC ID N°:13)

ATGAAAACATTCATACTCCTGCTcTGGGTACTGCTGCTCTGGGTTatcttctgcttc
 ccggtgccactgctcagcctgagcgcgactgccgagtgagcagctccgagtcaaggagaacttcgacaaggctcgttct
 ctgggacctggTACgcatggccaagaaggaccccgagggcctcTTTctgcaggacaacatcgtcgcggagtctcc
 gtggacgagaccggccaagatgagcggccacagccaaggggccgagtccttttgaataacTGGgacgtgtgcgcagac
 atggtgggcaccttcacagacaccgaggaccctgccaagtcaagatgaagTACTGGggcgtagcctccTAGctcc
 agaaaggaaatgatgaccactggatcgtcgcacagactacgacacgTATgccgtgcagtactcctgccgctcctgaac
 ctcgatggcacctgtgctgacagctactccttcgtgtttccgggaccccaacggcctgccccagaagcgcagaagattgt
 aaggcagcggcaggaggagctgtgcctggccaggcagTACaggctgatcgtccacaacgggTACTgcatggcaga
 tcagaaagaaaccttttgactataaagacgatgacgataagcatcaccatcaccatcacTAA

15

hRBP4 mutante 7: (SEC ID N°:14)

ATGAAAACATTCATACTCCTGCTcTGGGTACTGCTGCTCTGGGTTatcttctgcttc
 ccggtgccactgctcagcctgagcgcgactgccgagtgagcagctccgagtcaaggagaacttcgacaaggctcgcttct
 ctgggacctggTACgccatggccaagaaggaccccaggggcctcTTTctgcaggacaacatcgtcgggagttctcc
 gtggacgagaccggccaagatgagcggcacagccaaggggccgagtccgcttttgaataacTGGgacgtgtgcgcagac
 atggtgggcaccttcacagacaccgaggacctgccaagtcaagatgaagTACTGGggcgtagcctccTTTctcc
 agaaaggaaatgatgacctggatcgtcgcacagactacgacacgTAGgccgtgcagtactcctgccgcctcctgaa
 cctc gatggcacctgtgctgacagctactcctcgtglttcccggacccaacggcctgccccagaagcgagaagattg

taaggcagcggcaggaggagctgtgcctggccaggcagTACaggctgatcgtccacaacggTACTgcatggcag
 atcagaaaagaaccttttgactataaagacgatgacgataagcatcaccatcaccatcacTAA

5 **hRBP4 mutante 8: (SEC ID N°: 15)**

ATGAAAACATTCATACTCCTGCTcTGGGTACTGCTGCTCTGGGTTatcttctgcttc
 ccggtgccactgctcagcctgagcgcgactgccgagtgagcagctccgagtcaaggagaacttcgacaaggctcgcttct
 ctgggacctggTACgccatggccaagaaggaccccaggggcctcTTTctgcaggacaacatcgtcgggagttctcc
 gtggacgagaccggccaagatgagcggcacagccaaggggccgagtccgcttttgaataacTGGgacgtgtgcgcagac
 atggtgggcaccttcacagacaccgaggacctgccaagtcaagatgaagTACTGGggcgtagcctccTTTctcc
 agaaaggaaatgatgacctggatcgtcgcacagactacgacacgTATgccgtgcagtactcctgccgcctcctgaa
 cctc gatggcacctgtgctgacagctactcctcgtglttcccggacccaacggcctgccccagaagcgagaagattg
 aaggcagcggcaggaggagctgtgcctggccaggcagTAGaggctgatcgtccacaacggTACTgcatggcaga
 tcagaaaagaaccttttgactataaagacgatgacgataagcatcaccatcaccatcacTAA

10 **hRBP4 mutante 9: (SEC ID N°: 16)**

ATGAAAACATTCATACTCCTGCTcTGGGTACTGCTGCTCTGGGTTatcttctgcttc
 ccggtgccactgctcagcctgagcgcgactgccgagtgagcagctccgagtcaaggagaacttcgacaaggctcgcttct
 ctgggacctggTACgccatggccaagaaggaccccaggggcctcTTTctgcaggacaacatcgtcgggagttctcc
 gtggacgagaccggccaagatgagcggcacagccaaggggccgagtccgcttttgaataacTGGgacgtgtgcgcagac
 atggtgggcaccttcacagacaccgaggacctgccaagtcaagatgaagTACTGGggcgtagcctccTTTctcc
 agaaaggaaatgatgacctggatcgtcgcacagactacgacacgTATgccgtgcagtactcctgccgcctcctgaa
 cctc gatggcacctgtgctgacagctactcctcgtglttcccggacccaacggcctgccccagaagcgagaagattg
 aaggcagcggcaggaggagctgtgcctggccaggcagTACaggctgatcgtccacaacggTAGtgcgatggcaga
 tcagaaaagaaccttttgactataaagacgatgacgataagcatcaccatcaccatcacTAA

15 Tyr108 y Phe54 cubren el sitio de unión a retinol del hRBP4 mientras que el resto de mutaciones TAG son restos de proteína expuestos a disolvente. Cuando se expresa en células de mamífero, la terminación de la traducción en estos sitios TAG daría como resultado una proteína truncada que carece de la etiqueta His del extremo C. En consecuencia, el anticuerpo dirigido contra His no reconocerá la proteína truncada. Una traducción por lectura creará un producto de longitud completa con una etiqueta His. Por lo tanto, la detección mediante este anticuerpo dirigido contra His sirve como indicador de la actividad de lectura.

20 *Expresión de hRBP4:*

25 pRSRBP se transfectó simultáneamente en células HEK293F con pCMVpyS, pCMVpyB, pCMVpyC y pCMVpyD y las células se hicieron crecer en presencia de D-ornitina 5 mM tal como se ha descrito en el Ejemplo 1. A continuación, hRBP4 marcado con His se analizó mediante transferencia Western y anticuerpo dirigido contra His, y puesto que el hRBP4 truncado no contiene la etiqueta His, solamente el hRBP4 de longitud completa con PCL incorporado en el codón TAG se detectará en la transferencia Western. Así, la presencia de hRBP4 de longitud completa demuestra la incorporación de PCL. Tal como se indica en la Figura 6A, el anticuerpo dirigido contra His detectó una banda a 26 kDa, el tamaño esperado del hRBP4 de longitud completa, lo que indica la incorporación de PCL en el codón mutado con TAG. PCL se incorporó en nueve construcciones hRBP4 (Tabla 2) con tasas diferentes, donde las construcciones

mutantes 2, 6, y 9 tenían una tasa superior y las construcciones mutantes 7 y 8 la tasa inferior. El rendimiento de proteínas para 2, 6 y 9 estuvo entre 4 y 8 mg/ml, aproximadamente un 20 % del rendimiento para hRBP4 natural. Estas proteínas se purificaron y se sometieron a análisis por espectrometría de masas (MS). La masa obtenida para las proteínas fue consistente con lo esperado para la incorporación de PCL, y el análisis por MS en tándem confirmó la incorporación de PCL en el sitio esperado (Figura 6A a 11).

La Figura 6B muestra el SDS-PAGE y la Figura 6C muestra el espectro de masas de hRBP4 producido en células HEK293F en las que PCL se ha incorporado en hRBP4. hRBP4 se purificó del medio de las células HEK293F transfectadas simultáneamente en ausencia (A, carril 1) o presencia (A, carril 2) de D-ornitina, y se analizó mediante SDS-PAGE. La flecha indica el hRBP4 de longitud completa. La proteína purificada se analizó mediante espectrometría de masas (B): La masa observada es de 23166,0 Da, cercana a la masa esperada de 23168 Da.

La detección mediante espectrometría de masas de PCL incorporada de forma específica del sitio a hRBP4:

Un cultivo de 500 ml de células HEK293F transfectadas con los genes *pylS*, *pylT*, *pylC* y *pylD* y ADN del mutante 6 de hRBP4 (Tabla 2: hRBP4 Phe122PCL) se hizo crecer en medio suplementado con D-ornitina 5 mM. El medio se aplicó a dos columnas Ni-NTA para capturar toda la proteína hRBP4 de longitud completa. Cuando se combinaron las fracciones eluidas contenían 4,3 mg de proteína total. La masa observada de la proteína fue 23166,8 Da (esperado 23168 Da) consistente con la incorporación de un único resto PCL (no se muestran los datos). Una alícuota de la muestra se sometió a digestión trípica y a análisis por LC-MS.

El espectro MS/MS de la Figura 7 se podría asignar al péptido YWGVASF*LQK (SEC ID N°:17) mientras que el resto F* tuvo una masa consistente con la de PCL lo que confirma la incorporación específica del sitio de en el sitio TAG deseado en el resto 122. Las asignaciones del péptido YWGVASF*LQK se indican a continuación:

- Asignaciones para YWGVASF*LQK a m/z 647,86 (F* = PCL)
- Masa monoisotópica del péptido neutro Mr (calc): 1273,6819
- Modificaciones variables:
- F7 : PCL en F (F)
- Puntuación de iones: 60 Esperado: 0,0013
- Correspondencias (negrita): 22/74 fragmentos de iones usando los 32 picos más intensos

N°	b	b ⁺	b ⁺⁺	b ⁺⁺⁺	b ⁺⁺⁺⁺	Sec.	y	y ⁺	y ⁺⁺	y ⁺⁺⁺	y ⁺⁺⁺⁺	N°	
1	164,0706	82,5389				Y						10	
2	350,1499	175,5786				W	1111,6259	556,3166	1094,5993	547,8033	1093,6153	547,3113	9
3	407,1714	204,0893				G	9255465	463,2769	908,5200	454,7636	907,5360	454,2716	8
4	506,2398	253,6235				V	868,5251	434,7662	851,4985	426,2529	850,5145	425,7609	7
5	577,2769	289,1421				A	769,4567	385,2320	752,4301	376,7187	751,4461	376,2267	6
6	664,3089	332,6581		646,2984	323,6528	S	698,4195	349,7134	681,3930	341,2001	680,4090	340,7081	5
7	887,4410	444,2241		869,4304	435,2189	F	611,3875	306,1974	594,3610	297,6841			4
8	1000,5251	500,7662		982,5145	491,7609	L	388,2554	194,6314	371,2289	186,1181			3
9	1128,5837	564,7955	1111,5571	556,2822	1110,5731	Q	275,1714	138,0893	258,1448	129,5761			2
10						K	147,1128	74,0600	130,0863	65,5468			1

La Figura 8 muestra el análisis espectrométrico de la masa (TIC y EIC de 2+ iones de YWGVASF*LQK) de una digestión trípica de RBP4 Phe122PCL humano, e indica la incorporación de PCL en el sitio Phe122TAG diana. La Figura 9 muestra el análisis por espectrometría de masas de una digestión trípica de RBP4 Phe122PCL humano. El espectro de masas muestra 3+ y 2+ precursores de YWGVASF*LQK a m/z 425,57 (3+) y 637,85 (2+), respectivamente. (F* = PCL), indicando adicionalmente la incorporación de PCL en el sitio Phe122TAG diana.

Ejemplo 3: Detección de PCL biosintéticamente generado en lisado procedente de células de mamífero.

Este ejemplo demuestra que PCL se genera biosintéticamente en células de mamífero y se puede detectar como aminoácido libre en el lisado de dichas células. Además, este ejemplo sugiere que PCL se incorpora de manera análoga a pirrolisina y los otros 20 aminoácidos naturales, y que la incorporación de PCL no es el resultado de una modificación de la proteína posterior a la traducción.

Detección de PCL en el lisado celular

Cultivos de 60 ml de células HEK293F se hicieron crecer con o sin D-ornitina 5 mM y con los genes *pylT*, *pylS*, *pylB*, *pylC* y *pylD*. Las células se recogieron y se lisaron mediante sonicación en 250 ml de F₂O doblemente desionizado. Los residuos celulares se aglomeraron mediante centrifugación. La proteína se precipitó por adición de metanol frío hasta una concentración final del 80 % y se retiraron mediante centrifugación. A continuación, el lisado soluble se concentró mediante speedvac y se analizó mediante cromatografía líquida de alta presión combinada con espectrometría de masas para demostrar la presencia de PCL.

Las muestras secas se reconstituyeron en primer lugar en 100 ml de Fase móvil A/Fase móvil B 98/2 (fase móvil A. agua con ácido fórmico al 0,1 %; fase móvil B: acetonitrilo con ácido fórmico al 0,1 %). A continuación, las muestras se diluyeron 20 veces, y 2 ml se inyectaron en el HPLC. La separación del analito se realizó en una columna HPLC Agilent

zorbax 300SB_C18 inversa usando el siguiente gradiente de disolución: 0 min 2 %B; 5 min 2 %B; 60 min 100 %B; Caudal: 0,25 ml/min. La comparación de las trazas del HPLC (Figura 10A) muestra un pico a 4,13 min (indicado por un asterisco) que está presente en el lisado de las células transfectadas con los genes biosintéticos pylb, pylc y pyld y que han crecido en presencia de D-ornitina ((cromatograma de iones extraídos), traza del EIC inferior), pero no el lisado procedente de células que han crecido en ausencia de D-ornitina (traza del EIC superior). Un espectro de masas completo (Figura 10C) del pico HPLC a 4,13 min indicó una masa de 242,14943 Da consistente con la masa teórica de PCL (242,14992 Da), la versión desmetilada de PYL. La lisina (pico HPLC a 1,44 min) es igualmente abundante en ambas muestras (Figura 10B) y un espectro de masas completo de la lisina (Figura 10D) sirve como calibración interna e ilustra la precisión del método. Este análisis indica que PCL se genera a partir de D-ornitina como aminoácido libre detectable en el lisado celular.

Ejemplo 4: Ensayo de incorporación de los posibles precursores PCL y pirrolisina, con análogos sintéticos de pirrolisina, y con diferentes combinaciones de genes biosintéticos

Este ejemplo demuestra que D-ornitina es un precursor preferido de la biosíntesis de PCL tal como se determina mediante su incorporación específica de emplazamiento en la RBP4 humana. También se ilustra en este ejemplo la incorporación específica del sitio de determinados análogos de pirrolisina, incluyendo CYC, a sitios codificados mediante TAG en la proteína humana modelo RBP4, proteína 4 de unión a retinol humano usando células de mamífero. Este ejemplo también proporciona una visión de la especificidad de sustrato de la pylS ARNt sintetas.

Incorporación de N-e-ciclopentiloxicarbonil-L-lisina (CYC)

Construcciones de ADN para nueve mutantes hRBP4 TAG (Tabla 3) se transfectaron simultáneamente de forma individual a células HEK293F junto con ADN de pylT/pylS como se ha descrito en el Ejemplo 2 y se cultivaron en ausencia y presencia de N-e-ciclopentiloxicarbonil-L-lisina (CYC) 4 mM, un análogo de pirrolisina. Los medios de cultivo se recogieron y analizaron mediante transferencia Western con anticuerpos dirigidos contra His y contra hRBP4 (Figura 11A). La transferencia Western con el anticuerpo dirigido contra His reveló una proteína de 24 kDa para seis de las nueve construcciones de hRBP4 TAG mutante. El tamaño de la proteína es equivalente al tamaño del hRBP4 de tipo natural, lo que indica la producción de una proteína de longitud completa mediante la traducción con lectura. La actividad de lectura varió entre las construcciones hRBP4 mutantes (Tabla 3) y fue dependiente de CYC y pylS. En particular, la proteína hRBP4 de longitud completa no se pudo detectar para los mutantes 1, 5 y 7. Se observaron elevados rendimientos de proteína de longitud completa se observaron para los mutantes 2, 6 y 9, de forma similar a los resultados observados para la incorporación de PCL (Ejemplo 2, Figura 6A).

Tabla 3

Mutante de hRBP4	Número de mutante	Incorporación de CYC
Tyr51TAG	1	No detectado por Western
Phe62TAG	2	Fuerte
Trp93TAG	3	Muy débil
Tyr116TAG	4	Débil
Trp117TAG	5	Detectable
Phe122TAG	6	Fuerte
Tyr140TAG	7	No detectado
Tyr191TAG	8	Medio
Tyr199TAG	9	Fuerte

Las proteínas expresadas que contienen CYC se purificaron del medio por cromatografía Ni-NTA y se analizaron mediante SDS-PAGE seguido por tinción con azul de Coomassie. La Figura 11B muestra el análisis SDS-PAGE del mutante 2 de hRBP4 TAG purificado. La preparación purificada muestra una sola banda de proteína con un tamaño de 24 kDa. El espectro de masa de la proteína purificada fue consistente con una incorporación en un solo sitio de CYC (Figura 11C). El peso molecular esperado para la incorporación de CYC en un solo sitio en lugar de Phe62 fue (23114,6 Da (natural) - 165,2 Da (Phe) + 258,3 Da (CYC)) 23208,7 Da, y la observada 23182,0 Da. La masa observada sugiera la ciclación del resto Q del extremo N a ácido pirrolidonacarboxílico dando como resultado la pérdida de 18 Da y la presencia de tres enlaces disulfuro intactos que dan como resultado la pérdida de 6 Da (23208,7 Da - 6 Da - 18 Da = 23184,7 Da). El análisis por MS en tándem de la proteína purificada demostró que CYC se incorporó al sitio designado por TAG en la construcción hRBP4 (Figura 12). Se ha estimado que el rendimiento de la proteína debido a la transfección transitoria es de aproximadamente 5 mg/l. La fragmentación MS/MS de KDPEGLFLQDNIVAIEFSVDET-GQMSATAK (SEC ID N°: 18) es

Masa monoisotópica del péptido neutro Mr (calc): 3248,5646 Modificaciones variables:

F7 : Cys en F (F)

Q9 : Desamidado (NQ)

Nil: Desamidado (NQ)

M24: Oxidación (M), con pérdida neutra 0,0000

(se muestra en la tabla), 63,9983

Puntuación de iones: 113 Esperado: 4,6e-009

Correspondencias (negrita): 110/498 fragmentos de iones usando los 156 picos más intensos

Nº	b	b ⁺⁺	b ⁺	b ⁺⁺⁺	b ⁰	b ⁺⁺⁺	Sec.	y	y ⁺⁺	y ⁺	y ⁺⁺	y ⁺	y ⁺⁺⁺	nº
1	129,1022	65,0548	112,0757	56,5415										29
2	244,1292	122,5682	227,1026	114,0550	226,1186	113,5629	D	3121,4967	1561,2421	3104,4504	1552,7288	3103,4664	1552,2368	28
3	341,1819	171,0946	324,1554	162,5813	323,1714	162,0893	P	3006,4500	1503,7286	2989,4234	1495,2154	2988,4394	1494,7233	27
4	470,2245	235,6159	453,1980	227,1026	452,2140	226,6106	E	2909,3972	1455,2022	2892,3707	1446,6890	2891,3867	1446,1970	26
5	527,2460	264,1266	510,2195	255,6134	509,2354	255,1214	G	2780,3546	1390,6810	2763,3281	1382,1677	2762,3441	1381,6757	25
6	640,3301	320,6687	623,3035	312,1554	622,3195	311,6634	L	2723,3332	1362,1702	2706,3066	1353,6569	2705,3226	1353,1649	24
7	878,4982	439,7527	861,4716	431,2395	860,4876	430,7475	F	2610,2491	1305,6282	2593,2226	1297,1149	2592,2385	1296,6229	23
8	991,5823	496,2948	974,5557	487,7815	973,5717	487,2895	L	2372,0810	1186,5441	2355,0544	1178,0308	2354,0704	1177,5388	22
9	1120,6248	560,8181	1103,5983	552,3028	1102,6143	551,8108	Q	2258,9969	1130,0021	2241,9704	1121,4888	2240,9863	1120,9968	21
10	1235,6518	618,3295	1218,6252	609,8163	1217,6412	609,3243	D	2129,9243	1065,4808	2112,9278	1056,9675	2111,9438*	1056,4755	20
11	1350,6787	675,8430	1333,6522	667,3297	1332,6682	666,8377	N	2014,9274	1007,9673	1997,9008	999,4541	1996,9168	998,9620	19
12	1463,7628	732,3850	1446,7362	723,8718	1445,7522	723,3798	I	1899,9004	950,4539	1882,8739	941,9406	1881,8899*	941,4486	18
13	1562,8312	781,9192	1545,8047	773,4060	1544,8206	772,9140	V	1786,8164	893,9118	1769,7898	885,3986	1768,8058	884,9065	17
14	1633,8683	817,4378	1616,8418	808,9245	1615,8578	808,4325	A	1687,7480	844,3776	1670,7214	835,8643	1669,7374	835,3723	16
15	1762,9109	881,9591	1745,8844	873,4458	1744,9003	872,9538	E	1616,7108	808,8591	1599,6843	800,3458	1598,7003	799,8538	15
16	1909,9793	955,4933	1892,9528	946,9800	1891,9688	946,4880	F	1487,6683	744,3378	1470,6417	735,8245	1469,6577	735,3325	14
17	1997,0114	999,0093	1979,9848	990,4960	1979,0008	990,0040	S	1340,5998	670,8036	1323,5733	662,2903	1322,5893	661,7983	13
18	2096,0798	1048,5435	2079,0532	1040,0302	2078,0692	1039,5382	V	1253,5678	627,2875	1236,5413	618,7743	1235,5572	618,2823	12
19	2211,1067	1106,0570	2194,0802	1097,5437	2193,0961	1097,0517	D	1154,4994	577,7533	1137,4729	569,2401	1136,4888	568,7481	11
20	2340,1493	1170,5783	2323,1228	1162,0650	2322,1387	1161,5730	E	1039,4725	520,2399	1022,4459	511,7266	1021,4619	511,2346	10
21	2441,1970	1221,1021	2424,1704	1212,5889	2423,1864	1212,0968	T	910,4299	455,7186	893,4033	447,2053	892,4193	446,7133	9
22	2498,2184	1249,6129	2481,1919	1241,0996	2480,2079	1240,6076	G	809,3822	405,1947	792,3556	396,6815	791,3716	396,1894	8
23	2626,2770	1313,6422	2609,2505	1305,1289	2608,2665	1304,6369	Q	752,3607	376,6840	735,3342	368,1707	734,3502*	367,6787	7
24	2773,3124	1387,1599	2756,2859	1378,6466	2755,3019	1378,1546	M	624,3021	312,6547	607,2756	304,1414	606,2916	303,6494	6
25	2860,3445	1430,6759	2843,3179	1422,1626	2842,3339	1421,6706	S	477,2667	239,1370	460,2402	230,6237	459,2562	230,1317	5
26	2931,3816	1466,1944	2914,3550	1457,6811	2913,3710	1457,1891	A	390,2347	195,6210	373,2082	187,1077	372,2241	186,6157	4
27	3032,4293	1516,7183	3015,4027	1508,3050	3014,4187	1507,7130	T	319,1976	160,1024	302,1710	151,5892	301,1870	151,0972	3
28	3103,4664	1552,2368	3086,4398	1543,7235	3085,4558	1543,2315	A	218,1499	109,5786	201,1234	101,0653			2

29							K	147,1128	74,0600	130,0863	65,5468			1
----	--	--	--	--	--	--	---	----------	---------	----------	---------	--	--	---

Ensayo de incorporación de los posibles precursores PCL y pirrolisina

La producción de una proteína hRBP4 Phe122PCL de longitud completa (mutante 6) se midió en presencia del posible precursor de PCL D-ornitina, D-prolina, D-arginina, ácido D-glutámico, 4-hidroxil-D-prolina y ácido 2-pirrolidona-5-carboxílico (Figura 13A y B, Figure 14A). pRSRBP se transfeció simultáneamente en células HEK293F con pCMVpyS, pC-MVpyB, pCMVpyC y pCMVpyD y las células se hicieron crecer en presencia de 5 mM del posible precursor (Figura 14A) como se ha descrito en el Ejemplo 2. hRBP4 con la etiqueta His y por tanto de longitud completa con PCL incorporado se analizó posteriormente mediante transferencia Western y anticuerpo dirigido contra His y SDS-PAGE. Tal como se muestra en la Figura 13B. solamente la D-ornitina da lugar a una banda de proteína medible en SDS-PAGE de las muestras purificadas mediante Ni-NTA. La detección de las muestras no purificadas mediante transferencia Western (Figura 13A) indica solamente un bajo nivel de formación de proteína hRBP4 de longitud completa en presencia del resto de precursores, lo que indica claramente que D-ornitina es el precursor más eficaz para la biosíntesis e incorporación de PCL. De los análisis realizados por transferencia Western y SDS-PAGE, parece que la generación biosintética de PCL a partir de D-ornitina 5 mM (banda 2) y la posterior incorporación de la proteína es más eficaz que la incorporación del conocido sustrato PyIS CYC añadido a 5 mM (banda 9) al medio de células transfectadas con pCMVpyS. Este ejemplo demuestra que D-ornitina es el precursor preferido para la biosíntesis de PCL.

Ensayo de incorporación con análogos sintéticos de pirrolisina

En paralelo con el experimento anterior, también se determinó la producción de la proteína hRBP4 Phe122PCL de longitud completa (mutante 6) en presencia de una serie de análogos sintéticos de pirrolisina diseñados con un resto acetilo para su derivatización química tras la incorporación a la proteína (Figura 14B). Los análogos sintéticos se prepararon como se describe en el Ejemplo 19. pRSRBP se transfeció simultáneamente en células HEK293F con pCMVpyS que suministra una copia génica de ARNt de pylT y de ARNt de pylS sintetasa. Dependiendo de la solubilidad, los análogos sintéticos se añadieron al medio a una concentración final de 2 o 5 mM. La detección de hRBP4 con la etiqueta His y por tanto de longitud completa con PCL incorporado se analizó posteriormente mediante transferencia Western y anticuerpo dirigido contra His y SDS-PAGE. Tal como se muestra en la Figura 13B. solamente CYC (banda 9) da lugar a una banda de proteína medible en las muestras purificadas mediante Ni-NTA. Los análisis mediante transferencia Western de las muestras no purificadas (Figura 13A) sugiere que TU3000-016 (banda 15) es también un sustrato viable de la ARNt de pylS sintetasa. Sin embargo, la eficacia de la incorporación es baja y el compuesto no es estable, ya que un lote diferente (TU2982-150) experimentó degradación durante su medida mediante espectroscopía de resonancia magnética nuclear, dando como resultado una incorporación a hRBP4 (banda 10) aún más baja. Este ejemplo, de acuerdo con los informes publicados, demuestra que los análogos de pirrolisina que tienen un carbono sp² en el punto de unión del resto del anillo del análogo no se toleran como sustratos de la ARNt de pylS sintetasa.

Ensayos de incorporación con diferentes combinaciones de genes biosintéticos

También se ensayó la producción de una hRBP4 Phe122PCL de longitud completa con cultivos celulares con D-ornitina 5 mM y transfección simultánea con diferentes combinaciones de los genes biosintéticos pylB, pylC y pylD (Figura 13C). Todos los cultivos se transfectaron simultáneamente con pRSRBP que suministra una copia génica de ARNt de pylT y de ARNt de pylS sintetasa. La proteína hRBP4 de longitud completa solo se detectó mediante transferencia Western con el anticuerpo dirigido contra la etiqueta His cuando pylC y pylD se transfectaron simultáneamente. pylB, aunque igualmente necesario para la biosíntesis de PCL, no es esencial para la biosíntesis de PCL y su posterior incorporación. Esta observación se confirmó adicionalmente por la producción de mutantes PCL de longitud completa de la proteína del dominio Fc de mIgG 1 (Ejemplo 5) y EPO de ratón y ser humano (Ejemplo 6). Los tres ejemplos ilustran que solamente los genes pylC y pylD son esenciales para la biosíntesis de PCL y su posterior incorporación a la proteína.

Ejemplo 5: Incorporación específica de sitio de PCL generada biosintéticamente en el dominio Fc de la IgG 1 de ratón usando células de mamífero.

Este ejemplo demuestra que la incorporación específica de sitio de PCL generada biosintéticamente en células de mamífero utilizando los métodos proporcionados en el presente documento es el procedimiento general y no está limitado a la proteína hRBP4.

Expresión de Fc de mIgG1

Se generaron cuatro mutantes TAG n K333 (mutante n° 1), K336 (mutante n° 2), T394 (mutante n° 3) y L426 (mutante n° 4) del dominio Fc de la IgG 1 de ratón (Tabla 2 y Figura 17A). pRSFc n° 1-4 (Tabla 2) se transfectaron simultáneamente en células HEK293F con pCMVpyT, pCMVpyS, pCMVpyC y pCMVpyD y se hicieron crecer células en presencia de D-ornitina 5 mM (pCMVpyB no se añadió). La proteína del dominio Fc etiquetada con His se purificó a partir del medio mediante cromatografía Ni-NTA y se analizó en SDS-PAGE. Los tamaños de las bandas de proteína del gel para cada construcción son consistentes con su tamaño esperado para las proteínas de longitud completa (Figura 17A). La construcción de la expresión caracteriza una etiqueta His6 en el extremo C para la purificación y por tanto solo se recupera la proteína de longitud completa. Tal como se muestra en la Figura 17A, la expresión del mutante n° 1 depende de la adición de D-ornitina al medio de crecimiento. Los niveles de expresión de los cuatro mutantes son similares. El análisis por espectrometría de masas no se llevó a cabo debido a la glicosilación del dominio Fc cuando se produjo en células HEK293F.

Ejemplo 6: Incorporación específica de sitio de la PCL generada a eritropoyetina (EPO) utilizando células de mamífero.

En este ejemplo, PCL se incorpora específicamente en el sitio en la eritropoyetina, lo que demuestra adicionalmente que los métodos proporcionados en el presente documento son un procedimiento general para la incorporación de PCL generada biosintéticamente específica del sitio a las proteínas en células de mamífero.

Expresión de eritropoyetina de ratón (EPO)

La incorporación de PCL en proteínas mutantes EPO se llevó a cabo en células HEK293F. Se introdujeron mutaciones TAG en restos Lys o ARG expuestos en once superficies colocándolos separados de la interfaz de unión al receptor EPO. Los sitios de incorporación se muestran en la parte inferior y se relacionan en la Tabla 2. La proteína EPO de ratón se expresó tal como se describe en el Ejemplo 1 excepto que no se añadieron los genes pylB: pRSEPO n° 1-11 (Tabla 2) se transfectaron simultáneamente en células HEK293F con pCMVpyT, pCMVpyS, pCMVpyC y pCMVpyD y se hicieron crecer células en presencia de D-ornitina 5 mM. EPO etiquetada con His se purificó a partir del medio mediante cromatografía Ni-NTA y se analizó en SDS-PAGE. Las construcciones EPO contienen una etiqueta en el extremo C. Por tanto, solo la proteína con PCL incorporada satisfactoriamente producirá la proteína de longitud completa, se purificará mediante cromatografía Ni-NTA y dará por tanto una banda detectable en SDS-PAGE. Un SDS-PAGE de EPO de longitud completa indica la incorporación satisfactoria de PCL para cada once mutantes (Figura 17B). Los tamaños de las bandas de proteína del gel para cada construcción son consistentes con su tamaño esperado para las proteínas de longitud completa. Para el mutante n° 1, se ilustra que la formación de la proteína de longitud completa depende de la D-ornitina. Los niveles de expresión de las once proteínas mutantes son similares y entre 10 y 20 % del peso de la proteína que se expresa a -40 mg/l (Figura 17B). El análisis por espectrometría de masas no se llevó a cabo debido a la glicosilación de EPO en células HEK293F. Se proporcionan a continuación las secuencias de la proteína EPO humana madura (SEC ID N°:19) y de ratón (SEC ID N°:20), con los sitios de incorporación de PCL en negrita y los sitios de glicosilación subrayados. La numeración de los mutantes comienza desde el extremo N (véase también la Tabla 2).

EPO

Humana APPRLICDSRVLEERYLLEAKEAE NITTGCAEHCSLENITVPDTKVNIFYAW KRMEVGGQAVEVW QGLALLS
De ratón APPRLICDSRVLEERYLLEAKEAE NVTMGCAEGPRLSENITVPDTKVNIFYAW KRMEVEEQAI EVW QGLSLLS

EAVLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLRRLGAQ KEAISPPDAASAAPLRTITADTFKLF
 EAILQAQALLANSSQPPEPLQLHIDKAISGLRSLTSLRVLGAQ KELMSPPDTTPPAPLRTLTVDTFCKLF

RVYSNFLRGKCLKLYTGEA CRTGDR
 RVYANFLRGKCLKLYTGEVCRRGD R

- 5 **Ejemplo 7:** Plásmido para la incorporación de PCL generada biosintéticamente a proteínas usando células de *Escherichia coli*.

Se proporciona en este ejemplo una descripción del plásmido que, cuando se transforma en células de *Escherichia coli*, permite la incorporación de PCL en sitios codificados por TAG en una proteína diana expresada a partir de un segundo plásmido transformado simultáneamente en la misma célula de *Escherichia coli*.

Construcción de pArra-pylSTBCD

En la Figura 5 se muestra un plásmido para la incorporación de PCL en células de *E. coli* que se construyó como sigue:
 15 Un casete que codifica pylT en el promotor proK se sintetizó mediante la amplificación del promotor, ARNt, y la secuencia 3' procedente de pSUP con cebadores solapantes. A continuación se mezclaron las tres piezas con los cebadores terminales que codifican los sitios de restricción de ApaI y XhoI a fin de sintetizar la inserción completa. Tras la digestión con ApaI y XhoI, el casete se ligó en una estructura pSUP que se había preparado con las mismas enzimas para preparar pSUP-pylT. Las secuencias de codificación de pylS, pylB, pylC, y pylD procedentes de *M. mazei* se amplificaron a partir de los plásmidos pCMVU6 adecuados anteriormente preparados (Ejemplo 1) y se insertaron en pMH4 sin etiquetas (véase, Lesley SA, Kuhn P, Godzik A, Deacon AM, Mathews I, Kreusch A, Spraggon G, Klock HE, McMullan D, Shin T, Vincent J, Robb A, Brinen LS, Miller MD, McPhillips TM, Miller MA, Scheibe D, Canaves JM, Guda C, Jaroszewski L, Selby TL, Elsliger MA, Wooley J, Taylor SS, Hodgson KO, Wilson IA, Schultz PG, y Stevens RC, "Structural genomics of the *Thermotoga maritima* proteome implemented in a high-throughput structure determination pipeline", *Proc Natl Acad Sci USA* 2002; 99: 11664-11669). El promotor-CDS-terminador completo de cada uno se amplificó a continuación con los cebadores con diferentes sitios de restricción (KpnI-pylS-SbfI, NdeI-pylD-BglII, BglII-pylB-XhoI, XhoI-pylC-KpnI). El producto pylS PCR se digirió con KpnI y SbfI y se ligó en pSUP-pylT entre los sitios KpnI y PstI para dar pAra-pylST. El pylB, pylC, y los productos pylD se cortan con las respectivas enzimas y se ligan en una estructura de plásmido preparada con NdeI y KpnI. Tras la verificación del producto plásmido de esta ligadura de cuatro puntos mediante secuenciación y diagnóstico de la PCR, se amplificó a continuación el casete completo con cebadores para añadir sitios KpnI a ambos extremos del casete pylD-pylB-pylC. Este se digirió con KpnI y se ligó en pAra-pylST en el sitio KpnI para dar el pAra-pylSTBCD final (Figura 5). El plásmido final contiene pylD, pylB, pylC, y pylS cada uno bajo el control de promotores independientes del híbrido T7 e inducibles por arabinosa y con el terminador rrnB en la dirección 3' de cada uno (idéntico al contexto de pMH4) y el gen pylT bajo el promotor proK (similar al contexto pSUP/pSUPAR con una copia de ARNt individual) (Cellitti y col., "In vivo incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy", *J Am Chem Soc.* 23 jul 2008; 130(29):9268-81).

- 40 **Ejemplo 8:** Incorporación específica del sitio de PCL generada biosintéticamente a FAS-TE usando células de *Escherichia coli*.

Este ejemplo proporciona una descripción de la incorporación de PCL en sitios codificados con TAG de la ácido graso sintetasa humana (FAS-TE) expresada a partir de un segundo plásmido transformado simultáneamente en la misma célula de *Escherichia coli*.

45 El dominio de la tioesterasa que codifica los restos 2221 a 2502 de la ácido graso sintetasa humana (FAS-TE) se expresa a partir de un vector pMH4 con una etiqueta en el extremo N (MGDSKIHSHHHHHENLYFQ g) (SEC ID N°:21) (véase, Lesley SA, Kuhn P, Godzik A, Deacon AM, Mathews I, Kreusch A, Spraggon G, Klock HE, McMullan D, Shin T, Vincent J, Robb A, Brinen LS, Miller MD, McPhillips TM, Miller MA, Scheibe D, Canaves JM, Guda C, Jaroszewski L, Selby TL, Elsliger MA, Wooley J, Taylor SS, Hodgson KO, Wilson IA, Schultz PG, y Stevens RC, "Structural genomics of the *Thermotoga maritima* proteome implemented in a high-throughput structure determination pipeline", *Proc Natl Acad Sci USA* 2002; 99: 11664-11669). Los codones TAG se mutaron usando la clonación PIPE (véase, Klock, HE, Koesema EJ, Knuth MW, y Lesley, SA, "Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscreening to accelerate structural genomics efforts", *Proteins.* 1 de mayo de 2008; 71(2):982-94) como se describe en detalle en Cellitti y col., "In vivo incorporation of unnatural amino acids to probe structure, dynamics, and ligand binding in a large protein by nuclear magnetic resonance spectroscopy", *J Am Chem Soc.* 23 de julio de 2008;130(29):9268-81. Los mutantes que se ensayaron para la incorporación de PCL fueron Leu2222TAG/Leu2223Ile e Y2454TAG (Figura 18A). Las células HK100 se transformaron simultáneamente con un

plásmido pMH4-FAS-TE y pAra-pylSTBCD y se seleccionaron en placas LB+Kan+Cm. Se hicieron crecer cultivos líquidos a 37 °C en TB (Sigma)+Kan+Cm suplementado con 5 mM de D-ornitina (Sigma o Nova Biochem) antes de la inducción. A DO₅₉₅ = 0,8, las células se movieron a 30 °C y se indujeron 15-30 minutos después con arabinosa al 0,2 %. Las células se hicieron crecer durante aproximadamente 20 horas después de la inducción antes de la cosecha mediante centrifugación. Las células se lisaron en TBS+ glicerol al 5 % pH 8 mediante sonicación. La fracción de proteína soluble se purificó mediante cromatografía Ni-NTA (Qiagen) de acuerdo con el protocolo del fabricante. Los rendimientos son de 46-80 mg/l para FAS-TE Leu2222PCL/Leu2223Ile y 155-186 mg/l para FAS-TE Tyr2454PCL, comparable a entre 50 y 80 % de los rendimientos de la proteína natural. El tamaño molecular en SDS-PAGE y el peso molecular de las proteínas determinado por espectrometría de masas es consistente con la incorporación de sitio único de PCL en la localización deseada (Figura 18B).

La secuencia de FAS-TE (SEC ID N°:22) con los dos sitios de incorporación de PCL (negrita y subrayado) se proporcionan a continuación (para Leu2222PCL, el resto Leu2223 está mutado para Ile2223):

MGSDKIH^{HHHH}HENLYFQGS^{LL}VNPEGPTLMRLNSVQSSERPLFLVHPIEGSTTVF
HSLASRLSIPTYGLQCTRAAPLDSIHSLAAYYIDCIRQVQPEGPYRVAGYSYGACVA
FEMCSQLQAQQSPAPTHNSLFLFDGSPTYVLA^{YT}QSYRAKLT^{PG}SEAEAETE^{AIC}FF
VQQFTDMEHNRVLEALLPLKGLEERVA^{AA}VDLIKSHQGLDRQELSFAARSFYYK
LRAAEQYTPKAKYHGNVMLLRAK^{TGGA}YGEDLGADYNLSQVCDGKVS^{VH}VIEG
DHRTLLEGSGLESII^{SI}HSSLA

Ejemplo 9: Incorporación específica del sitio de PCL generada biosintéticamente en FKBP12 usando células de Escherichia coli.

Este ejemplo describe la incorporación de PCL en sitios codificados con TAG de FKBP12 expresada a partir de un segundo plásmido transformado simultáneamente en la misma célula de Escherichia coli.

FKBP12 se expresa a partir del vector pET (Novagen) con una etiqueta en el extremo N (MGSSHHHHHLEVL^{FQGP}) (SEC ID N°:23). Se introdujo un codón TAG en la posición Ile90 utilizando la clonación PIPE (véase Klock, HE, Koesema EJ, Knuth MW, Lesley, SA, "Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscreening to accelerate structural genomics efforts", Proteins. 1 de mayo de 2008;71(2):982-94). células BL21(DE3) (Invitrogen) se transfectoron simultáneamente con pET-FKBP12 y pAra-pylSTBCD y se seleccionaron en placas LB+Kan+Cm. se hicieron crecer cultivos líquidos a 37 °C en TB+Kan+Cm suplementado con 5 mM de D-ornitina. A DO₅₉₅ = 0,4, las células se movieron a 30 °C y se indujeron 30 minutos después con IPTG 1 mM. Se hicieron crecer células unas 20 horas más antes de la cosecha. Las células se lisaron y purificaron como para FAS-TE. La proteína purificada se dializó a continuación en TBS y se cortó con proteasa HRV-3C (2U por 1 mg de FKBP12) durante 48 horas a 4 °C para eliminar la etiqueta His. El material del corte se dejó fluir a través de una columna Ni-NTA, se recogió, se concentró, y se hizo avanzar sobre una columna de exclusión por tamaño S75 (GE Healthcare) para purificación adicional. La proteína final se concentró adicionalmente a 15-20 mg/ml antes de la cristalización. El resultado para FKBP Ile90PCL fue de 120 mg/l en bruto. La FKBP Ile90PCL se va a usar para obtener una estructura mediante rayos X de una proteína que contiene PCL. La Figura 19A-C muestra la SDS-PAGE y los datos de espectrometría de masas y los resultados de la cristalización dan como resultado para la PCL biosintéticamente incorporada en FKBP12. La masa obtenida fue de 12085,6 Da, que es consistente con el valor esperado de 12084 Da para la incorporación del sitio único de la PCL.

La secuencia de FKBP12 (SEC ID N°:24) con los sitios de incorporación de PCL (negrita y subrayado) se proporciona a continuación:

MGSSHHHHHLEVL^{FQGP}VQVETISPGDGRTPFKRGQTCVVHYTG^{MLEDG}GKKF
DSSRDRNKPFK^FMLGKQE^{VI}RGWEEGVAQMSV^GQRAKLTISP^{DY}AYGATGHP^{GII}
PPHATLVFDVELL^{KLE}

Ejemplo 10: Incorporación específica del sitio de PCL generada biosintéticamente en el factor de crecimiento de fibroblastos 21 (FGF21) usando células de Escherichia coli.

Se proporciona en este ejemplo la incorporación de PCL en 20 sitios codificados por TAG por separado del factor de crecimiento de fibroblastos humano, FGF21 expresado a partir de un segundo plásmido transformado

simultáneamente en la misma célula de Escherichia coli.

El factor de crecimiento de fibroblastos 21, FGF21 se expresó a partir del vector pSpeedET (véase, Klock, HE, Koesema EJ, Knuth MW, Lesley, SA, "Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscreening to accelerate structural genomics efforts", Proteins. 1 de mayo de 2008;71(2):982-94) con una etiqueta en el extremo N (MGDSKIH¹¹HHHHHENLYFQ g) (SEQ ID NO:21) y que codifica los restos 33-209 de la proteína humana traducida. Los codones TAG para la incorporación del análogo de PYL y la posterior unión a PEG se introdujeron en la construcción aa33-209 de FGF21 (SEC ID N°:25) en las siguientes 20 posiciones individuales: Ser35, Gln39, Arg47, Gln56, Arg64, Asp74, Lys84, Lys87, Lys97, Arg100, Arg105, His115, Arg124, Glu129, Lys150, Arg154, Leu167, Leu170, Leu181 y Gln184. La secuencia de la construcción FGF21 (33-209 aa) se muestra a continuación con los sitios de incorporación en las 20 construcciones destacados en negrita y subrayados.

MGSSHHHHHHS **SGENLYFQGD** **SSPLLQFGGVRQRYLYTDD** **AQQTEAHLEI**
REDGTVGGAA **DQSPELLQL** **KALKPGVIQI** **LGVKTSRFLC** **QRPDGALYGS**
LHFDPEACSF **RELLLEDGYN** **VYQSEAHGLP** **LHLPGNKSPH** **RDPAPRGP**
FLPLPGLPPA **LPEPPGILAP** **QPPDVGSSDP** **LSMVGPSQGR** **SPSYAS**

Tanto las células HK100 como las células BL21(DE3) se transformaron simultáneamente con un plásmido pSpeedet-FGF21 y pAra-pylSTBCD y se seleccionaron en LB+Kan+Cm. Se hicieron crecer cultivos líquidos como en el Ejemplo 8, añadiendo 5 mM de D-ornitina antes de la inducción. Se ensayó la DO₅₉₅ para la interrupción a 30°C entre 0,2 y 1,0 con una inducción posterior con arabinosa al 0,2 % o 1 mM IPTG 15-30 minutos después y una DO₅₉₅ = 0,4-2,0. Se hicieron crecer las células durante aproximadamente 20 horas después de la inducción antes de la cosecha. Se lisaron las células en TBS con glicerol al 5 %, Triton X-114 al 1 %, o desoxicolato al 2,5 %. A continuación, el aglomerado insoluble se volvió a suspender en TBS + clorhidrato de guanidina 6M pH 8. La proteína se purificó sobre una resina Ni-NTA y se volvió a plegar tanto en la columna como después de la elución a partir de la columna. La etiqueta se eliminó posteriormente con proteasa TEV, y se purificó el producto mediante cromatografía Ni-NTA, intercambio iónico, cromatografía de exclusión por tamaño.

Los datos representativos que muestran la incorporación de PCL a FGF21 se muestran en la Figura 20. Los resultados preliminares de la expresión se registran en la Tabla 4. Las proteínas FGF21 de longitud completa y trunca se purificaron gracias a que la construcción presenta una etiqueta His en el extremo N. La medida en la que se utiliza el codón TAG como codón de detención (dando como resultado la proteína trunca) y se utiliza para incorporar PCL (dando como resultado la proteína de longitud completa) fue dependiente del sitio de incorporación diferente (Figura 20). Para tres de los mutantes (Tabla 4) no se pudo detectar la proteína FGF21, y para los restantes 17 mutantes, los rendimientos totales de la proteína obtenidos variaron entre 5,7 y 143 mg/l con una estimación para los rendimientos de las proteínas de longitud completa deseadas entre 4 y 56 mg/l. Para todos los mutantes, se verificó la incorporación de PCL mediante espectrometría de masas.

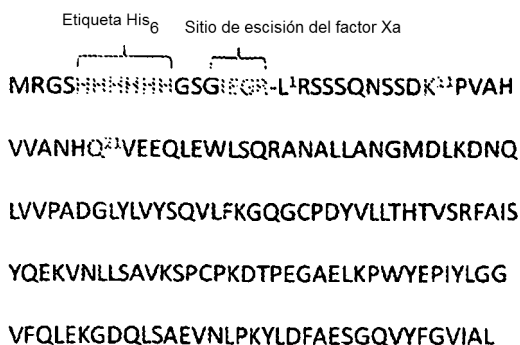
Tabla 4

Mutante	Rendimiento de la proteína total [mg/l]	¿MS Verificada?	% de truncamiento	Rendimiento estimado de la proteína de longitud completa [mg/l]
R64	9	sí	15	7,7
L181	48	sí	50	24,0
S35	9	sí	30	6,3
Q39	5,7	no	30	no det.
R47	9	sí	30	6,3
Q56	22	sí	10	19,8
D74	16	sí	15	13,6
R100	31	sí	75	7,6
R105	10	sí	60	4,0
H115	16	sí	50	8,0
R124	25	sí	40	15,0
E129	15	no	80	no det.
R154	52	sí	50	26,0
L167	143	no	90	no det.

L170	47	sí	60	18,8
Q184	75	sí	50	37,5
K84	75	sí	30	52,5
K87	21	sí	75	5,1
K97	60	sí	10	54,0
K150	113	sí	50	56,3

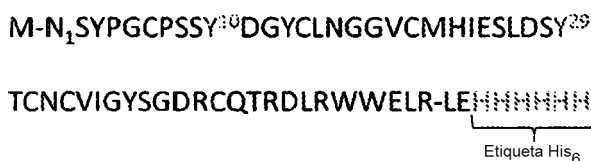
Ejemplo 11: Incorporación de PCL a mTNF-a.

5 Para expresar el mutante mTNF-a Gln21PCL, se transformaron simultáneamente células de E. coli BL21(DE3) con pAra-pylSTBCD y el gen mutante respectivo mTNF-a en un vector plásmido pET22b. Se hicieron crecer las células transformadas en presencia de 5 mM de D-ornitina en medio TB a 37°C y se indujeron con 1 mM de IPTG y arabinosa al 0,2 % (p/v) cuando la DO₆₀₀ alcanzó 0,5. A continuación las células se agitaron de forma continua a 30°C durante 12-16 h y se recogieron. El aglomerado celular se almacenó a -20°C hasta el uso. La estructura cristalina de rayos X de mTNF-a mostró los sitios de incorporación de PCL Lys 11 y Gln21 como se indica a continuación en la secuencia de proteína de mTNF-a recombinante que contiene una etiqueta His₆ en el extremo N seguido por un sitio de escisión del (GIEGR) (SEC ID N°:26):



15 **Ejemplo 12:** Incorporación de PCL a mEGF.

20 Para expresar el mutante mEGF-Tyr10PCL, se transformaron simultáneamente células de E. coli BL21(DE3) con pAra-pylSTBCD y el mutante respectivo mEGF en un vector plásmido pET22b. Se hicieron crecer las células transformadas en presencia de 5 mM de D-ornitina en medio TB a 37°C y se indujeron con 1 mM de IPTG y arabinosa al 0,2 % (p/v) cuando la DO₆₀₀ alcanzó 0,5. A continuación las células se agitaron de forma continua a 30°C durante 12-16 h y se recogieron. El aglomerado celular se almacenó a -20°C hasta el uso. La estructura cristalina de rayos X de mEGF mostró los sitios de incorporación de PCL Tyr10 y Tyr29 como se indicó a continuación se proporciona en la secuencia de proteína de mEGF con una etiqueta His₆ en el extremo C seguido por un sitio de escisión del (SEC ID N°:27):



25

Ejemplo 13: Incorporación de pirrolisina (Pyl) o PCL en mTNF-a.

30 Para insertar Pyl en mTNF-a, se transformaron células de E. coli BL21(DE3) con el gen mTNF-a mutante en un vector plásmido pET22b con el codón Gln21 (CAA) mutado para un codón de detención (TA g) así como pAra-pylSTBCD que contiene pylS de M. mazei, pylT, pylB, pylC, y pylD. Para incorporar exclusivamente PCL en mTNF-a pAra-pylSTBCD mutante se sustituyó con with pAra-pylSTCD, que carece del gen de la presunta metiltransferasa PylB. Se hicieron crecer las células transformadas en presencia de 5 mM de D-ornitina en medio TB a 37°C y se indujeron con 1 mM de IPTG y arabinosa al 0,2 % (p/v) cuando la DO₆₀₀ alcanzó 0,5. A continuación las células se agitaron de forma continua a 30°C durante 12-16 horas y se recogieron. El aglomerado celular se almacenó a -20°C. Tras descongelar el aglomerado celular durante 15 minutos en hielo, las células se volvieron a suspender en tampón de lisis (Tris/HCl 20 mM. NaCl 50 mM, pH 8,0) a 3 ml por gramo de peso en húmedo. Se añadió lisozima a 1 mg/ml y las células se sonicaron durante 2 minutos en hielo. Se centrifugó el lisado a 30.000 x g durante 20 minutos a 4°C para aglomerar el desecho celular. Se añadió 1 ml de suspensión de Ni-NTA al 50 % (Qiagen) al lisado aclarado y se mezcló suavemente agitando a 4 °C durante 60 minutos. La mezcla de lisado-Ni-NTA se introdujo en una columna y se recogió el flujo a su

40

través. Tras lavar la resina con 20 ml de imidazol 25 mM en PCS (pH 8,0), se eluyó la proteína con 2,5 ml de imidazol 250 mM en PBS (pH 8,0) y se intercambió el tampón en PBS, pH 7.4 utilizando una columna PD-10 (GE Healthcare). La expresión de mTNF-a Gln21TAG en presencia o ausencia de pylB así como en presencia o ausencia de D-ornitina se analizó mediante electroforesis en gel de SDS-poliacrilamida (SDS-PAGE) (Figure 21A). En la Figura 21A la banda 1 es el peso molecular normalizado de SeeBlue Plus2 Pre-Stained Standard; La banda 2 es la expresión en presencia de pylB y D-ornitina; La banda 3 es la expresión en presencia de pylB sin D-ornitina añadida; La banda 4 es la expresión en ausencia de pylB con 5mM de D-ornitina añadida; y la banda 5 es la expresión en presencia de pylB y D-ornitina. Los datos ilustran que la expresión de mTNF-a de longitud completa depende de la presencia de D-ornitina y es similar en ausencia o presencia del gen pylB.

Ejemplo 14: Incorporación de pirrolisina (Pyl) o PCL en mEGF

Células de E. coli BL21(DE3) se transformaron simultáneamente con pAra-pylSTBCD y los genes mEGFTyr10TAG y Tyr29TAG en un vector plásmido pET22b. Ambas construcciones se expresaron en presencia de 5 mM de D-ornitina en medio TB a 37°C con la adición de 1 mM de IPTG y arabinosa al 0,2 % (p/v) cuando la DO₆₀₀ alcanzó 0,5. A continuación, la temperatura se redujo a 30°C, y las células se recogieron 16 horas después de la inducción. El aglomerado celular se volvió a suspender en 20 ml de Tris/HCl 20 mM (pH 8,5) y se sonicó durante 5 minutos. Tras centrifugación a 30.000 x g durante 20 minutos, se descartó el sobrenadante. El aglomerado se volvió a suspender mediante sonicación en 20 ml de Tris/HCl 20 mM (pH 8,5) que contenía Triton-X100 al 2 % (v/v). Tras otro ciclo de centrifugación a 30.000 x g durante 20 minutos, el aglomerado se solubilizó mediante sonicación en 10 ml de urea 8 M, Tris/HCl 20 mM, p-mercaptoetanol 10 mM, pH 8,5. los residuos celulares insolubles se eliminaron mediante centrifugación (30.000 x g durante 20 minutos) y el sobrenadante se diluyó 2 veces con tampón de repliegue (Tris/HCl 100 mM, glutatión reducido 4 mM, glutatión oxidado 0,4 mM, etanol al 20 % (v/v), pH 8,5). La muestra diluida se dializó a continuación con tampón de repliegue durante la noche a 4°C utilizando un casete de diálisis *slide-alyzer* (corte de peso molecular 3500 Da, Pierce). La proteína insoluble se eliminó mediante centrifugación a 30.000 x g durante 20 minutos, y el sobrenadante se suplementó con p-mercaptoetanol hasta una concentración final de 2 mM. Se añadió 1 ml de suspensión de Ni-NTA al 50 % (Qiagen) a la proteína replegada y se mezcló suavemente agitando a 4 °C durante 60 minutos. La mezcla de proteína-Ni-NTA se introdujo en una columna y se recogió el flujo a su través. Tras lavar la resina con 20 ml de imidazol 25 mM en PCS (pH 8,0), se eluyó la proteína con 2,5 ml de imidazol 250 mM en PBS (pH 8,0). Finalmente, se intercambió el tampón de la proteína en PBS (pH 7,4) utilizando una columna PD-10 (GE Healthcare). Se investigó la pureza de la preparación de proteína mediante SDS-PAGE (Figura 21B). En la Figura 21B, La banda 1 es el peso molecular normalizado de SeeBlue Plus2 Pre-Stained Standard y la banda 2 es la proteína mutante mEGF Tyr10TAG tras la purificación de Ni-NTA.

Además, se obtuvieron los espectros de ESI-MS de mEGF Tyr10TAG para la proteína en presencia o ausencia de la expresión de PylB, que se obtuvieron usando los métodos descritos anteriormente para mTNFa. (Figura 21C). El espectro de masas inferior en la Figura 21C ilustra que la incorporación de Pyl se produce predominantemente en presencia del gen pylB (masa esperada de mEGF Tyr10Pyl = 7310 Da), mientras que el espectro de masas superior en la Figura 21C ilustra que la incorporación de Pyl se produce solo en ausencia de pylB (masa esperada de mEGF Tyr10Pyl = 7296 Da). De esta manera, la proteína observada en la Figura 21B, banda 2 es mEGF con Pyl incorporada tal como se confirmó adicionalmente mediante el análisis LC-MS/MS. De igual forma, la proteína observada en la Figura 21A, banda 2 es probablemente mTNFa con Pyl incorporada, mientras que la banda 4 es probablemente mTNFa con Pyl incorporada.

Adicionalmente, la cuantificación de la relación PCUPYL según el análisis de masas de los cromatogramas iónicos extraídos de LC-MS/MS de la incorporación de PCL y Pyl en muestras de mEGF Tyr10TAG que expresan todos los genes (pylS, pylT, pylB, pylC, y pylD de M. mazei) y mostraron que Pyl es 5 a 10 veces más abundante que PCL, mientras que la cuantificación de la relación PCL/Pyl procedente del análisis de masas de los cromatogramas iónicos extraídos de LC-MS/MS de la incorporación de PCL y Pyl en muestras de mTNF Gln21TAG que expresan todos los genes (pylS, pylT, pylB, pylC, y pylD) mostró que PCL es aproximadamente 7 veces más abundante que Pyl. La cuantificación en ausencia del gen PylB muestra solo la proteína PCL, lo que ilustra que la incorporación de Pyl es estrictamente dependiente de pylB, lo que sugiere además que PylB es sin duda la metiltransferasa requerida para la biosíntesis de Pyl.

Ejemplo 15: Incorporación de otros análogos y precursores:

Células HK100 (derivadas de Genehogs; Invitrogen) se transfectaron simultáneamente con pAra-pylSTBCD, pAra-pylSTCD, pAra-pylSTC, pAra-pylSTD, o pSUPAR-pylST y pMH4-FASTE-L2222TAG-L2223I. Se hicieron crecer células en cultivos de 25 ml de caldo terrífico (TB) (Sigma) a 37 °C a una DO 595-0,6. Las células se movieron a 30 °C y se añadieron los análogos o los compuestos precursores a cada cultivo individual a la concentración indicada en la Figura 15. Los compuestos evaluados fueron N-s-ciclopentiloxycarbonil-L-lisina (CYC; Sigma), D-ornitina (Chem-Impex), PCL-A (véase el Ejemplo 36-1, Compuesto 3647-125), PCL-B (véase el Ejemplo 36-2, Compuesto 3793-011), P2C (véase el Ejemplo 36-2, Compuesto 3647-164), P5C (véase el Ejemplo 35-1, compuesto 3793-007) y Lys-Ns-D-ornitina (véase el Ejemplo 35-2, Compuesto 3793-031). A continuación se indujeron las células 20 minutos después con arabinosa al 0,2 %. Tras 18-20 horas, se recogieron las células mediante centrifugación. Se lisaron las células mediante sonicación, se purificaron en Ni-NTA (Qiagen) en condiciones naturales y se evaluaron en un gel

SDS-PAGE seguido por la tinción de Coomasie (0,25 ml de Ni-NTA, elución de 0,75 ml, 20 ml en gel). La Figura 15A muestra la proteína purificada (FAS- TE) a partir de células que crecen con los genes *pyl* indicados presentes (*pylS/pylT* o *pylB/pylC/pylD/pylS/pylT*) y alimentan los compuestos indicados. Se utilizaron Cyc y D-ornitina (D-Orn) como controles positivos. La no adición de compuestos se presenta como control negativo. Los volúmenes de las muestras de gel en las bandas 1-11 y en 12-18 son cada uno internamente consistentes.

Las bandas 2 and 3 muestran la proteína obtenida a partir de células que crecen solo con los genes *pylS* y *pylT* y tanto con PCL- B (Lys-P2C) o PCL-A (Lys- P5C), respectivamente. Ambos compuestos se incorporaron a las proteínas, demostrando por tanto la capacidad de *PylS* de utilizar los dos compuestos. Las bandas 5 a 11 muestran la evaluación de la síntesis de proteínas en células con el conjunto completo de los genes *pyl* biosintéticos (*pylB/pylC/pylD/pylS/pylT*) en presencia del precursor Lys-Ns-D-Orn (banda 10), o precursores que tenían solo un anillo de pirrolina: P2C (banda 6) y P5C (banda 8). Los precursores que tenían solo un anillo de pirrolina: P2C (banda 6) y P5C (banda 8) no fueron suficientes para soportar la biosíntesis de PCL, los que sugirió que estos no eran intermedios en la ruta. Sin embargo, Lys-Ns-D-Orn (banda 10) dio como resultado la expresión de la proteína con PCL incorporada en el codón ámbar, demostrando que esta es un intermedio en la ruta biosintética de PCL. La Tabla 5 proporciona la cantidad de proteína obtenida (Bradford), la masa observada y la relación de PCL:PYL:CYC obtenida.

Tabla 5

Banda	Prueba	Bradford (mg/25 ml)	Masa LC-MS intacta (observada)	Relación PCL:PYL:CYC a partir de LC-MS
1	negativo	0,071	-	-
2	PCL-B (Lys-P2C)	2,703	33270,8	PCL solo
3	PCL-A (Lys-P5C)	0,539	33256,8	PCL solo
4	positivo	0,796	33272,8	CYC solo
5	negativo	0,131	33275,2	2:1:-
6	P2C	0,103	33274,0	4:1:-
7	positivo	1,838	33257,2, 33368,4	PCL solo
8	P25	0,075	Ruido (33248)	
9	positivo	0,795	33248 (PCL), 33360	
10	Lys-Ns-D-ornitina	1,66	33248 (PCL)	
11	positivo	0,51	33248 (PCL)	

Para determinar si Lys-Ns-D-Orn era un intermedio en la ruta biosintética en la dirección 5' de cualquiera de los genes requeridos (*pylC* y *pylD*), Células HK100 (derivadas de Genehogs; Invitrogen) se transfectaron simultáneamente con pSUPAR-*pylST*. pMH4-FASTE-L2222TAG-L2223I, y cualquiera de pAra-*pylSTB*, pAra-*pylSTC*, pAra-*pylSTD*, pAra-*pylSTCD* o pAra-*pylST*- BCD. La Figura 15B muestra que solo se requiere *pylD* (banda 15) para la formación de PCL a partir de Lys-Ns-D-Orn, sugiriendo que este es un intermedio en la dirección 5' de *pylC* en la ruta biosintética. Además, la evaluación de RMN confirmó que Lys-Ne-D-Orn es un sustrato para *PylD*.

Ejemplo 16: Derivatización de PCL incorporada en hRBP4 con 2-amino benzaldehído, 2-aminoacetofenona y 2-amino-5-nitro-benzofenona.

Se proporciona en este ejemplo un marcado de PCL con 2-amino-benzaldehído (2-ABA), 2-aminoacetofenona y 2-amino-5-nitro-benzofenona. Se cree que la reacción sigue el esquema general que se muestra en la Figura 22. Se usaron la espectrometría de masas y la RMN para evaluar las estructuras formadas. La Figura 23 indica los conjugados de proteína formados a partir de tres restos diferentes de 2-amino-benzaldehído y presenta los cambios de masa esperados y observados.

Detección espectrométrica de masas del sitio de la PCL específicamente modificado por 2-ABA en hRBP4

Se expresó la proteína de unión a retinol (hRBP4) en células HEK293F tal como se describe en el Ejemplo 2. Se incorporó la PCL en vez de Phe122 (mutante nº 6 de hRBP4) tal como se verificó mediante espectrometría de masas (Figura 6A-11). 10 ml de disolución madre de hRBP4 Phe122PCL se mezclaron con 89 ml de tampón de acetato de sodio 200 mM, pH 5,0 y 1 ml de una disolución 1 M de 2-ABA y se incubó durante 16 horas a temperatura ambiente. Las concentraciones finales en la mezcla de reacción fueron 17 mM de la proteína hRBP4 Phe122PCL y 10 mM de 2-amino-benzaldehído (2-ABA). En la Figura 24 se muestra el espectro de masas de la hRBP4 derivatizada, donde la masa obtenida correspondió al aducto de 2-ABA de PCL (23269,2 Da). La hRBP4 sin modificar tiene una masa de 23166,8 Da y por tanto la masa observada aumenta a 102,4 Da (esperada +103 Da) lo que demuestra que hRBP4 ha sido modificada con 2-ABA. Al menos un 96 % de las intensidades de los picos en el espectro de masas es debida al aducto de 2-ABA de PCL, lo que indica que la reacción llegó casi hasta su finalización.

Se llevó a cabo el análisis LC-MS de la digestión triptica de la proteína hRBP4 Phe122PCL derivatizada con 2-ABA y el análisis MS/MS (Figura 25A) identificó el péptido YWGVASF*LQK esperado (SEC ID N°:17), donde F* tenía una masa consistente con la de la PCL modificada con 2-ABA tal como se muestra en la Figura 23. La reacción se completó de tal manera que el resto de PCL sin derivatizar no fue detectable. Se proporciona a continuación el espectro MS/MS asignado de YWGVASF*LQK (F* = aducto de PCL-2-ABA):

Masa monoisotópica del péptido neutro Mr (calc): 1376.7241 Modificaciones de la variable:

F7 : aducto de PCL-2-ABA a F (F)

Puntuación de iones: 4 4 Esperado: 0,059

Correspondencias (negrita): 13/74 fragmentos de iones usando los 28 picos más intensos

Nº	b	b ⁺	b ⁺	b ⁺⁺	b ⁺	b ⁺⁺	Sec.	y	y ⁺	y ⁺	y ⁺⁺	y ⁺	y ⁺⁺	Nº
1	164,0706	82,5389												10
2	350,1499	175,5786					W	1214,6681	607,8377	1197,6415	599,3244	1196,6575	598,8324	9
3	407,1714	204,0893					G	1028,5887	514,7980	1011,5622	506,2847	1010,5782	505,7927	8
4	506,2398	253,6235					V	971,5673	486,2873	954,5407	477,7740	953,5567	477,2820	7
5	577,2769	289,1421					A	872,4989	436,7531	855,4723	428,2398	854,4883	427,7478	6
6	664,3089	332,6581			646,2984	323,6528	S	801,4618	401,2345	784,4352	392,7212	783,4512	392,2292	5
7	990,4832	495,7452			972,4727	486,7400	F	714,4297	357,7185	697,4032	349,2052			4
8	1103,5673	552,2873			1085,5567	543,2820	L	388,2554	194,6314	371,2289	186,1181			3
9	1231,6259	616,3166	1214,5993	607,8033	1213,6153	607,3113	Q	275,1714	138,0893	258,1448	129,5761			2
10							K	147,1128	74,0600	130,0863	65,5468			1

La Figura 25B es el TIC y EIC de iones 2+ de YWGVASF*LQK (F* = PCL y aducto de PCL-2-ABA), donde la comparación de los EIC (cromatograma de iones extraído) para las especies derivatizada y sin derivatizar (no detectable) indica el agotamiento de la reacción. La Figura 25C es el análisis por espectrometría de masas de hRBP4 Phe122PCL derivatizado con 2-ABA que muestra los precursores 3+ y 2+ de YWGVASF*LQK a m/z 459,92 (3+) y 689,37 (2+) respectivamente. (F* = aducto de PCL-2-ABA). este ejemplo demuestra que las reacciones observadas con 2-ABA se producen específicamente en el sitio con el resto de PCL incorporado en el sitio TAG deseado en el resto 122.

Eficacia de reacción como una función del pH:

se evaluó la eficacia de reacción como una función del pH, (Figura 26), haciendo reaccionar 17 mM de la proteína mutante hRBP4 PCL con o sin 10 mM de 2-amino-benzaldehído (2-ABA) en 200 mM de tampón de acetato de sodio, pH 5,0, o con 10 mM de 2-amino-benzaldehído (2-ABA) 10x tampón PBS, pH 7,4 durante 12 horas a temperatura ambiente. Las mezclas de los espectros de masas de la reacción mostraron que al menos un 87 % de la intensidad total del pico corresponde a la de la proteína modificada con un resto de 2-ABA (+102 Da tal como se esperaba) (Figura 26A, B y C). La mezcla de reacción a pH 7,4 contiene aproximadamente un 13 % de proteína sin reaccionar (Figura 26C) mientras que solo se detectó un 4,2 % de proteína sin reaccionar en la reacción a pH 5,0 (Figura 26B) sugiriendo una reactividad ligeramente mayor de la PCL a pH 5,0 en comparación con el pH 7,4. La reacción es específica de la presencia de PCL, de tal manera que hRBP4 con OMePhe incorporada en vez de Phe62 no se modificó mediante la presencia de 10 mM 2-ABA (Figura 26D, E y F).

Eficacia de reacción como una función de la concentración de reactivo a proteína y con otros restos de tipo 2-ABA:

La eficacia de reacción como una función de la relación de concentración de reactivo a proteína, y la reactividad con reactivos de tipo 2- ABA se evaluó haciendo reaccionar de proteína hRBP4 PCL mutante 17 mM con 2-amino-benzaldehído (2-ABA) 0,1 mM, 2-amino-acetofenona (2-AAP) 0,1 mM o 2-amino-5-nitro-benzofenona (2-ANBP) 0,1 mM en tampón de acetato de sodio 200 mM a pH 5,0. Después de 12 horas a temperatura ambiente, el espectro de masas de la mezcla de reacción mostró la masa esperada de la proteína conjugada (Figura 27). Para 2-ABA, la intensidad relativa del pico conjugado correctamente fue del 88 % de la intensidad total; solo el 4,2 % de la proteína permaneció sin reaccionar (Figura 27A). El 93 % de 2-AAP pareció reaccionar; el 4,5 % no reaccionó (Figura 27B). Para 2-ANBP solo reaccionó un 5,4 % (Figura 27C) debido presumiblemente a la baja solubilidad del reactivo, a medida que 2-ANBP precipita inmediatamente tras la adición de la disolución que contiene la proteína mutante hRBP4 PCL.

Este ejemplo demuestra que las proteínas modificadas con PCL reaccionan con diferentes análogos de 2-amino-benzaldehído y se derivatizan en el sitio de la PCL incorporada. En todos los casos, la masa medida de la proteína modificada fue consistente con la masa esperada para las estructuras que se representan en la Figura 23. Los datos demuestran también que las reacciones de conjugación proceden con una elevada eficacia y cercana al agotamiento a una relación baja de reactivo a proteína de solo 6 a 1. Los datos demuestran también que cada muestra de proteína se derivatiza solo una vez. Sin embargo, se observó que la unión de múltiples moléculas de reactivo se obtuvo (Figura 28A) a relaciones muy altas de reactivo a proteína (4700 veces) y a un pH de 7,5. La precipitación de una muestra idéntica a pH 5,0 sugiere también múltiples reacciones. de forma similar hRBP4 modificada con OMePhe en vez de Phe62 (6,5 mM) mostró un modelo similar de conjugados (Figura 28B) al observado en la Figura 28A cuando se hace reaccionar con un exceso molar de 15400 veces de 2-ABA a pH 7,5. Esto demuestra que bajo estas condiciones (exceso molar grande de reactivo sobre proteína), la unión no depende de la presencia de una cadena lateral de PCL sino que implica probablemente cadenas laterales de lisina. Sin embargo, Las reacciones de ejemplo ilustran que los restos de PCL incorporados en las proteínas se pueden derivatizar de forma específica y casi

cuantitativa haciéndose reaccionar con 2-amino-benzaldehído que contiene moléculas añadidas en un exceso molar pequeño.

Ejemplo 17: Derivatización de PCL incorporada en FAS-TE con 2-amino-acetofenona

16 mM de FAS-TE Tyr2454PCL producida en *Escherichia coli* como en el Ejemplo 8 se hicieron reaccionar con 2-AAP 1 mM durante 16 horas a temperatura ambiente y 24 horas a 4 grados °C a pH 5,0 en 200 mM de tampón de acetato de sodio. La Figura 29A muestra el espectro de masas de FAS-TE Tyr2454PCL sin reaccionar mientras que la Figura 29B muestra el espectro de masas de la mezcla de reacción: se produce un 100 % de la intensidad del pico observable a 33318,8 Da, 116,8 Da más grande que el material sin reaccionar y comprendiendo errores de aumento de masa de 116 Da esperados para FAS-TE Tyr2454PCL modificados con 2-AAP. De igual forma, a pH 7,4 la reacción avanza hasta un agotamiento del 95 % (Figura 29C (sin reaccionar) y Figura 29D (con reacción)).

Ejemplo 18: Derivatización de PCL incorporada en hRBP4 con 2-amino-acetofenona- PEG8

Este ejemplo demuestra que PCL incorporada en hRBP4 se puede derivatizar hasta finalización en un sitio único con un derivado de polietilenglicol (PE g) de 2-amino-acetofenona. En este ejemplo, el PEG contiene 8 unidades de etilenglicol y su estructura se muestra en la Figura 31. El ejemplo demuestra también que hRBP4 natural a pH 5,0 y pH 7,5 no reacciona con 2-AAP-PEG8 con relaciones de reactivo a proteína de hasta 2300.

se preparó hRBP4 mutante nº 6 como en el Ejemplo 2. La 2-amino-acetofenona PEG8 (TU3205-044) se preparó tal como se describe en el Ejemplo 20. Se dejó que las reacciones continuaran durante 14 horas a temperatura ambiente y 72 horas a 4 grados °C antes de que se obtuviera el espectro de masas de las mezclas de reacción.

Para las reacciones llevadas a cabo a pH 7,5, se diluyeron 10 ml de hRBP4 Phe122PCL (0,22 mg/ml, en PBS, pH 7,5) con 10 ml de 10x PBS. Se añadieron 0,2 ml o 2 ml de una disolución madre 100 mM de 2-AAP-PEG8 (en agua) a una concentración final de 1 y 9,1 mM respectivamente. La concentración de proteína fue de 4,7 mM y 4,3 mM, respectivamente, dando como resultado un exceso molar de 210 o 2100 de reactivo a proteína. La Figura 32A muestra el espectro de masas de la mezcla de reacción a la relación 210 a 1, indicando un agotamiento aproximado del 95 % de la reacción y el aumento de masa observado de 556 Da fue idéntico al valor esperado (Figura 31). se obtuvo un agotamiento del 100 % para la reacción a un exceso molar de 2100 (no se muestran los datos).

De igual forma, Se llevaron a cabo las reacciones a pH 5,0. En este caso, se diluyeron 10 ml de hRBP4 Phe122PCL (0,22 mg/ml, en PBS, pH 7.5) con 90 ml de tampón de acetato de sodio 200 mM a pH 5.0. Se añadieron 1 ml o 10 ml de una disolución madre 100 mM de 2-AAP-PEG8 (en agua) a una concentración final de 1 y 9,1 mM respectivamente. la concentración de proteína fue de 0,94 y 0,86 mM, respectivamente, dando como resultado un exceso molar de 1050 o 10500 de reactivo a proteína. La Figura 32B muestra el espectro de masas de la mezcla de reacción a la relación 1050 a 1. Se obtuvo un agotamiento del 100 % para ambas reacciones y el aumento de masa observado de 556 Da fue idéntico al valor esperado (Figura 31).

Para ensayar la reactividad de 2-AAP-PEG8 con la proteína hRBP4 natural, las reacciones del ensayo se ajustaron de forma similar a las anteriores. Para las reacciones a pH 7.5 la concentración final de proteína natural fue de 20 mM y la concentración de 2-AAP-PEG8 fue de 1 or 9,1 mM resultante en una relación molar de 46 y 460 a 1. Las reacciones a pH 5 se llevaron a cabo a una concentración de proteína 4 mM y 1 y 9,1 mM 2-AAP-PEG (230 y 2300 a 1 de reactivo a proteína). En las cuatro reacciones, solo la proteína hRBP4 natural sin modificar se observó en la masa esperada (Figura 32D-F). Este ejemplo demuestra que la reacción de acoplamiento de 2-AAP-PEG es muy específica para la presencia de un resto de PCL en la proteína diana.

Ejemplo 19: Derivatización de PCL incorporada en FAS-TE con 2-amino-acetofenona-PEG de diferente peso molecular

Este ejemplo demuestra la generalidad de la reactividad de las cadenas laterales de PCL con 2-amino-acetofenona PEG. Este ejemplo demuestra adicionalmente que los 2-AAP-PEG de longitudes suficientes para ser útiles para la modificación de proteínas bioterapéuticas se pueden conjugar con las proteínas modificadas con PCL.

16 mM de FAS-TE Tyr2454PCL producida en *Escherichia coli* como se describe en el Ejemplo 8 se hicieron reaccionar con 1 mM de TU3205-044 (2-AAP-PEG8) 2-AAP durante 16 horas a temperatura ambiente a pH 5,0 en 200 mM de tampón de acetato de sodio. La Figura 33A muestra el espectro de masas de FAS-TE Tyr2454PCL sin reaccionar, mientras que la Figura 33B es el espectro de masas de la mezcla de reacción. En la Figura 33B, se produce un 100 % de la intensidad del pico observable a 33758,4 Da, que es 556 Da más grande que la del material sin reaccionar. Esta diferencia de masa es la esperada para FAS-TE Tyr2454PCL modificada con 2-AAP-PEG8 (Figura 31).

En un ejemplo adicional, PCL incorporada en FAS-TE a Tyr2454 se derivatizó con tres diferentes 2-AAP- PEG con PM de 0,5kDa, 2.4 kDa y 23 kDa. Se produjo el mutante PCL en el sitio único FAS-TE en *Escherichia coli* a un rendimiento de aproximadamente 160 mg/l, (que corresponde a un rendimiento en peso de -80 %) tal como se describe en el Ejemplo 8. Se hicieron reaccionar alícuotas de FAS- TE Tyr2454PCL (0,16 mM en PBS, pH 7,5) con TU3205-044 (0,5

kDa 2-AAP-PE), TU3205-048 (2-AAP-PEG de 2,4 kDa) y TU3205-052 (2-AAP-PEG de 23 kDa) a relaciones molares de 10:1 1 o 100:1 1 durante 6 días a 4°C temperatura ambiente. Las estructuras de los PEG y sus síntesis se describen en el Ejemplo 37.

5 Antes del análisis de SDS-PAGE (Figura 34) se eliminaron los reactivos de PEG en exceso uniendo las proteínas FAS-TE etiquetadas con His a las perlas de Ni-NTA y se repitió el lavado de las perlas con tampón PBS. Para el análisis espectrométrico del gel y de la masa, se eliminó el tampón en exceso procedente de las perlas Ni-NTA y se eluyó la proteína con tampón imidazol. Específicamente, se añadieron 50 ml de perlas de Ni-NTA a 50 ml de la mezcla de reacción, se incubaron durante 2 horas y la mezcla de reacción y el tampón se separaron mediante centrifugación. 10 A continuación se lavaron las perlas 3 veces con 1 ml de PBS; 70 ml de 250 mM de tampón imidazol, se añadió Tris 20 mM, a pH 8 a las perlas para eluir la proteína para el análisis espectrométrico de gel y masa. La PEGilación con 2-AAP-PEG de 0,5 kDa no se pudo resolver mediante SDS-PAGE sino que se verificó mediante espectrometría de masas. La extensión de la reacción fue aproximadamente del 57 % para la reacción 100:1 1 a temperatura ambiente y del 43 % para la reacción 100:1 a 4°C (no se muestran los datos). Para los 2-AAP-PEG más grandes, los productos PEGilados se pueden resolver mediante SDS-PAGE (Figura 34). Todas las reacciones fueron incompletas y 15 continuaron hasta aproximadamente un 25-30 % para 2-AAP-PEG de 2,4 kDa y 23 kDa. La conjugación de sitio único en el sitio de la PCL y la extensión de la reacción se confirmaron mediante espectrometría de masas para 2-AAP-PEG de 2,4 kDa (Figura 33). Los datos de espectrometría de masas para la proteína derivatizada con 2-AAP-PEG de 23 kDa no se pudieron obtener debido a la falta de homogeneidad del PEG.

20 Los rendimientos de reacción notificados son una estimación inferior debido a que la rigurosidad de la extracción puede haber favorecido la extracción de FAS-TE sin reaccionar ya que la PEGilación puede haber disminuido la afinidad de la etiqueta His (como fue el caso de la PEGilación de FAS-TE Leu2222PCL; no se muestran los datos). Sin embargo, la detección del material PEGilado tras la extracción demuestra la estabilidad del enlace PCL-2-AAP-PEG.

25 Ejemplo 20: Derivatización de PCL incorporada en FGF21 con 2-amino-acetofenona-PEG de diferentes pesos moleculares

30 El ejemplo muestra que FGF21 con PCL incorporada en diversas posiciones se puede PEGilar con reactivos 2-AAP-PEG. El ejemplo muestra también que los mutantes FGF21 PEGilados se pueden separar de FGF21 de longitud completa sin reaccionar y FGF21 truncada mediante una combinación de cromatografía de intercambio iónico y exclusión por tamaño.

35 FGF21 con PCL incorporada en la posición de la Lisina 84 se expresó en Escherichia coli y se replegó y purificó tal como se describe en el Ejemplo 10 excepto que la proteína no se sometió a escisión de la proteasa TEV. El depósito de proteína fue de aproximadamente 6,6 mg/ml en PBS y contenía aproximadamente un 20 % de proteína FGF21 truncada en el resto 84. 5 ml del depósito FGF21 se mezclaron con 45 ml de tampón acetato de sodio 200 mM a pH 5.0 y 0.5 ml de una disolución madre de TU3205-044 100 mM (2-AAP-PEG8, véase el Ejemplo 20). La relación molar final fue de 1 mM de 2-AAP-PEG8 para aproximadamente 30 mM de FGF21. Se dejó que la reacción continuara durante 16 40 horas a temperatura ambiente y que se agotara.

45 La Figura 35A muestra el espectro de masas de FGF21 sin reaccionar, mientras que la Figura 35B es el espectro de masas de la mezcla de reacción. la reacción de PEGilación continuó hasta el agotamiento y dio como resultado una proteína con 21792,4 Da, que es 556 Da más grande que la del material sin reaccionar. Se esperaba esta diferencia de masas para FGF21 modificada con 2-AAP-PEG8.

50 Diversos mutantes FGF21 PCL (expresados y purificados como en el Ejemplo 10) se PEGilaron con 2-AAP- PEG 23 kDa (TU3205-052, véase el Ejemplo 37-3). Usualmente, las concentraciones de proteína fueron entre 100 y 400 mM a la vez que se añadía 2-AAP-PEG de 23 kDa hasta una concentración final de 1 mM en tampón PBS, pH 7,4. Las reacciones se incubaron a 4°C durante 3 días. Se muestran los SDS-PAGE de las reacciones de PEGilación con siete mutantes FGF21 PCL representativos en la Figura 36A, a la vez que se muestran ocho proteínas FGF21 PEGiladas purificadas separadas de FGF21 sin reaccionar de longitud completa (FL) y FGF21 truncada (TR) en la Figura 36B.

55 Ejemplo 21: Derivatización de PCL incorporada en EPO con 2-amino-acetofenona- PEG

60 Se incorporó PCL en diversas posiciones en EPO de ratón tal como se describe en el Ejemplo 6. Tras la purificación con Ni-NTA, se purificaron adicionalmente con construcciones mEPO (construcciones mutantes nº 6, nº 9 y nº 10) se purificaron adicionalmente mediante una columna de filtración en gel S-300 en PBS. Las proteínas mEPO purificadas se concentraron hasta aproximadamente 1 mg/ml. se añadió 2-AAP-PEG 23kDa activada (TU3205-052, véase el Ejemplo 37-3) a las proteínas purificadas en 1 mM y se incubó en las condiciones que se indican en la tabla 6.

Tabla 6

nº	construcciones	pH	temp
1	EPO6	pH 7,5	4C
2	EPO6	pH 5,5	4C

3	EPO6	pH 8,5	4C
4	EPO9	pH 7,5	4C
5	EPO9	pH 5,5	4C
6	EPO9	pH 8,5	4C
7	EPO10	pH 7,5	4C
8	EPO10	pH 5,5	4C
9	EPO10	pH 8,5	4C
10	EPO6	pH 7,5	4C
11	EPO6	pH 7,5	22C
12	EPO6	pH 7,5	30C
13	EPO6	pH 7,5	37C

5 Todas las construcciones mEPO se hicieron avanzar en forma de monómero por la columna. Cuando 2-AAP-PEG de 23 kDa se incubó con las proteínas purificadas, EPO se PEGilaron en un sitio único, ya que migraron como una banda de 65 kDa en SDS-PAGE (Figura 37). La eficacia de la reacción varió desde un 10 % a un 15 % dependiendo de las condiciones, resultando un pH de 5,5 con un mayor grado de PEGilación con respecto a pH 7 y pH 8,5. Además, temperaturas mayores de 4° C no aumentan significativamente la extensión de la PEGilación.

Ejemplo 22: Derivatización de PCL incorporada en hRBP y FAS-TE con D-manosamina.

10 Este ejemplo demuestra el acoplamiento directo de un aminoazúcar, D-manosamina, a PCL incorporada en dos proteínas dianas. Este ejemplo sugiere un esquema de reacción general para la glicosilación de PCL que contiene proteínas (Figura 38).

15 Se preparó RBP4 Phe122PCL humana (mutante nº 6) en células HEK293F como en el Ejemplo 2. 10 ml de 170 mM de disolución madre se añadieron a 89 ml de tampón 10x PBS, pH 7,5 y se mezclaron con 1 ml de D-manosamina 1 M. Se incubó la proteína hRBP4 Phe122PCL (17 mM) con 10 mM de D-manosamina durante 14 a temperatura ambiente (no se observó reacción) seguido por 48 horas a 37°C. La Figura 39 muestra el espectro de masas de la mezcla de reacción después de un periodo de incubación de 37°C. Además, el pico esperado de hRBP4 a 23165.6 Da (esperado 23166 Da), se observó un pico presuntamente asignado al aducto de D-manosamina a 23300,0 Da. El aumento de la masa de 164,4 Da está cercano, pero no es idéntico al aumento de 161,1 Da esperado para el presunto producto de reacción que se representa en la Figura 38. Además, parte de la proteína degradada a especie se detectó a 21007,8 Da.

25 En un segundo ejemplo, FAS-TE modificado con PCL en la posición 2222 (11 mM) se expresó en Escherichia coli y se purificó como en el Ejemplo 8, se incubó con D-manosamina 10 mM durante 72 horas a temperatura ambiente. En la Figura 40A se muestra el espectro de masas de la muestra sin reaccionar y contiene la señal esperada a 33250,4 Da que corresponde a la proteína sin reaccionar y un pico a 33360,4 Da atribuido a la proteína contaminante en la muestra. La espectrometría de masas de la mezcla de reacción muestra la señal esperada a 33250,4 Da que corresponde a la proteína sin reaccionar (Figura 40B). Es visible un pico adicional a 33408,8 Da, 158,4 Da más grande que el material de partida, presumiblemente que el del producto de reacción que se muestra en la Figura 37. La comparación de estas dos muestras sugiere una conversión de FAS-TE Leu2222PCL al aducto de D-manosamina con un rendimiento de aproximadamente 50 % en estas condiciones de reacción.

Ejemplo 23: Reticulación covalente de FGF21 mediada por PCL

35 Este ejemplo demuestra que las proteínas se pueden dimerizar covalentemente mediante un reticulador específico de PCL bifuncional. FGF21 con PCL incorporada en la posición de la Lisina 84 se expresó en Escherichia coli y se replegó y purificó tal como se describe en el Ejemplo 10, excepto que la proteína no se sometió a escisión por la proteasa TEV. FGF21 Lys84PCL se derivatizó con el reticulador TU633-010 (Figura 43A; véase el Ejemplo 37-4 para la síntesis). El depósito de proteína fue de aproximadamente 6,6 mg/ml en PBS y contenía aproximadamente un 20 % de proteína FGF21 truncada en el resto 84. 5 ml del depósito de FGF21 se mezclaron con 45 ml de tampón acetato de sodio 200 mM a pH 5,0. se añadieron 0,1 ml una disolución madre 5 mM de reticulador TU633-010 (en DMSO) hasta una concentración final de 10 mM de reticulador y aproximadamente 30 mM de FGF21. De forma similar 50 ml de disolución madre de FGF21 en PBS, pH 7,4 se hizo reaccionar con 1 ml de 5 mM de TU633-010 a una concentración de 100 mM del reticulador y aproximadamente 300 mM de FGF21. Se preparó una muestra de FGF21 diluida en 200 mM de tampón acetato de sodio como control y se trató de idéntica forma. Después de 16 horas a temperatura ambiente, se analizaron alícuotas de las reacciones y de la muestra del control mediante espectrometría de masas y análisis SDS-PAGE sin purificación. El espectro de masas obtenido para la muestra a pH 5,0 (Figura 43B) muestra de forma clara un pico a la masa esperada del dímero covalente (43037,2 Da; 53 % de intensidad relativa de todos los picos de FGF21), toda la masa esperada de FGF21 con un extremo del reticulador unido (21820,0 Da; 33 %) y FGF21 sin reaccionar (21235,2 Da; 14 %). Para la muestra a pH 7,4, no se detectó el dímero covalente; se hizo reaccionar un

28 % de FGF21 con el reticulador mientras que la mayoría de las proteínas está sin reaccionar (no se muestran los datos).

5 El ajuste del reactivo a la concentración de la proteína aumentó adicionalmente el rendimiento del dímero covalente. Específicamente, 10 ml del depósito de FGF21 se mezclaron con 90 ml de tampón acetato de sodio 200 mM a pH 5,0. se añadieron 0,3 ml una disolución madre 5 mM de reticulador TU633-010 (en DMSO) hasta una concentración final de 15 mM de reticulador y aproximadamente 30 mM de FGF21. Se incubó una muestra durante 4 días a temperatura ambiente mientras que una segunda muestra se incubó a 4°C. Se prepararon también muestras en 10x PBS, pH 7,5 y se incubaron de idéntica forma. Para la muestra a pH 5,0 incubada a temperatura ambiente, el pico del dímero FGF21 covalente en la masa esperada de 43034,4 Da es la especie dominante, mientras que FGF21 sin reaccionar no se detectó a 21233,6 Da. algunos FGF21 modificados con un extremo del reticulador unido se detectan a 21818,4 Da. (Figura 44A). La reacción no progresa en la misma extensión a pH 5,0 y 4°C debido a que aproximadamente un 19 % de FGF21 permanece sin reaccionar; el 40 % está modificado con un extremo del reticulador unido mientras que aproximadamente un 41 % de la intensidad del pico del espectro de masas es el del dímero covalente. Las reacciones a pH 7,5 no dan como resultado ningún dímero covalente tal como se indica mediante SDS-PAGE (Figura 44B).

Modificación específica de sitio de proteínas que contienen pirrolina-carboxi-lisina (PCL) con otras moléculas diversas.

20 En otras realizaciones, el acoplamiento de los conjugados de 2-aminobenzaldehído (ABA) y conjugados de 2-aminoacetofenona (AAP) proporcionado en el presente documento a proteínas que contienen PCL se llevó a cabo en 10 x solución salina tamponada con fosfato (PBS) a pH 7,0 y 25°C. la reacción de conjugación se inició mediante la adición de 10 mM de proteína que contiene PCL y 100 mM de conjugado ABA/AAP. La formación completa del conjugado de proteína se verificó mediante espectrometría de masas con ionización mediante electropulverización (ESI-MS) o de desorción/ionización laser asistida por matriz (MALDI). Se analizó el acoplamiento de los conjugados ABA/AAP ADN mediante el ensayo de cambio de geles utilizando un Gel NuPAGE Bis-Tris al 4-12 % (Invitrogen, Carlsbad, CA). Tras el acoplamiento cuantitativo, el conjugado de proteína se dializó en tampón fosfato de sodio 10 mM (pH 7,5) y se concentró a 100 mM utilizando una Unidad de Filtro de Centrifuga Amicon Ultra-4 con un corte de 10 kDa (Millipore Corporation, Bedford, MA). A continuación se añadió una disolución preparada recientemente de 200 mM de NaCNBH₃ (disueltos en 10 mM de tampón fosfato, pH 7,5), hasta una concentración final de 20 mM. Tras dejar que continuara la reacción de reducción durante 2 - 4 horas a 25°C, la reacción se inactivó rápidamente mediante la adición de seis volúmenes de tampón fosfato de sodio 10 mM (pH 7,5). Utilizando una columna NAP-5 o una columna PD10 (GE Healthcare, Piscataway, NJ), El conjugado de proteína reducido se intercambiò finalmente por el tampón en el tampón deseado. Los ejemplos no limitantes de acoplamiento de dichos conjugados de 2-aminobenzaldehído (ABA) y conjugados de 2-aminoacetofenona (AAP) con diversas proteínas se proporcionan a continuación.

Ejemplo 24: Acoplamiento del reactivo de biotina

40 Para demostrar que la biotina se puede acoplar a proteínas que tienen uno o más restos de PCL incorporados en las anteriores, se conjugó mEGF-Tyr10PCL (véase el Ejemplo 12) con biotina usando un reactivo de ABA-biotina (X3626-140, Ejemplo 40). El acoplamiento de la biotina se llevó a cabo como sigue, 500 mM de ABA-biotina se añadieron a 10 mM de mEGF-Tyr10PCL en solución salina tamponada con fosfato (PBS, pH 7,0) y DMSO al 0,5 % (v/v) y se hicieron reaccionar a 25°C durante 16 horas. Se verificó la formación completa del conjugado de biotina mediante ESI-MS (Figura 47A; masa esperada de la proteína sin acoplar = 7296; masa esperada de la proteína acoplada = 7902).

45 Tras el acoplamiento cuantitativo, una disolución preparada recientemente de 200 mM de NaCNBH₃ (disueltos en PBS, pH 7,0) se añadió hasta una concentración final de 20 mM. Tras dejar que continuara la reacción de reducción durante 3 horas a 25°C, la reacción se inactivó rápidamente mediante la adición de seis volúmenes de PBS (pH 7,0). Se eliminó el NaCNBH₃ excesivo mediante diálisis frente al PBS (pH 7,0) a 4°C utilizando un casete de diálisis Slide-A-Lyzer (corte de pesos moleculares de 3.500 Da, Pierce). El conjugado de biotina reducida se concentró usando una Unidad de Filtro de Centrifuga Amicon Ultra-4 con un corte de 3,5 kDa (Millipore Corporation). Tras la electroforesis a través de un gel NuPAGE Bis-Tris al 4-12 % (Invitrogen), se transfirió la proteína biotinilada sobre una membrana de difluoruro de polivinilideno (PVDF) utilizando un Sistema de Transferencia de Geles IBlot (Invitrogen). A continuación se detectó el conjugado de biotina con un anticuerpo de cabra dirigido contra biotina conjugado con peroxidasa de rábano picante (HRP) (dilución 1: 100, Cell Signaling Technologies) y se visualizó utilizando una Hyperfilm ECL (GE Healthcare) (Figura 47B). mEGF-Tyr10PCL desacoplada y fluoresceína-conjugada con mEGF-Tyr10PCL sirvieron como controles negativos. La banda 1 20 pmol de conjugado de mEGF-Tyr10PCL-ABA-biotina; Banda 2, 8 pmol de conjugado de mEGF-Tyr10PCL-ABA-biotina; Banda 3, 2 nmol de mEGF-Y10PCL; Banda 4, 20 pmol de conjugado de mEGF-Tyr10PCL-ABA-fluoresceína.

Ejemplo 25: Moléculas de acoplamiento fluorescentes

65 Para demostrar que una molécula fluorescente se puede acoplar a proteínas que tienen uno o más restos de PCL incorporados en las anteriores, se conjugó mEGF-Tyr10PCL (véase el Ejemplo 12) con fluoresceína usando un reactivo de ABA-fluoresceína (véase 3793-050, Ejemplo 42). El acoplamiento de la fluoresceína se llevó a cabo como sigue, 1 mM de ABA-fluoresceína se añadió a 10 mM de mEGF-Tyr10PCL en solución salina tamponada con fosfato

(PBS, pH 7,0) y DMSO al 0,5 % (v/v) y se hicieron reaccionar a 25°C durante 16 horas. Se verificó la formación del conjugado de fluoresceína mediante ESI-MS (Figura 47C; masa esperada de la proteína sin acoplar = 7296; masa esperada de la proteína acoplada = 8062).

5 El conjugado de fluoresceína se redujo con 20 mM de NaCNBH₃ durante 3 horas a 25°C. tras inactivar rápidamente la reacción de reducción con seis volúmenes de PBS (pH 7,0), Se eliminó el NaCNBH₃ residual mediante diálisis frente al PBS (pH 7,0) a 4°C utilizando un casete de diálisis Slide-A-Lyzer (corte de pesos moleculares de 3.500 Da). A continuación se conjugó el concentrado hasta 1 mM utilizando una Unidad de Filtro de Centrifuga Amicon Ultra-4 con un corte de 3,5 kDa. Se obtuvieron los espectros de absorbancia en el intervalo de 350 - 700 nm de 1 mM de mEGF-Tyr10PCL-ABA-fluoresceína y de 10 mM ABA-fluoresceína utilizando un SpectraMax Plus (Molecular Devices).
10 Ambos proporcionaron una absorbancia máxima a 500 nm.

A continuación se registraron los espectros de fluorescencia en un fluorómetro SpectraMax GEMINI (Molecular Devices). Se obtuvieron un espectro de emisión para 1 mM de mEGF-Tyr10PCL-ABA-fluoresceína y de 10 mM de ABA-fluoresceína manteniendo la longitud de excitación a 490 nm, analizando a la vez la longitud de onda de emisión desde 510 nm a 750 nm utilizando un tamaño de 2 nm por etapa. Ambos proporcionaron una máximo de emisión a 522 nm. Se obtuvieron un espectro de excitación para 1 mM de mEGF-Tyr10PCL-ABA-fluoresceína y de 10 mM de ABA-fluoresceína manteniendo la longitud de excitación a 522 nm, analizando a la vez la longitud de onda de excitación desde 300 nm a 510 nm utilizando un tamaño de 2 nm por etapa.
15
20

Ejemplo 26: Polisacáridos de acoplamiento

Para demostrar que los polisacáridos se pueden acoplar a proteínas que tienen uno o más restos de PCL incorporados en los anteriores, se conjugó mEGF-Tyr10PCL (véase el Ejemplo 12) con un disacárido usando un reactivo de ABA-disacárido (3793-050; Ejemplo 42, PM 546,52). Se llevó a cabo la reacción de acoplamiento mediante la adición de 1 mM de ABA-disacárido a 10 mM de proteína mutante mEGF-Tyr10PCL en PBS y 1 % (v/v) de DMSO a pH 7,0. Se dejó continuar la reacción a temperatura ambiente durante 16 horas y se analizó mediante ESI-MS (Figura 47D). El espectro de masas muestra el pico principal en la masa esperada para la proteína conjugada (7825 Da). La masa de la proteína no acoplada es de 7296 Da.
25
30

Ejemplo 27: Acoplamiento de inmunomoduladores: conjugados de mono-nitrofenil hapteno

Para demostrar que se pueden acoplar inmunomoduladores a proteínas que tienen uno o más restos de PCL incorporados en las anteriores, mTNF-Gln21PCL y mEGF-Tyr10PCL (véanse los Ejemplos 11 y 12) se conjugaron con un mono-nitrofenil hapteno utilizando un reactivo de ABA-mono-nitrofenil hapteno (3793-001, Ejemplo 38-8). El acoplamiento del conjugado de mono-nitrofenil hapteno se llevó a cabo tal como se ha descrito anteriormente. La Figura 48A es el espectro de masas ESI de 3793-001 conjugado a mTNF-Gln21PCL (masa esperada de proteína no acoplada = 19275; masa esperada de la proteína acoplada = 19614), La Figura 48B es el espectro de masas ESI de 3793-001 conjugado a mEGF-Tyr10PCL (masa esperada de proteína no acoplada = 7296; masa esperada de la proteína acoplada = 7635).
35
40

Ejemplo 28: Acoplamiento de inmunomoduladores: conjugados de di-nitrofenil hapteno

Para demostrar que se pueden acoplar inmunomoduladores a proteínas que tienen uno o más restos de PCL incorporados en las anteriores, mTNF-Gln21PCL y mEGF-Tyr10PCL (véanse los Ejemplos 11 y 12) se conjugaron con un di-nitrofenil hapteno utilizando el di-nitrofenil hapteno (TU3627-088, Ejemplo 38-7). El acoplamiento del conjugado de di-nitrofenil hapteno se llevó a cabo tal como se ha descrito anteriormente. La Figura 48C es el espectro de masas ESI de TU3627-088 conjugado a mTNF-Gln21PCL (masa esperada de proteína no acoplada = 19275; masa esperada de la proteína acoplada = 19688), y la Figura 48D es el espectro de masas ESI de TU3627-088 conjugado a mTNF-Gln21PCL (masa esperada de proteína no acoplada = 7296; masa esperada de la proteína acoplada = 7709).
45
50

Ejemplo 29: Acoplamiento de inmunomoduladores: agonistas de TLR7.

Para demostrar que se pueden acoplar inmunomoduladores a proteínas que tienen uno o más restos de PCL incorporados en las anteriores, mEGF-Y10PCL (véase el Ejemplo 12) se conjugó con un agonista de TLR7 utilizando un reactivo agonista de ABATLR7 (véase X3678-114; Ejemplo 38-3). Se llevó a cabo la reacción de acoplamiento mediante la adición de 100 mM del agonista ABA-TLR7 a 10 mM de proteína mutante mEGF-Y10PCL en tampón acetato de sodio 200 mM y 1 % (v/v) de DMSO a pH 4,5. Se dejó continuar la reacción a temperatura ambiente durante 16 horas y se analizó mediante ESI-MS (Figura 49A). El espectro de masas muestra el pico principal en la masa esperada para la proteína conjugada (8763 Da), siendo la masa de la proteína no acoplada de 7296 Da. Ejemplo 30: Acoplamiento de inmunomoduladores: péptidos PADRE
55
60

Para demostrar que se pueden acoplar inmunomoduladores y péptidos a proteínas que tienen uno o más restos de PCL incorporados en las anteriores, mTNF-Gln21PCL y mEGF-Tyr10PCL (véanse los Ejemplos 11 y 12) se conjugaron con péptidos PADRE. Se llevó a cabo el acoplamiento de péptidos PADRE tal como se describe en el párrafo inmediatamente siguiente al subtítulo "Modificación específica de sitio de Pirrolina-Carboxi-Lisina (PCL) que
65

contiene proteínas con otras moléculas diversas". La Figura 50A es un análisis por espectrometría de masas MALDI-TOF de mTNF-Gln21PCL conjugada con PX2-PADRE (véase 3465-143; Ejemplo 38-11) -a pH 5,0, mientras que la Figura 50B es un análisis por espectrometría de masas MALDI-TOF de mTNF-Q21PCL conjugada con PX2-PADRE a pH 7,5. La masa esperada de la proteína sin acoplar es de 19275 Da y la masa esperada de la proteína acoplada es de 20842 Da. La Figura 50C es un análisis por espectrometría de masas ESI de mTNF-Gln21PCL conjugada con BHA-exPADRE (véase 3647-104; Ejemplo 38-10). Aquí, la masa esperada de la proteína sin acoplar es de 19275 Da y la masa esperada de la proteína acoplada es de 21317 Da. Además, La Figura 51 es un espectro de masas ESI que muestra el acoplamiento de BHA-exPADRE con mEGF-Tyr10PCL (masa esperada de la proteína sin acoplar 7296 Da; masa esperada de la proteína acoplada 9338 Da).

Ejemplo 31: Acoplamiento de inmunomoduladores: Fosfolípidos

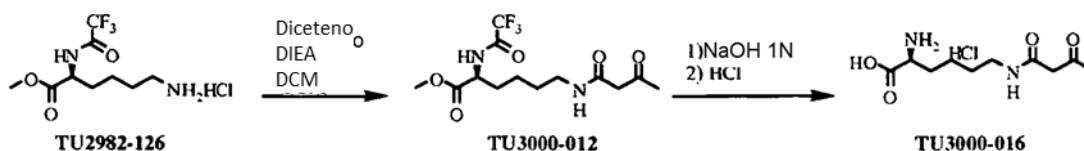
Para demostrar que se pueden acoplar inmunomoduladores y fosfolípidos a proteínas que tienen uno o más restos de PCL incorporados en las anteriores, se conjugó mEGF-Tyr10PCL (véase el Ejemplo 12) con un fosfolípido (DOPE) usando un reactivo de fosfolípido ABA (TU3627-092; Ejemplo 43-1). El acoplamiento de ABA-DOPE a mEGF-Tyr10PCL (PM = 7296 Da) se llevó a cabo en 20 mM HEPES (pH 7.0) y DMSO al 1 % (v/v) a 25°C durante 16 horas. La reacción de conjugación se inició mediante la adición de 10 mM de mEGF-Tyr10PCL y 100 mM de ABA-DOPE. La formación del conjugado de proteína se verificó mediante análisis de espectrometría de masas con ionización por pulverización (ESI-MS) de la conjugación de DOPE con mEGF-Y10PCL (Figura 49B; masa esperada de la proteína sin acoplar = 7296; masa esperada de la proteína acoplada = 8227 Da).

Ejemplo 32: Acoplamiento de oligonucleótido a proteína: Péptidos CpG

Para demostrar que se pueden acoplar oligonucleótidos e inmunomoduladores CpG a proteínas que tienen uno o más restos de PCL incorporados en las anteriores, mTNF-Gln21PCL y mEGF-Tyr10PCL (véanse los Ejemplos 11 y 12) se conjugaron con oligonucleótidos CpG utilizando tanto reactivo CPG BHA-BG1 (3647-057; Ejemplo 38-12) o reactivo CpG BHA-BG2 (véase 3597-167; Ejemplo 38-14). Se llevó a cabo el acoplamiento del oligonucleótido CpG tal como se describe en el párrafo inmediatamente siguiente al subtítulo "Modificación específica de sitio de Pirrolina-Carboxi-Lisina (PCL) que contiene proteínas con otras moléculas diversas". El acoplamiento de BHA-BG1 1 (7,4 kDa) y BHA-BG2 (7,4 kDa) con mTNF- Gln21PCL (19,3 kDa) se confirmó mediante el ensayo de cambio de geles (Figura 52A), y el acoplamiento de BHA-BG2 (7,4 kDa) con mEGF- Tyr10PCL (7,2 kDa) se confirmó también mediante el ensayo de cambio de geles (Figura 52B).

Ejemplo 33: Síntesis de análogos de pirrolisina reactivos

Ejemplo 33-1: Síntesis del clorhidrato del ácido (S)-2-amino-6-(3-oxobutanamido)hexanoico (TU3000-016)

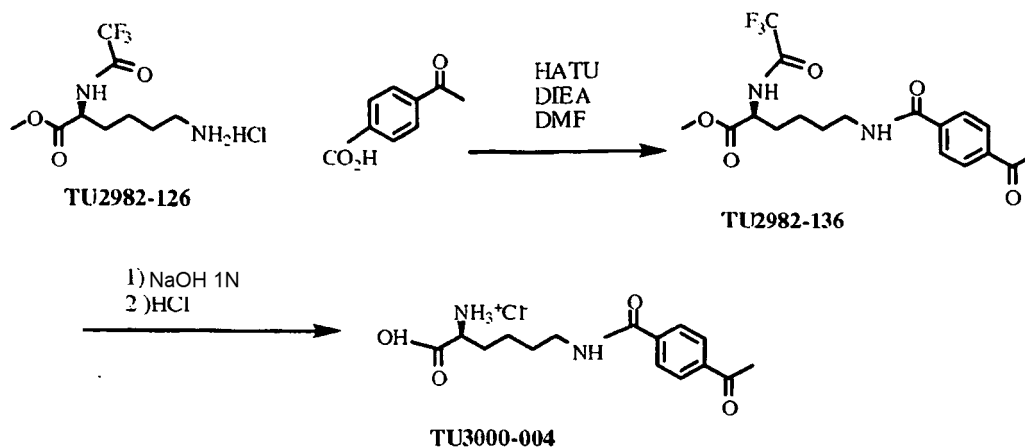


Se preparó el clorhidrato de 6-amino-2-(2,2,2-trifluoroacetamido)hexanoato de (S)-metilo (TU2982-126) de acuerdo con el procedimiento descrito en Bing Hao y col., ChemBio 2004, 11, 1317-24 y su material suplementario excepto que se llevó a cabo la eliminación del grupo Cbz en presencia de HCl para dar como resultado la sal de HCl.

Para el clorhidrato de 6-amino-2-(2,2,2-trifluoroacetamido)hexanoato de (S)-metilo (0,943 g, 3,22 mmol), N,N-diisopropil etilamina (DIEA, 1,39 ml) y diclorometano (DCM, 10 ml) en un vial de vidrio de 40 ml se añadió diceteno (0,37 ml) en atmósfera de N₂, y la reacción se agitó a temperatura ambiente durante 16 horas. La mezcla de reacción se diluyó con acetato de etilo (EtOA), se lavó sucesivamente con H₂O, HCl 1 N, H₂O, disolución acuosa saturada de Na₂CO₃, y disolución acuosa saturada de NaCl, se secó con Na₂SO₄, se filtró y se concentró a presión reducida. El residuo se purificó mediante cromatografía instantánea en SiO₂, dando como resultado 6-(3-oxobutanamido)-2-(2,2,2-trifluoroacetamido)hexanoato (S)-metilo (TU3000-012) como un aceite de color amarillo. MS (ESI+): calc. 341,12, encontrado 341,10 (MH⁺). H-RMN (400 MHz, CDCl₃): 1,33 (2H, m), 1,55 (2H, m), 1,89 (2H, m), 2,259 (3H, s), 3,31 (2H, m), 3,410 (2H, s), 3,783 (3H, s), 4,543 (1H, dt, J = 4,4, 8,0 Hz), 7,091 (1H, br.s), 7,279 (1H, br.d, J = 7,6 Hz), F-RMN (376 MHz, CDCl₃): -75,609.

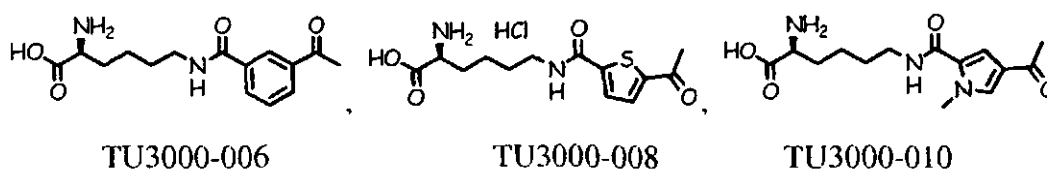
6-(3-oxobutanamido)-2-(2,2,2-trifluoroacetamido)hexanoato de (S)-metilo (0,736 g, 2,16 mmol) se trataron con NaOH acuoso 1 N (6,5 ml) y 10 ml H₂O a temperatura ambiente durante 18 horas. El análisis de LC-MS se utilizó para revelar que se completó la reacción. La mezcla de reacción se concentró a presión reducida. El residuo se trató con un exceso de HCl acuoso 1 N, y se concentró a sequedad a presión reducida, Síntesis del clorhidrato del ácido (S)-2-amino-6-(3-oxobutanamido)hexanoico (TU3000-016). MS (ESI+): calc. 231,13, encontrado 231,10 (MH⁺).

Ejemplo 33-2: Síntesis del cloruro de (S)-5-(4-acetilbenzamido)-1-carboxipentan-1-aminio (TU3000-004)



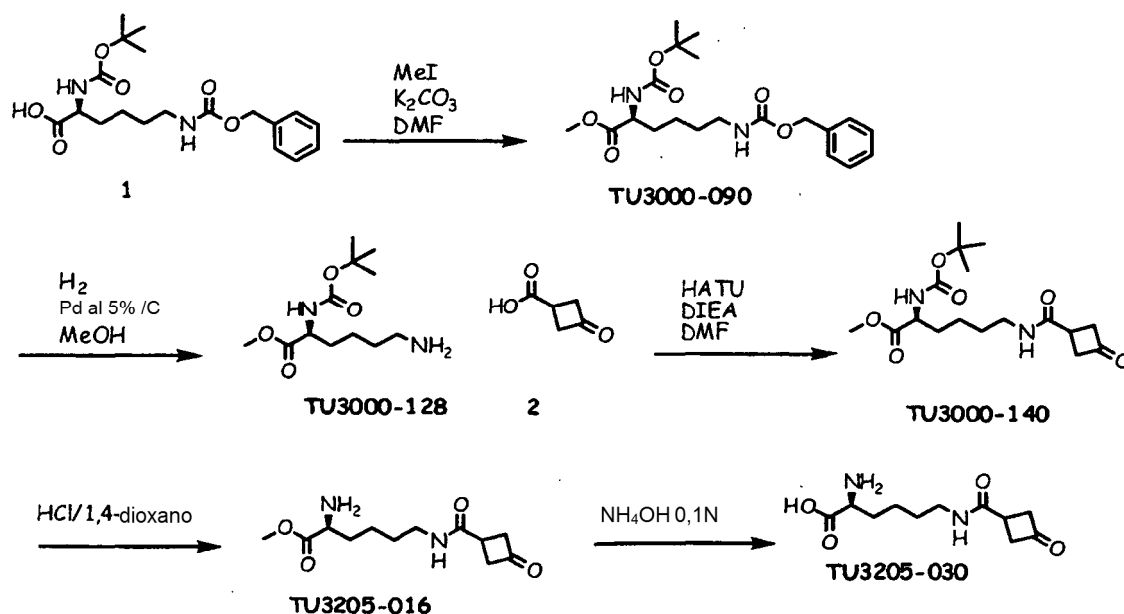
- 5 Clorhidrato de 6-amino-2-(2,2,2-trifluoroacetamido)hexanoato de (S)-metilo (TU2982-126) (303 mg, 1,03 mmol), Ácido 4-acetilbenzoico (190 mg), HATU (418 mg), DIEA (523 ml) y DMF (8 ml) se combinaron y agitaron a temperatura ambiente durante 18 horas. La mezcla de reacción se diluyó con EtOA, y se lavó sucesivamente con H₂O, HCl 1 N, H₂O, disolución acuosa saturada de Na₂CO₃, y disolución acuosa saturada de NaCl, se secó con Na₂SO₄, se filtró y se concentró a presión reducida. El residuo se purificó mediante cromatografía instantánea en SiO₂, dando como
- 10 resultado el clorhidrato de 6-amino-2-(2,2,2-trifluoroacetamido)hexanoato de (S)-metilo (TU2982-136). MS (ESI⁺): calc. 403,14, encontrado 403,20 (MH⁺). H-RMN (400 MHz, CDCl₃): 1,44 (2H, m), 1,68 (2H, m), 1,90 (1H, m), 2,00 (1H, s), 2,64 (3H, s), 3,31 (2H, m), 3,50 (2H, m), 3,80 (3H, s), 4,60 (1H, dt, J = 4,8, 7,6 Hz), 6,35 (1H, br.s), 7,21 (2H, br.d, J = 7,64 Hz), 7,86 (2H, d, J = 8,4 Hz), 8,01 (2H, d, J = 8,8 MHz), F-RMN (376 MHz, CDCl₃): -75,625.
- 15 6-(4-Acetilbenzamido)-2-(2,2,2-trifluoroacetamido)hexanoato de (S)-metilo (TU2982-136) (0,473 g) se trató con NaOH acuoso 1 N (2,36 ml) en 10 ml de MeOH a temperatura ambiente durante 18 horas. Se usó el análisis de LC-MS para revelar la hidrólisis completa del éster de metilo, manteniendo a la vez el resto de trifluoroamida intacto. Se calentó la reacción a 60°C durante 5 horas, en cuyo momento la hidrólisis de la amida fue casi completa, tal como se indicó mediante el análisis de LC-MS. La mezcla de reacción se enfrió y se concentró a presión reducida. El residuo se trató
- 20 con un exceso de HCl acuoso 1 N, y se concentró a sequedad a presión reducida, dando como resultado cloruro de ((S)-5-(4-acetilbenzamido)-1-carboxipentan-1-aminio (TU3000-004) como un sólido de color ligeramente amarillo. MS (ESI⁺): calc. 293,14, encontrado 293,20 (MH⁺).

- 25 *Ejemplo 33-3: síntesis del cloruro de ((S)-5-(5-acetil-1H-pirrol-2-carboxamido)-1-carboxipentan-1-aminio (TU3000-006), cloruro de (S)-5-(3-acetilbenzamido)-1-carboxipentan-1-aminio (TU3000-008), y del cloruro de (S)-5-(4-acetil-1-metil-1H-pirrol-2-carboxamido)-1-carboxipentan-1-aminio (TU3000-010)*



- 30 el cloruro de (S)-5-(5-acetil-1H-pirrol-2-carboxamido)-1-carboxipentan-1-aminio (TU3000-006), el cloruro de (S)-5-(3-acetilbenzamido)-1-carboxipentan-1-aminio (TU3000-008), y el cloruro de (S)-5-(4-acetil-1-metil-1H-pirrol-2-carboxamido)-1-carboxipentan-1-aminio (TU3000-010) se prepararon de la misma manera que en TU3000-004 pero utilizando los ácidos correspondientes en vez del ácido 4-acetilbenzoico. El ácido utilizado fue el ácido 4-acetil-1-metil-1H-pirrol-2-carboxílico, el ácido 5-acetil-1H-pirrol-2-carboxílico y el ácido 3-acetilbenzoico para TU3000-006, TU3000-008, y TU3000-010, respectivamente. TU3000-006: MS (ESI): m/z = 293,20 [M+H]; TU3000-008: MS (ESI): m/z = 299,10 [M+H]; y TU3000-010: MS (ESI): m/z = 296,20 [M+H].
- 35

Ejemplo 33-4: Síntesis del ácido (S)-2-amino-6-(3-oxociclobutanocarboxamido)hexanoico (TU3205-030).



5 Yodometano (2,0 ml), K₂CO₃ (5,60 g), ácido (S)-6-(benciloxycarbonilamino)-2-(terc-butoxicarbonilamino)hexanoico (1) (NovaBiochem, A29340), y DMF anhidro (20 ml) se combinaron y agitaron a temperatura ambiente durante 2 horas. la mezcla se sometió a tratamiento acuoso, dando como resultado 6-(benciloxycarbonilamino)-2-(terc-butoxicarbonilamino)hexanoato de (S)-metilo (TU3000-090) como un aceite transparente. MS (ESI+): calc. 417,20, encontrado 417,20 (MNa⁺), calc. 295,16, encontrado 295,209 ((M-Boc)H⁺). H-RMN (400MHz, CDCl₃):1,372 (2H, m), 1,425 (9H, s), 1,515 (2H, m), 1,623 (1H, m), 1,785 (1H, br.s), 3,175 (2H, m), 3,723 (3H, s), 4,284 (1H, m), 4,848 (1H, br.s), 5,084 (3H, solapando br.s y s), 7,344 (5H, m). C-RMN (100 MHz, CDCl₃): 22,340, 28,268, 29,307, 32,313, 40,569, 52,272, 53,099, 66,597, 79,897, 128,061. 128,100, 128,471, 136,519, 155,435, 156,426, 173,230.

15 6-(Benciloxycarbonilamino)-2-(terc-butoxicarbonilamino)hexanoato de (S)-metilo (8,60 g) se hidrogenó en 150 ml de MeOH con paladio al 5 % sobre carbono activado (1,08 g) con 1 atm de hidrógeno a temperatura ambiente durante 3 horas. El catalizador gastado se eliminó mediante filtración a vacío a través de un parche de celite que se lavó con MeOH. El filtrado combinado y el lavado se concentraron a presión reducida, dando como resultado 6-amino-2-(terc-butoxicarbonilamino)hexanoato de (S)-metilo (TU3000-128) como un aceite viscoso claro. MS (ESI+): calc. 261,17, encontrado 261,20 (MH⁺). H-RMN (600MHz, CDC13):1,328 (2H, m), 1,375 (9H, s), 1,390 (2H, m), 1,55 (1H, m), 1,74 (1H, m), 2,623 (2H, d, J = 6,9Hz), 3,671 (3H, s), 4,237 (1H, m), 5,007 (1H, m).

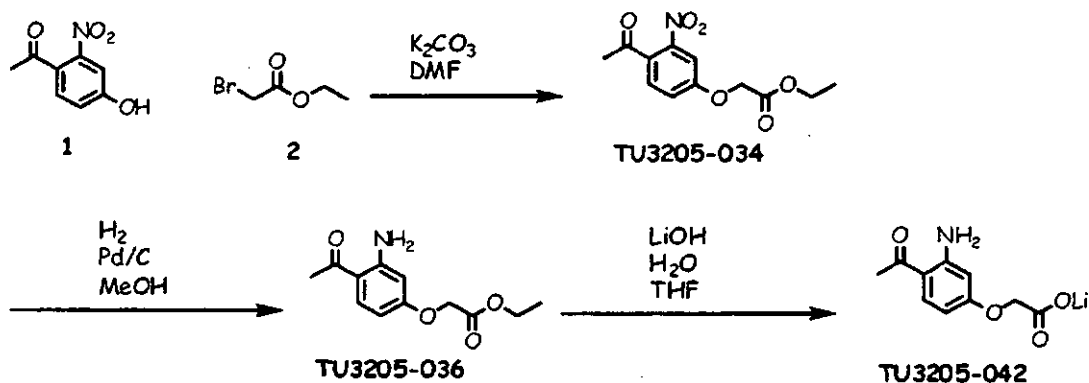
25 Se preparó una disolución madre 1 M de 6-amino-2-(terc-butoxicarbonilamino)hexanoato de (S)-metilo en 20 ml de DMF, y se usó. Una alícuota de 2 ml del 6-amino-2-(terc-butoxicarbonilamino)hexanoato de (S)-metilo de disolución madre, ácido 3-oxociclobutanocarboxílico (2) (Parkway *BX-102, 283mg), HATU (800mg), DIEA (1,0 ml) y 8 ml DMF se combinaron en un vial de vidrio de 40ml, y se agitó a temperatura ambiente durante la noche. El análisis de LC-MS reveló una reacción completa. La mezcla de reacción se diluyó con EtOAc, se lavó sucesivamente con agua, disolución saturada acuosa NaCl, se secó con MgSO₄, se filtró y se concentró a presión reducida. El material bruto se purificó mediante cromatografía instantánea en gel de sílice (hexanos/EtOAc), dando como resultado el 2-(terc-butoxicarbonilamino)-6-(3-oxociclobutanocarboxamido)hexanoato de (S)-metilo (TU3000-140) como un aceite viscoso transparente. MS (ESI+): calc. 379,18, encontrado 379,20 (MNa⁺), calc. 257,15, encontrado 257,20 ((M-Boc)H⁺). H-RMN (400MHz, DMSO-d₆):1,245 (4H, m), 1,371 (9H, s), 1,579 (2H, m), 3,065 (3H, m), 3,135 (4H, m), 3,608 (3H, s), 3,895 (1H, m), 7,209 (1H, d, J = 8,0Hz), 8,123 (1H, t, J = 5,4Hz), C-RMN (100 MHz, DMSO-d₆): 22,850, 27,185, 28,111, 28,551, 30,209, 38,276, 50,859, 51,639, 53,412, 53,529, 78,137, 155,507, 172,992, 173,167, 205,576.

35 (S)-2-(terc-butoxicarbonilamino)-6-(3-oxociclobutanocarboxamido)hexanoato de metilo (0,501 g) se trató con 20 ml de HCl 4 M en 1, 4-dioxano a temperatura ambiente durante 20 minutos y el disolvente se eliminó a presión reducida. El aceite viscoso resultante se capturó en 10 ml de CH₃CN, y se sembró con pequeñas cantidades de cristales del compuesto preparado en una escala más pequeña. Los cristales resultantes se recogieron mediante filtración a vacío, se lavaron con CH₃CN, y se secaron a presión reducida, dando como resultado el 2-amino-6-(3-oxociclobutanocarboxamido)hexanoato de (S)-metilo (TU3205-016) como un sólido incoloro. MS (ESI+): calc. 257,15, encontrado 257,20 (MH⁺). H-RMN (400MHz, DMSO-d₆):1,267 (1H, m), 1,419 (3H, m), 1,753 (2H, m), 3,114 (7H, m), 3,752 (3H, m), 4,023 (1H, t, J = 6,4Hz), 8,169 (1H, t, J = 5,5Hz), 8,365 (3H, br.s). C-RMN (100 MHz, DMSO-d₆): 21,499, 27,160, 28,390, 29,548, 38,067, 50,858, 51,721, 52,727, 169,949, 173,034, 205,586.

(S)-amino-6-(3-oxociclobutanocarboxamido)hexanoato de metilo (0,297 g) y 18 ml de H₂O se colocaron en un vial de vidrio de 40 ml. A la disolución transparente resultante se añadieron 2,2 ml de NH₄OH 1N, y la reacción se agitó a temperatura ambiente durante 22 horas, momento en el que el análisis por LC-MS reveló una reacción completa. La mezcla de reacción se congeló y liofilizó, dando como resultado el ácido (S)-2-amino-6-(3-oxociclobutanocarboxamido)hexanoico (TU3205-030) como cristales incoloros. MS (ESI): m/z 243,20 [M+H].

Ejemplo 34: Síntesis de intermedios de pirrolisina reactivos

10 **Ejemplo 34-1:** Síntesis de 2-(4-acetil-3-aminofenoxi)acetato de litio (TU3205-042)

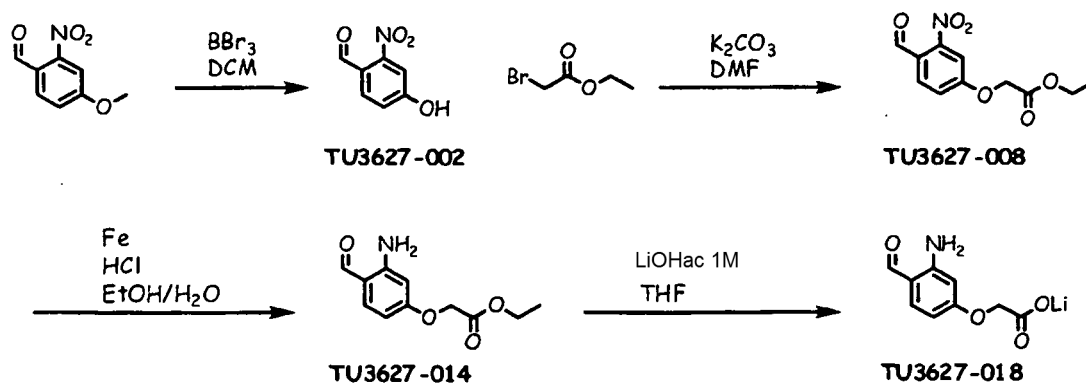


15 1-(4-Hidroxi-2-nitrofenil)etanona (Carbocore, 181 mg, 1,00 mmol), 2-Bromoacetato de etilo (183 mg, 1,10 mmol), carbonato de potasio (138 mg, 1,00 mmol) y DMF (5 ml) se combinaron en un vial de vidrio de 20 ml y se agitaron a 60°C durante 2 horas, en cuyo momento, el análisis de LC-MS mostró una reacción completa limpia. La mezcla de reacción se diluyó con agua y se extrajo con EtOAc, se lavó con disolución saturada de NaCl. Se repitió la reacción utilizando 1-(4-Hidroxi-2-nitrofenil)etanona (0,802 g, 4,43 mmol), El 2-bromoacetato de etilo (0,813 g, 4,87 mmol), carbonato de potasio (0,612 g, 4,43 mmol) y DMF (25 ml), y se elaboró de la misma manera. Los extractos de EtOAc combinados se secaron con MgSO₄ anhidro, se filtraron y se concentraron a presión reducida, dando como resultado acetato de 2-(4-acetil-3-nitrofenoxi)acetato de etilo (TU3205-034) como un aceite de color amarillo oscuro MS (ESI⁺): calc. 268,07, encontrado 268,10 (MH⁺). H-RMN (400MHz, CDCl₃): 1,317 (3H, t, J = 7,2 Hz), 2,512 (3H, s), 4,294 (2H, c, J = 7,2 Hz), 4,722 (2H, s), 7,192 (1H, dd, J = 8,4 Hz, 2,4 Hz), 7,455 (1H, d, J = 8,4 Hz), 7,462 (1H, d, J = 2,8 Hz), C-RMN (100 MHz, CDCl₃): 14,133, 29,695, 61,922, 65,499, 110,105, 119,735, 129,461, 130,154, 147,793, 159,356, 167,474, 198,352.

20 2-(4-Acetil-3-nitrofenoxi)acetato de etilo (TU3205-034) (111 mg) en MeOH (5 ml) se hidrogenó utilizando paladio al 10 % sobre carbón activo (11 mg) a presión atmosférica de H₂ a temperatura ambiente. Después de 30 minutos, el análisis de LC-MS reveló una reacción completa limpia. La reacción se repitió usando TU3205-034 (1,45 g), paladio al 10 % sobre carbón activo (140 mg), y MeOH (80 ml). Las dos mezclas de reacción se combinaron, y el catalizador gastado se eliminó mediante filtración a través de un parche de celite. El filtrado se concentró a presión reducida, dando como resultado el 2-(4-acetil-3-aminofenoxi)acetato de etilo (TU3205-036) como un sólido amarillo oscuro. MS (ESI⁺): calc. 238,10, encontrado 238,10 (MH⁺). RMN H (400 MHz, CDCl₃): 1,295 (3H, t, J = 7,0 Hz), 2,507 (3H, s), 4,269 (2H, c, J = 7,2 Hz), 4,603 (2H, s), 6,047 (1H, d, J = 2,8 Hz), 6,236 (1H, dd, J = 8,8 Hz, 2,8 Hz), 6,396 (2H, br.s), 7,647 (1H, d, J = 9,2 Hz), C-RMN (100 MHz, CDCl₃): 14,135, 27,658, 61,532, 64,939, 100,237, 104,030, 113,586, 134,286, 152,468, 162,294, 168,310, 199,055.

35 Se disolvió 2-(4-acetil-3-aminofenoxi)acetato de etilo (TU3205-036) (1,02 g) en THF (17 ml) y se trató con LiOH acuoso (1 M, 4,3 ml) a temperatura ambiente durante 1 hora. El análisis de LC-MS mostró una reacción completa limpia. La mezcla de reacción se concentró a presión reducida, dando como resultado acetato de litio 2-(4-acetil-3-aminofenoxi) (TU3205-042) como un sólido amarillento. MS (ESI⁺): calc. para el ácido libre 210,10, encontrado 210,10 (MH⁺).

40

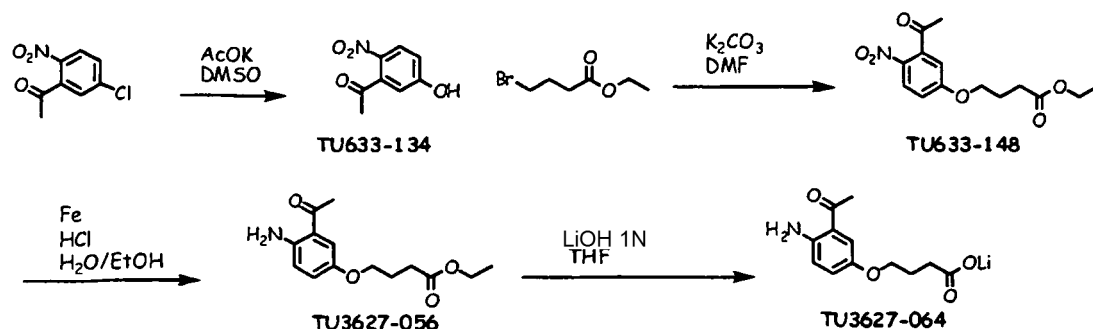
Ejemplo 34-2: Síntesis de 2-(3-amino-4-formilfenoxi)acetato de litio (TU3627-018).

5 Para 4-metoxi-2-nitrobenzaldehído (CarboCore, CO-0119, 5,5 g) y 200 ml de DCM e un matraz de fondo redondo de 500 ml se añadió tribromuro de boro (24 g) gota a gota con enfriamiento en un baño de hielo. La reacción se agitó a la misma temperatura durante 30 minutos y a continuación a temperatura ambiente durante 3 horas. La mezcla de reacción se vertió cuidadosamente en agua helada y se dejó reposar a temperatura ambiente durante 3 días. La mezcla acuosa se extrajo con EtOAc, se lavó con disolución saturada acuosa, NaCl, se secó con Na_2SO_4 , se filtró y se concentró a presión reducida. El material bruto se purificó mediante cromatografía instantánea en gel de SiO_2 usando un gradiente lineal de 20 a 60 % de EtOAc en hexanos (Rf. 0,32, 50 % de EtOAc en hexanos), dando como resultado TU3627-002 como cristales de color naranja. MS (ESI+): calc. 168,02, encontrado 168,10 (MH+). H-RMN (400MHz, DMSO-d_6): 8 7,211 (1H, dd, J = 2,4, 8,4Hz), 7,362 (1H, d, J = 2,4Hz), 7,866 (1H, d, 8,8Hz), 9,998 (1H, s), 11,466 (1H, s). C-RMN (100 MHz, DMSO-d_6): δ 110,726, 119,683, 120,770, 132,616, 151,225, 162,650, 187,885.

15 TU3627-002 (1,87 g), Bromoacetato de etilo (Aldrich, 1,5 ml), carbonato de potasio (2,32 g) y DMF se combinaron y agitaron a temperatura ambiente durante 18 horas. La mezcla de reacción se repartió entre EtOAc y agua. Se separó la capa orgánica, se lavó con disolución saturada acuosa NaCl, se secó con MgSO_4 , se filtró y se concentró a presión reducida. El material bruto se purificó mediante cromatografía instantánea en gel de SiO_2 usando un gradiente lineal de 5 a 35 % de EtOAc en hexanos (Rf. 0,32, 35 % de EtOAc en hexanos), dando como resultado TU3627-008 como un aceite amarillo oscuro. MS (ESI+): calc. 254,1, encontrado 254,1 (MH+). H-RMN (400MHz, CDCl_3): 8 1,304 (3H, t, J = 7,0Hz), 4,285 (2H, c, J = 7,2Hz), 4,767 (2H, s), 7,233 (1H, dd, 2,0, 8,4Hz), 7,522 (1H, d, J = 2,4Hz), 7,967 (1H, 8,4Hz), 10,286 (1H, s). C-RMN (100 MHz, CDCl_3): δ 14,082, 61,971, 65,482, 69,302, 110,427, 119,520, 124,321, 131,525, 151,286, 161,66, 167,172, 186,841.

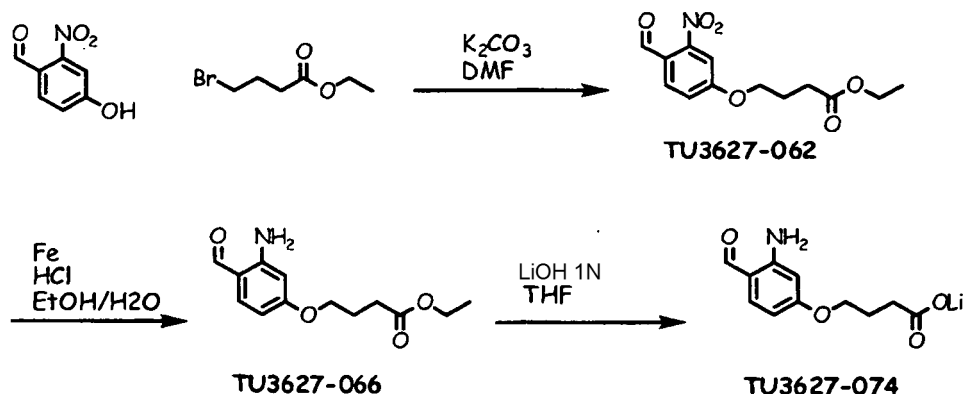
25 TU3627-008 se redujo mediante el método descrito por (C. A. Merlic y col., J.Org.Chem., 1995, 60, 3365-69). Específicamente, TU3627-008 (1,717g), polvo de hierro (3,79 g), EtOH (45 ml), H_2O (11 ml) y HCl conc. (180 ml) se combinaron y calentaron a reflujo durante 2 horas. La mezcla de reacción se filtró, y el filtrado se concentró. El residuo se purificó mediante cromatografía instantánea en gel de sílice usando un gradiente lineal de 5 a 60 % de disolvente B en disolvente A (disolvente A: 5 % de NEt_3 en hexanos; disolvente B: 5 % de NEt_3 en EtOAc) (Rf. 0,34, 35 % de EtOAc en hexanos), dando como resultado TU3627-014 como cristales amarillo claro. MS (ESI+): calc. 224,10, encontrado 224,10 (MH+). H-RMN (400MHz, DMSO-d_6): 8 1,216 (3H, t, J = 7,2Hz), 4,173 (2H, c, 7,2Hz), 4,777 (2H, s), 6,168 (1H, d, J = 2,4Hz), 6,248 (1H, dd, J = 2,4, 8,8Hz), 7,183 (2H, br.s), 7,441 (1H, d, J = 8,4 hz), 9,646 (1H, s). C-RMN (100 MHz, DMSO-d_6): δ 13,976, 60,711, 64,354, 98,521, 103,988, 113,256, 137,699, 152,665, 162,849, 168,176, 191,670.

35 TU3627-014 (0,608 g) en 5ml de THF se trató con una alícuota de 2,72 ml de una disolución acuosa de LiOH 1M. a temperatura ambiente. Después de 20 minutos, el análisis de LC-MS reveló una reacción completa. La mezcla de reacción se concentró a presión reducida hasta sequedad, dando como resultado el 2-(3-amino-4-formilfenoxi)acetato de litio (TU3627-018) como un sólido amarillo. MS (ESI): m/z 196,1 [M+H]. H-RMN (400MHz, DMSO-d_6): 4,132 (2H, s), 6,110 (1H, d, J = 2,0Hz), 6,145 (1H, dd, J = 2,0, 8,8Hz), 7,155 (2H, br.s), 7,33 (1H, d, J = 8,8 hz), 9,576 (1H, s). C-RMN (100 MHz, DMSO-d_6): 67,675, 98,354, 105,031, 112,386, 137,042, 152,914, 164,589, 169,349, 191,056.

Ejemplo 34-3: Síntesis de 4-(3-acetil-4-aminofenoxi)butanoato de litio (TU3627-064).

- 5 Una mezcla de 4'-cloro-2'-nitroacetofenona (Bionet n° de cat.3W0333, 1,00g) y DMSO se calentó a 170°C durante 20 minutos con irradiación microondas. La mezcla de reacción se repartió entre EtOAc y agua. Se separó la capa orgánica, se lavó con disolución saturada acuosa, NaCl, se secó con Na₂SO₄, se filtró y se concentró a presión reducida. Se llevó a cabo otra reacción de la misma manera, y los productos brutos procedentes de las dos reacciones se combinaron para la purificación mediante una cromatografía instantánea en gel de sílice (EtOAc). Se obtuvo el compuesto del título como un sólido de color naranja. MS (ESI+): calc. 182,15, encontrado 182,10 (MH⁺). H-RMN (400MHz, DMSO-d₆): 8,2,474 (3H, s), 6,837 (1H, d, J = 2,4Hz), 6,973 (1H, dd, J = 2,4, 8,8Hz), 8,078 (1H, d, J = 8,8Hz), 11,307 (1H, s). C-RMN (100 MHz, DMSO-d₆): δ 30,155, 113,264, 116,446, 127,386, 136,211, 140,986, 163,417, 200,099.
- 10
- 15 Una mezcla de 4'-hidroxi-2'-nitroacetofenona (1,12 g), 4-bromobutanoato de etilo (1,33 g), carbonato de potasio (0,94 g) y 4 ml de DMF se agitaron a 60°C durante 6,5 horas. La mezcla de reacción se repartió entre EtOAc y agua. Se separó la capa orgánica, se lavó con disolución saturada acuosa, NaCl, se secó con Na₂SO₄. se filtró y se concentró a presión reducida. El producto bruto se purificó mediante cromatografía instantánea en gel de sílice (hexanos/EtOAc), dando como resultado 1, el compuesto del título como cristales de color amarillo. MS (ESI+): calc. 296,11, encontrado 296,20 (MH⁺). H-RMN (400MHz, CDC13): 8,1,255 (3H, t, J = 7,2Hz), 2,141 (2H, quint, J = 6,6Hz), 2,506 (2H, t, J = 7,2Hz), 2,512 (3H, s), 4,116 (2H, t, J = 6,0Hz), 4,143 (2H, c, J = 7,2Hz), 6,757 (1H, d, J = 2,8Hz), 6,973 (1H, dd, J = 2,8, 9,2Hz), 8,125 (1H, d, J = 8,8Hz). C-RMN (100 MHz, CDC13): δ 14,182, 24,136, 30,282, 30,360, 60,616. 67,907, 112,305, 115,150, 127,023, 138,036, 141,205, 163,580, 172,730, 200,054.
- 20
- 25 TU633-148 (1,48 g) se redujo mediante el método descrito por Merlic (C.A.Merlic y col., J.Org.Chem., 1995, 60, 3365-69), tal como se describe en el presente documento. TU3627-056 se obtuvo como un aceite amarillo claro tras la cromatografía instantánea en gel de sílice usando un gradiente lineal de 5 a 60 % de disolvente B en disolvente A (disolvente A: 5 % de NEt₃ en hexanos; disolvente B: 5 % de NEt₃ en EtOAc). MS (ESI+): calc. 266,13, encontrado 266,20 (MH⁺). H-RMN (400MHz, CDC13): 8,1,258 (3H, t, J = 7,2Hz), 2,085 (2H, quint, J = 6,8Hz), 2,512 (2H, t, J = 7,2Hz), 2,556 (3H, s), 3,955 (2H, t, J = 6,0Hz), 4,146 (2H, c, J = 7,2Hz), 6,066 (2H, br.s), 6,632 (1H, d, J = 8,8Hz), 6,954 (1H, dd, J = 2,8, 8,8Hz), 7,186 (1H, d, J = 2,8Hz). C-RMN (400MHz, CDC13): δ 14,218, 24,688, 27,990, 30,729, 60,434, 67,790, 115,933. 118299, 118,621, 123,550, 144,593, 149,271, 173,213, 200,283.
- 30
- 35 TU3627-056 (0,50 g) en 3 ml de THF se trataron con una alícuota de 1,9 ml de una disolución acuosa de LiOH 1M a temperatura ambiente durante 18 horas. El análisis de LC-MS reveló una reacción completa limpia. La mezcla de reacción se concentró a presión reducida hasta sequedad, dando como resultado el 4-(3-acetil-4-aminofenoxi)butanoato de litio (TU3627-064) como un sólido de color amarillo. MS (ESI): m/z 238,10 [M+H]. H-RMN (400MHz, DMSO-d₆): 1,823 (2H, quint, J = 6,8Hz), 2,008 (2H, t, J = 6,8Hz), 2,489 (3H, s), 3,871, (2H, t, J = 6,8Hz), 6,694 (1H, d, 9,2Hz), 6,808 (2H, s), 6,962 (1H, dd, J = 2,8, 8,8Hz), 7,177 (1H, d, J = 2,8Hz). RMN C (100MHz, DMSO-d₆): 26,079, 28,014, 34,164, 68,394, 114,896, 116,388, 118,059, 123,861, 145,529, 147,995, 176,119, 199,756.
- 40

Ejemplo 34-4: Síntesis de 4-(3-amino-4-formilfenoxi)butanoato de litio (TU3627-074).

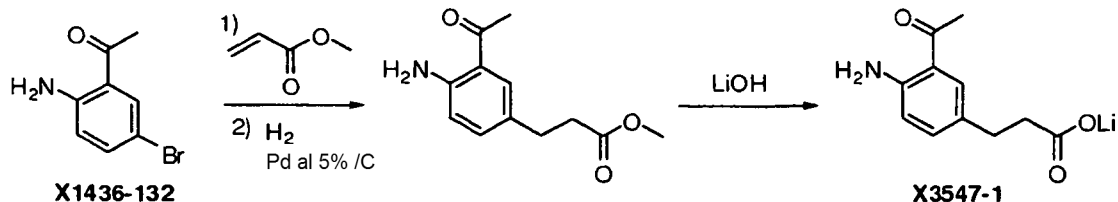


5 Una mezcla de 2-nitro-4-hidroxibenzaldehído (2,85 g), 4-bromobutanoato de etilo (3,66g), carbonato de potasio (2,84 g) y DMF (20 ml) se combinaron y agitaron a temperatura ambiente durante 50 horas. La mezcla de reacción se repartió entre H₂O y EtOAc. La capa orgánica se separó y se lavó con disolución saturada acuosa de NaHCO₃. Las capas acuosas combinadas se extrajeron con EtOAc. Las capas orgánicas conminadas se lavaron con ácido cítrico diluido, la disolución saturada acuosa de NaCl, se secó con MgSO₄, se filtró y se concentró. El residuo se purificó mediante cromatografía instantánea en gel de sílice usando un gradiente lineal de 20 a 40 % de EtOAc en hexanos (Rf: 0,35, 35 % de EtOAc en hexanos), dando como resultado el compuesto del título como un aceite de color amarillo. MS (ESI+): calc. 282,09, encontrado 282,10 (MH⁺). H-RMN (400MHz, CDCl₃): 8 1,261 (3H, t, J = 7,2Hz), 2,170 (2H, quint, J = 6,8Hz), 2,528 (2H, t, J = 7,2Hz), 4,153 (2H, d, J = 7,2Hz), 4,167 (2H, t, 6,4Hz), 7,217 (1H, dd, J = 2,4, 8,4Hz), 7,498 (1H, d, J = 2,4Hz), 7,964 (1H, d, J = 8,8Hz), 10,277 (1H, s).RMN-C (100MHz, CDC13): δ 14,190, 24,112, 27,712, 30,291, 32,443, 32,771, 60,648, 68,075, 110,024, 119,397, 123,405, 131,455, 151,572, 162,916, 172,710, 186,943.

TU3627-062 (3,89 g) se redujo mediante el método descrito por Merlic (C.A.Merlic y col., J.Org.Clem., 1995, 60, 3365-69), tal como se describe en el presente documento. TU3627-066 se obtuvo como un aceite amarillo claro tras la cromatografía instantánea en gel de sílice usando un gradiente lineal de 20 a 60 % de disolvente B en disolvente A (disolvente A: 5 % de NEt₃ en hexanos; disolvente B: 5 % de NEt₃ en EtOAc). Rf: 0,51, 50 % de EtOAc en hexanos). MS (ESI+): calc. 252,12, encontrado 252,20 (MH⁺). H-RMN (400MHz, DMSO-d₆): 8 1,177 (3H, t, J = 7,2Hz), 1,962 (2H, quint, J = 6,8hz), 2,441 (2H, t, J = 7,2Hz), 3,979 (2H, t, J = 6,4Hz), 4,064 (2H, c, J = 7,2Hz), 6,206 (1H, s), 6,219 (1H, d, J = 8,8Hz), 7,411 (1H, d, J = 8,8Hz), 9,623 (1H, s). C-RMN (100 MHz, DMSO-d₆): δ 14,041, 24,028, 30,002, 59,837, 66,459, 98,075, 104,394, 112,822, 137,558, 152,871, 163,809, 172,404, 191,535.

TU3627-066 (1,00 g) en 6 ml de THF se trató con 3,98 ml de disolución acuosa de LiOH 1M a temperatura ambiente durante 4 horas. La mayoría del disolvente se eliminó a presión reducida, y la mezcla turbia resultante se diluyó con agua doblemente desionizada, se congeló y se liofilizó, dando como resultado el 4-(3-amino-4-formilfenoxi)butanoato de litio (TU3627-074) como un sólido amarillo gris. MS (ESI+): m/z 224,20 (MH⁺). H-RMN (400MHz, DMSO-d₆): 1,852 (2H, quint, J = 6,8Hz), 2,040 (2H, t, 6,8Hz), 3,949 (2H, t, J = 6,8Hz), 6,191 (1H, dd, 2,4, 8,8Hz), 6,239 (1H, d, J = 2,4Hz), 7,208 (2H, br.s), 7,367 (1H, d, J = 8,8Hz), 9,597 (1H, s). C-RMN (100 MHz, DMSO-d₆): 25,624, 33,706, 67,799, 97,997, 104,677, 112,589, 137,427, 153,003, 164,195, 176,575, 191,394.

Ejemplo 34-5: Síntesis de 3-(3-acetil-4-aminofenil)propanoato de litio (X3547-1).



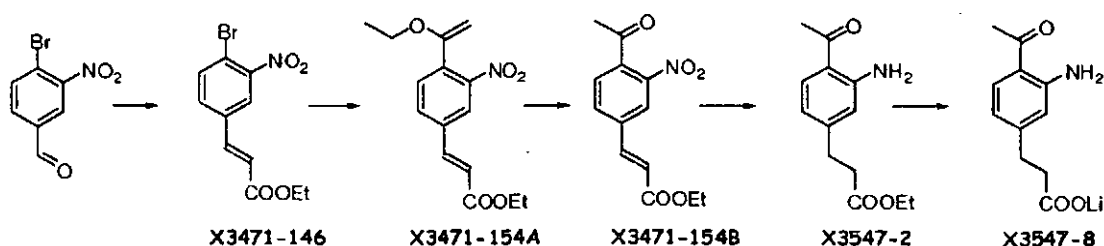
A una mezcla de 1-(2-amino-5-bromofenil)etanona (642 mg), Pd(OAc)₂ (33,7 mg), y P(otolil)₃ (137 mg) en DMF anhidro (10 ml) en un tubo de presión se añadió acrilato de metilo (351 ml) y TEA (1,4 ml). La mezcla se lavó con N₂ durante 3 minutos y a continuación se selló y se calentó a 110 °C durante 4 horas. La mezcla de reacción se enfrió a temperatura ambiente y a continuación se repartió entre acetato de etilo y agua. La capa acuosa se extrajo una vez con acetato de etilo, y las capas orgánicas combinadas se lavaron con salmuera, se secaron (Na₂SO₄), se filtraron, el disolvente se eliminó a vacío. El residuo bruto se purificó mediante cromatografía instantánea en gel de sílice (EtOAc/hexanos), dando como resultado el producto. RMN ¹H (400 MHz, MeOD) : 8 7,96 (d, J = 2,0 Hz, 1H), 7,63 (d, J = 16,0 Hz, 1H), 6,67 (dd, J = 8,8 Hz, J' = 2,0 Hz, 1H), 6,77 (d, J = 8,8 Hz, 1H), 6,29 (d, J = 16,0 Hz, 1H), 3,76 (s, 3H), 2,59 (s, 3H);

MS-ESI+220,24 (MH⁺).

3-(3-acetil-4-aminofenil)acrilato de (E)-metilo (219 mg) se redujo en un 5 % Pd/C (21,9 mg) en MeOH con balón de hidrógeno a temperatura ambiente, dando como resultado el producto (cuantitativo) tras la filtración y la concentración. El producto se usó para la siguiente etapa sin purificación adicional. RMN ¹H (400 MHz, MeOD) : 8,759 (d, J = 2,0 Hz, 1H), 7,13 (dd, J = 8,4 Hz, J' = 2,0 Hz, 1H), 6,67 (d, J = 8,4 Hz, 1H), 3,64 (s, 3H), 2,81 (dd, J = 7,6 Hz, J' = 7,6 Hz, 2H), 2,60 (dd, J = 7,6 Hz, J' = 7,6 Hz, 2H), 2,54 (s, 3H); MS-ESI+ 222,25 (MH⁺).

3-(3-acetil-4-aminofenil)propanoato de metilo (221 mg) y LiOH 4 M (0,275 ml) se añadieron a 3 ml de THF/agua (v/v = 3/1). El disolvente se eliminó bajo presión reducida después que se completó la reacción, Dando como resultado 3-(3-acetil-4-aminofenil)propanoato de litio (X3547-1). RMN ¹H (400 MHz, MeOD) : 8,762 (d, J = 2,0 Hz, 1H), 7,17 (dd, J = 8,4 Hz, J' = 2,0 Hz, 1H), 6,67 (d, J = 8,4 Hz, 1H), 2,80 (dd, J = 7,6 Hz, J' = 9,0 Hz, 2H), 2,55 (s, 3H), 2,41 (dd, J = 7,6 Hz, J' = 9,0 Hz, 2H); MS-ESI+ 207,22 (MH⁺).

15 *Ejemplo 34-6: Síntesis de 3-(4-acetil-3-aminofenil)propanoato de litio (X3547-8).*



Una disolución de trifetilfosforanilidina acetato de etilo (1,742 g) disuelto en CH₃CN (15 ml) se añadió con agitación a una disolución de CH₃CN (10 ml) que contenía 4-bromo-3-nitrobenzaldehído (1,150g). La mezcla de reacción se mantuvo a reflujo durante la noche. Después que se enfriara la mezcla, el disolvente se eliminó a presión reducida, dando como resultado un sólido bruto. Se obtuvo el producto X3471-146 puro como un sólido de color blanco tras la cromatografía instantánea en columna de gel de sílice (hexanos/EtOAc, 9:1).

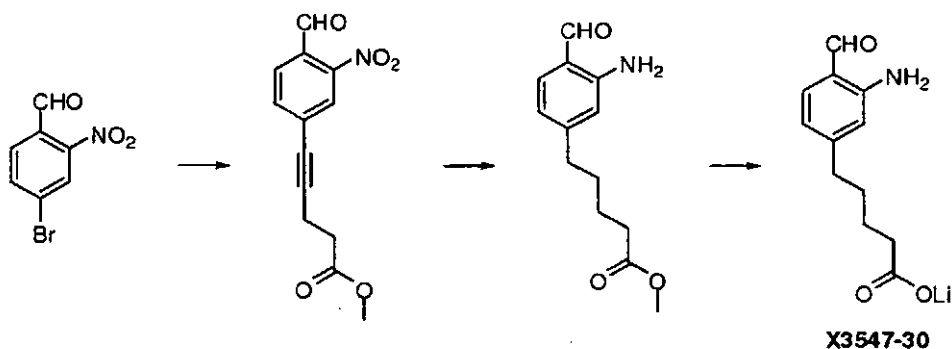
El producto X3471-146 (632 mg) procedente de la etapa anterior y PdCl₂(PPh₃)₂ (148 mg) se disolvieron en DMF (5,0 ml) con N₂ en un tubo Schlenk. Se añadió tributil(1-etoxivinil)estannano (711 ml) con agitación. La mezcla se calentó a 100°C durante la noche. Después que se enfriara la mezcla, el disolvente se eliminó a presión reducida, se diluyó con DCM, se lavó con agua y salmuera. Tras la eliminación del DCM, se purificó el residuo mediante cromatografía en columna de gel de sílice (15 % -25 % de EtOAc en hexanos) para dar el compuesto X3471-154.

El compuesto X3471-154A se trató con 20 ml de HCl 1N a temperatura ambiente durante 4 horas. La eliminación del disolvente dio como resultado el (E)-3-(4-acetil-3-nitrofenil)acrilato de etilo X3471-154B. RMN ¹H (400 MHz, MeOD) : 8,831 (d, J = 1,6 Hz, 1H), 8,04 (dd, J' = 1,6 Hz, J'' = 8,0 Hz, 1H), 7,76 (d, J = 16,0 Hz, 1H), 7,67 (d, J = 8,0 Hz, 1H), 6,74 (d, J = 16,0 Hz, 1H), 4,27 (q, J = 7,2 Hz, 2H), 2,57 (s, 3H), 1,34 (t, J = 7,2 Hz, 3H); ESI-MS (m/z) 264,07 (MH⁺).

El compuesto X3471-154B (263 mg) se redujo a 3-(4-acetil-3-aminofenil)propanoato de etilo X3547-2 con 5 % de Pd/C (26 mg) en 5,0 ml de MeOH en hidrógeno a 1 atm. ESI-MS (m/z) 236,30 (MH⁺).

El compuesto X3547-2 se trató con la mezcla de LiOH 4 M (0,248 ml), THF (3,0 ml) y agua (1,0 ml) a temperatura ambiente. después que se completó la reacción, La mezcla de reacción se liofilizó, dando como resultado 3-(4-acetil-3-aminofenil)propanoato de litio (X3547-8). MS (ESI+): m/z 208,10 (MH⁺).

Ejemplo 34-7: Síntesis de 5-(3-amino-4-formilfenil)pentanoato de litio (X3547-30).

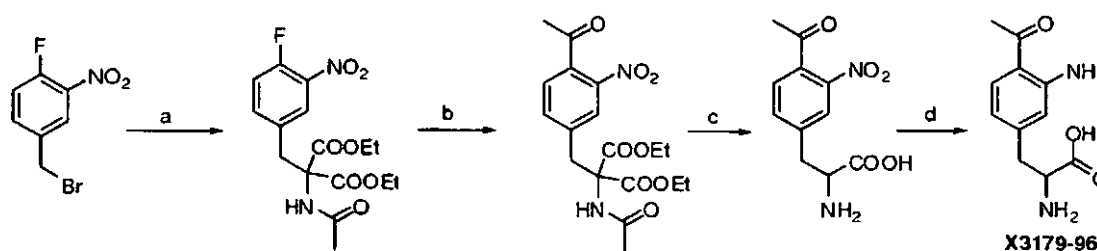


Una mezcla de 4-bromo-2-nitrobenzaldehído (460 mg), Pd(PPh₃)₂Cl₂ (70 mg), y CuI (38 mg), metil pent-4-inoato (269 mg) en TEA (5,0 ml) se agitó a temperatura ambiente hasta que se completó la reacción. El residuo bruto se purificó mediante cromatografía instantánea en gel de sílice (EtOAc al 15 % en hexanos), dando como resultado el metil 5-(4-formil-3-nitrofenil)pent-4-inoato. ESI-MS m/z 262,23 (MH⁺).

5 5-(4-formil-3-nitrofenil)pent-4-inoato de metilo (522 mg) se redujo en un 5 % de Pd/C (53 mg) en MeOH con un balón de hidrógeno a temperatura ambiente. La filtración seguida por concentración a presión reducida dio como resultado el 5-(3-amino-4-formilfenil)pentanoato de metilo. MS-ESI m/z 236,28 (MH⁺).

10 5-(3-amino-4-formilfenil)pentanoato de metilo (447 mg) se trató con LiOH (88 mg) en 8,0 ml de THF/agua (v/v = 3/1). La eliminación del disolvente después que se completó la reacción dio como resultado el 5-(3-amino-4-formilfenil)pentanoato de litio (X3547-30). MS (ESI⁺): 222,10 (MH⁺). RMN H (400 MHz, MeOD) : δ 9,72 (s, 1H), 7,37 (d, J = 8,0 Hz, 1H), 6,57 (s, 1H), 6,55 (d, J = 8,0 Hz, 1H), 2,56 (m, 2H), 2,19 (m, 2H), 1,64 (m, 4H).

15 *Ejemplo 34-8: Síntesis del ácido 3-(4-Acetil-3-aminofenil)-2-aminopropanoico (X3179-96).*



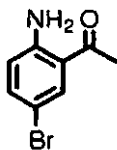
20 Hidruro de sodio (60 % en aceite mineral, 1,74 g) se lavó con hexanos y se suspendió en DMF (12 ml). Acetamidomalonato de dietilo (10,4 g) en DMF (30 ml) y 4-(bromometil)-1-fluoro-2-nitrobenzoceno (10,2 g) en DMF (10 ml) se añadieron sucesivamente y la mezcla de reacción se agitó durante 4 horas a temperatura ambiente, seguido por la eliminación del disolvente a presión reducida. El residuo se purificó mediante cromatografía instantánea en gel de sílice (EtOAc al 10-20 % en DCM), dando como resultado el 2-acetamido-2-(4-fluoro-3-nitrobenzil)malonato de dietilo. RMN ¹H (400 MHz, MeOD) : δ 7,77 (d, J = 8,4 Hz, 1H), 7,42 (s, 1H), 6,54 (d, J = 8,4 Hz, 1H), 4,25 (q, J = 7,2 Hz, 4H), 3,65 (s, 2H), 2,03 (s, 3H), 1,28 (t, J = 7,2 Hz, 6H); RMN ¹³C : δ 172,95, 168,23, 157,33, 154,72, 138,68, 134,19, 128,45, 119,31, 68,64, 63,88, 38,04, 22,31, 14,38; ESI-MS (m/z) 371,33 (MH⁺).

30 A 2-acetamido-2-(4-fluoro-3-nitrobenzil)malonato de dietilo (7,41 g) y nitroetano (6,0 ml) en EtOAc se añadió DBU (9,0 ml) y la mezcla se agitó a temperatura ambiente durante 16 horas. La mezcla de reacción se concentró y el residuo se disolvió en metanol (35 ml). Disolución acuosa de H₂O₂ (30 %, 10,2 ml) y disolución acuosa de NaHCO₃ al 10 % (10,2 ml) se añadieron y la mezcla se agitó a temperatura ambiente durante 16 horas. La mezcla de reacción se concentró a presión reducida. El residuo se disolvió en EtOAc, se lavó con HCl 1 N, salmuera, se secó con MgSO₄, se filtró y se concentró a presión reducida. El residuo se purificó mediante cromatografía instantánea en gel de sílice (EtOAc al 10-20 % en DCM), dando como resultado el 2-acetamido-2-(4-acetil-3-nitrobenzil)malonato de dietilo. RMN ¹H (400 MHz, MeOD) : δ 7,73 (s, 1H), 7,55 (d, J = 8,0 Hz, 1H), 7,47 (d, J = 8,0 Hz, 1H), 4,25 (q, J = 7,2 Hz, 4H), 3,72 (s, 2H), 2,53 (s, 3H), 2,03 (s, 3H), 1,28 (t, J = 7,2 Hz, 6H); RMN ¹³C : δ 172,99, 168,22, 153,68, 147,48, 140,96, 137,31, 136,82, 128,99, 126,79, 68,62, 63,95, 54,01, 38,58, 22,32, 14,36; ESI-MS (m/z) 395,38 (MH⁺).

40 Se disolvió 2-acetamido-2-(4-acetil-3-nitrobenzil)malonato de dietilo (2,0 g) en 10 ml de HCl al 37 %, se calentó a 100 °C durante la noche, y se enfrió. El sólido resultante se recogió por filtración, dando como resultado el ácido 3-(4-acetil-3-nitrofenil)-2-aminopropanoico. RMN ¹H (400 MHz, D₂O) : δ 8,01 (s, 1H), 7,53 (d, J = 7,6 Hz, 1H), 7,50 (d, J = 7,6 Hz, 1H), 4,07 (m, 1H), 3,20-3,34 (m, 2H), 2,52 (s, 3H); RMN ¹³C : δ 201,39, 170,78, 174,89, 140,02, 137,84, 136,28, 129,59, 126,46, 54,46, 36,56, 30,01; ESI-MS (m/z) 253,22 (MH⁺).

45 El ácido 3-(4-acetil-3-nitrofenil)-2-aminopropanoico se redujo en 5 % de Pd sobre carbono en MeOH con 1 atm H₂. La filtración seguida por concentración a presión reducida dio como resultado el ácido 3-(4-acetil-3-aminofenil)-2-aminopropanoico (X3179-96). MS (ESI⁺): m/z 223,22 (MH⁺). RMN ¹H (400 MHz, MeOD) : δ 7,78 (d, J = 8,4 Hz, 1H), 6,66 (s, 1H), 6,54 (d, J = 8,4 Hz, 1H), 4,25 (dd, J = 8,4 Hz, J'' = 5,2 Hz, 1H), 3,23 (dd, J' = 14,4 Hz, J'' = 5,2 Hz, 1H), 3,02 (dd, J = 14,4 Hz, J'' = 8,4 Hz, 1H), 2,54 (s, 3H); RMN ¹³C : δ 202,14, 171,20, 152,93, 142,51, 134,40, 118,78, 118,23, 116,91, 54,63, 37,42, 27,90.

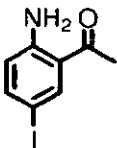
Ejemplo 34-9: Síntesis de 1-(2-Amino-5-bromofenil)etanona (X1436-132).



5 Un matraz cargado con 1-(2-aminofenil)etanona (135 mg) y DCM (2,5 ml) se enfrió a -10°C y NBS (178 mg) se añadió en varias porciones durante 30 min. La mezcla de reacción se diluyó con 10 ml de DCM, se lavó con disolución saturada acuosa de NaHCO_3 , se secó con Na_2SO_4 , se filtró y se concentró a presión reducida, dando como resultado el compuesto del título. MS (ESI+): m/z 215,06 (MH+). RMN ^1H (400 MHz, CDCl_3): δ 7,79 (s, 1H), 7,31 (d, J = 8,8 Hz, 1H), 6,54 (d, J = 8,8 Hz, 1H), 6,30 (br, 2H), 2,55 (s, 3H).

10

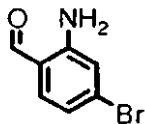
Ejemplo 34-10: Síntesis de 1-(2-amino-5-yodofenil)etanona (X1436-134).



15 Un matraz cargado con 1-(2-aminofenil)etanona (405 mg) y DCM (20 ml) se enfrió en un baño de agua helada y se añadió NIS (675 mg) en tres porciones. La reacción se agitó a 0°C . La mezcla de reacción se diluyó con 30 ml de DCM, se lavó con disolución saturada acuosa de NaHCO_3 , se secó con Na_2SO_4 , se filtró y se concentró a presión reducida. El material bruto se purificó mediante RP-HPLC preparativa, dando como resultado el compuesto del título. MS (ESI+): m/z 262,06 (MH+). RMN ^1H (400 MHz, CDCl_3): δ 7,94 (s, 1H), 7,44 (d, J = 8,8 Hz, 1H), 6,43 (d, J = 8,8 Hz, 1H), 6,32 (br, 2H), 2,53 (s, 3H).

20

Ejemplo 34-11: Síntesis de 2-Amino-4-bromobenzaldehído (X3547-158).



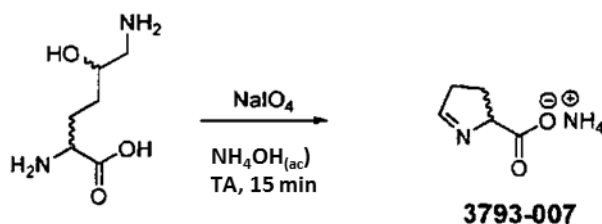
25 Polvo de hierro (1,20 g), agua (2,0 ml), y HCl (12 N, 40 ml) se añadieron consecutivamente a una disolución de 2-nitrobenzaldehído (230 mg) en etanol (8,0 ml). Tras agitar a 95°C durante 90 minutos, La mezcla de reacción se filtró caliente. Tras un lavado con etanol, los filtrados se combinaron y el disolvente se eliminó a vacío. El material bruto se purificó mediante cromatografía instantánea en gel de sílice (40:55:5 de acetato de hexandilato/trietilamina), dando como resultado el 2-amino-4-bromobenzaldehído como cristales de color amarillo. MS (ESI+): m/z 199,96 (MH+). RMN ^1H (400 MHz, MeOD) : δ 9,77 (s, 1H), 7,40 (d, J = 8,4 Hz, 1H), 6,95 (d, J = 1,6 Hz, 1H), 6,78 (dd, J = 8,4 Hz, J' = 1,6 Hz, 1H).

30

Ejemplo 35: Síntesis de precursores biosintéticos de PCL y pirrolisina

35

Ejemplo 35-1: Síntesis de DL-1-pirrolina-5-carboxilato de amonio (3793-007).

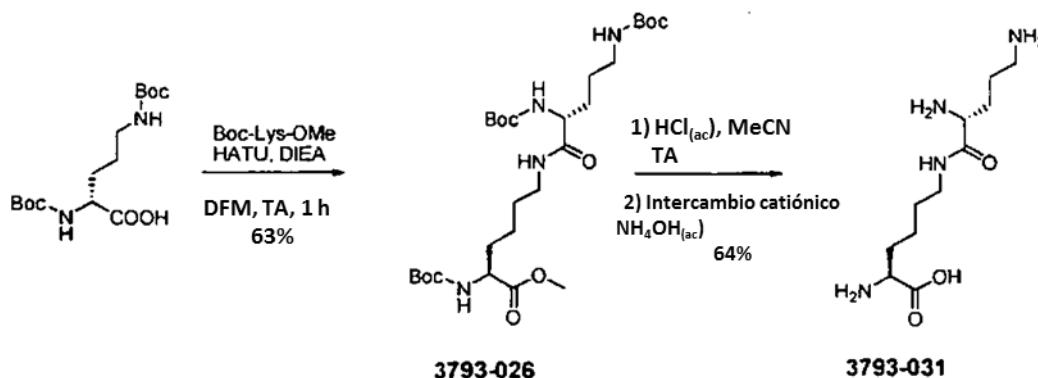


40 Sal de HCl de H-DL-8-DL-hidroxiLys-OH (118 mg, 0,5 mmol) se aplicó a un cartucho SPE de intercambio catiónico, y el aminoácido se eluyó con hidróxido de amonio. Se añadió NaIO_4 (107 mg) en el eluido de amonio y la mezcla de reacción se agitó a temperatura ambiente durante 30 minutos. La mezcla de reacción se liofilizó, y el producto bruto se

acidificó a pH 2,6 con HCl, y se purificó mediante SPE de intercambio catiónico. Se descartaron los primeros 10 ml de eluyente de H₂O se descartaron y los siguientes 30 ml de eluyente de H₂O se recogieron y se neutralizaron inmediatamente mediante NH₄OH_{aq} 1 N) (0,5 ml). Tras la liofilización, se obtuvo el producto deseado como un polvo amarillo claro. ESI-MS calculada para C₅H₇NO₂ [MH]⁺: 114,1; observada: 114,1.

5

Ejemplo 35-2: Síntesis de H-L-Lys-N^ε-(D-Orn)-OH (3793-031)

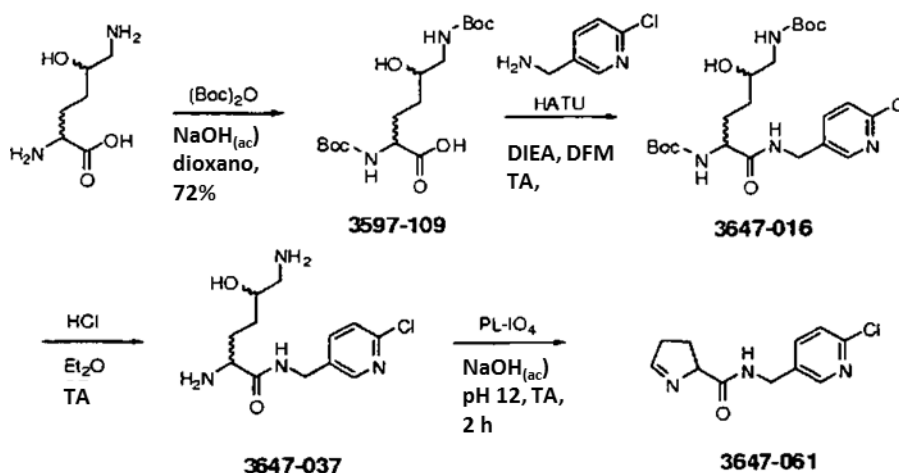


10 Boc-D-Orn(Boc)-OH (1,097 g) se trató con HATU (1,255 g) y DIEA (1,53 ml) en DMF (5 ml) durante 1 hora. Una
 11 disolución de DMF (5 ml) de Boc-Lys-OMe (sal de acetato, 1,057 mmol) y DIEA (627 ml), se añadió a continuación a la
 12 mezcla de reacción, y la reacción se agitó a temperatura ambiente durante 2 horas. La mezcla de reacción se repartió
 13 entre NaCl_(aq) al 10 % (30 ml) y EtOAc (60 ml). La capa de EtOAc se lavó con ácido cítrico al 5 % (30 ml), NaHCO_{3(aq)} al
 14 10 % (30 ml), y salmuera (30 ml). la capa de EtOAc se secó con Na₂SO_{4(s)}, se filtró y se concentró a presión reducida.
 15 El residuo se purificó mediante cromatografía instantánea en columna de gel de sílice con una columna de 80 g de
 16 sílice y un gradiente de elución de 0-15 % MeOH/DCM, dando como resultado Boc-Lys(Boc-D-Orn(Boc))-OMe
 17 (3793-026) como un aceite incoloro. RMN ¹H (400 MHz, CDCl₃): 6,54 (s a, 1H), 5,17 (br, 2H), 4,76 (s a, 1H), 4,4
 18 (cuartete, J = 6,7 Hz, 2H), 3,72 (s, 3H), 3,34-3,02 (m, 4H), 1,83-1,30 (m, 37H). RMN ¹³C (100 MHz, CDCl₃): 173,2,
 19 172,3, 156,6, 155,8, 155,5, 79,8, 79,2, 53,3, 53,0, 52,3, 39,2, 38,8, 32,0, 30,4, 29,1, 28,4, 28,3, 26,4, 22,5. ESI-MS
 20 calculado para C₂₇H₅₀N₄O₉ [MH]⁺: 575,4; observada: 575,4.

A una disolución acuosa al 65 % de acetonitrilo (5 ml) de Boc-Lys(Boc-D-Orn(Boc))-OMe (214 mg) se añadió HCl
 21 concentrado (1 ml), y la mezcla de reacción se agitó vigorosamente durante 2 horas. La mezcla de reacción se liofilizó,
 22 y se purificó mediante cromatografía de intercambio catiónico utilizando un cartucho de SCX SPE con NH₄OH 1 N en
 23 20 % de MeCN_{aq}. El éster de metilo se hidrolizó en el eluato de NH₄OH durante la noche. Tras la liofilización,
 24 H-L-Lys-N^ε-(D-Orn)-OH (3793-031) se obtuvo como un polvo de color blanco. RMN ¹H (400 MHz, D₂O): 3,33-3,30 (m,
 25 2H), 3,29-3,22 (m, 1H), 3,20-3,12 (m, 1H), 2,86 (t, J = 7,0 Hz, 2H), 1,65-1,48 (m, 8H), 1,36-1,25 (m, 2H). RMN ¹³C (100
 26 MHz, D₂O): 181,5, 176,9, 55,5, 54,3, 39,4, 38,9, 33,2, 31,4, 28,1, 24,5, 22,1. ESI-MS calculado para C₁₁H₂₄N₄O₃ [MH]⁺:
 27 261,2; observada: 261,2.

30

Ejemplo 35-3: Síntesis de N-((6-cloropiridin-3-il)metil)-1-pirrolin-5-carboxamida (3647-061).



35

t-Butilpirocarbonato (6,55 g) en dioxano (10 ml) se añadió gota a gota a DL-8-DL-hidroxiLys (1,99 g) en 1 N NaOH_(aq) (50 ml) y la reacción se agitó a temperatura ambiente durante la noche. La mezcla de reacción se acidificó a pH 3,0 con HCl_(aq) 1 N y se extrajo con EtOAc (200 ml) dos veces. Se combinaron las capas de EtOAc, se secaron sobre Na₂SO₄, se filtró y se concentró a presión reducida, dando como resultado el producto bruto. El producto bruto se purificó mediante cromatografía instantánea en columna de gel de sílice con gradiente de elución de 0-10 % de MeOH/DCM, dando como resultado Boc-DL-8-DL-hidroxiLys(Boc) (3597-109) como un sólido de color blanco. ESI-MS calculado para C₁₆H₃₀N₂O₇ [MNa]⁺: 385,2; observada: 385,2.

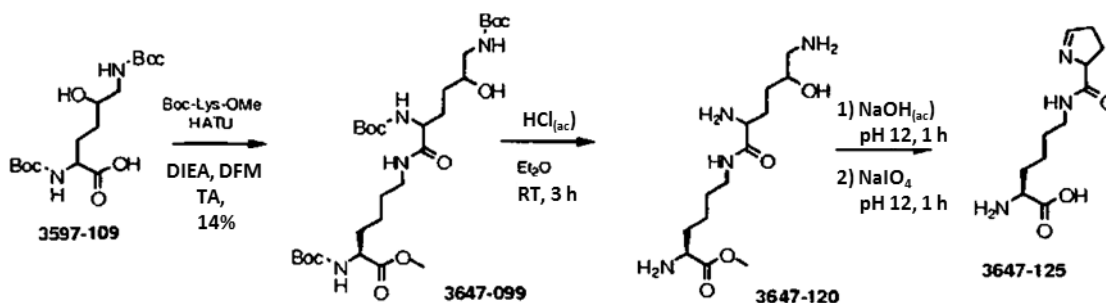
Boc-DL-8-DL-hidroxiLys(Boc) (435 mg) se trataron con HATU (456 mg) y DIEA (523 ml) en 3,5 ml DMF durante 1 hora. La disolución resultante se añadió a 6-cloropiridin-3-ilmetilamina (143 mg) en 1,5 l de disolución de DMF a temperatura ambiente y la reacción se agitó durante 2 días. La mezcla de reacción se extrajo con NaCl_(aq) al 10 % (15 ml) y EtOAc (30 ml). La capa de EtOAc se lavó con ácido cítrico al 5 % (15 ml), NaHCO₃ al 10 % (15 ml), y salmuera (15 ml). la capa de EtOAc se secó con Na₂SO₄, se filtró y se concentró a presión reducida, dando como resultado el producto bruto. El producto bruto se purificó mediante cromatografía instantánea en columna de gel de sílice con gradiente de elución de 0-15 % de MeOH/DCM, dando como resultado el producto (3647-016) como un sólido de color blanco. ESI-MS calculada para C₂₂H₃₅N₄O₆Cl [MH]⁺: 487,2; observada: 487,3.

El producto (3647-016) (263 mg) del anterior se agitó en una disolución de Et₂O (7 ml) y HCl_(aq) 1 N (7 ml) a temperatura ambiente durante la noche. La mezcla de reacción se evaporó y liofilizó, y el residuo se purificó mediante un cartucho SPE de intercambio catiónico, dando como resultado el producto (3647-037) como un sólido de color blanco. ESI-MS calculada para C₁₂H₁₉N₄O₂Cl [MH]⁺: 287,1; observada: 287,2.

El producto (3647-037) (69,6 mg) procedente del anterior se disolvió en 2 ml de H₂O (2 ml), y el pH se ajustó a 12 con NaOH_(aq) 1 N. se añadió la resina PL-IO₄ (Varian, 381mg, 0,48 mmol) y la reacción se agitó a temperatura ambiente durante 2 horas. La resina se eliminó mediante filtración y la mitad del filtrado se purificó mediante HPLC preparativa, dando como resultado el producto (3647-061) deseado como un sólido de color blanco. ESI-MS calculada para C₁₁H₁₂N₃OCl [MH]⁺: 238,1; observada: 238,1.

Ejemplo 36: Síntesis de compuestos modelo de PCL

Ejemplo 36-1: Síntesis de H-L-Lys-Ne-(DL-1-pirrolina-5-carbonil)-OH (3647-125).

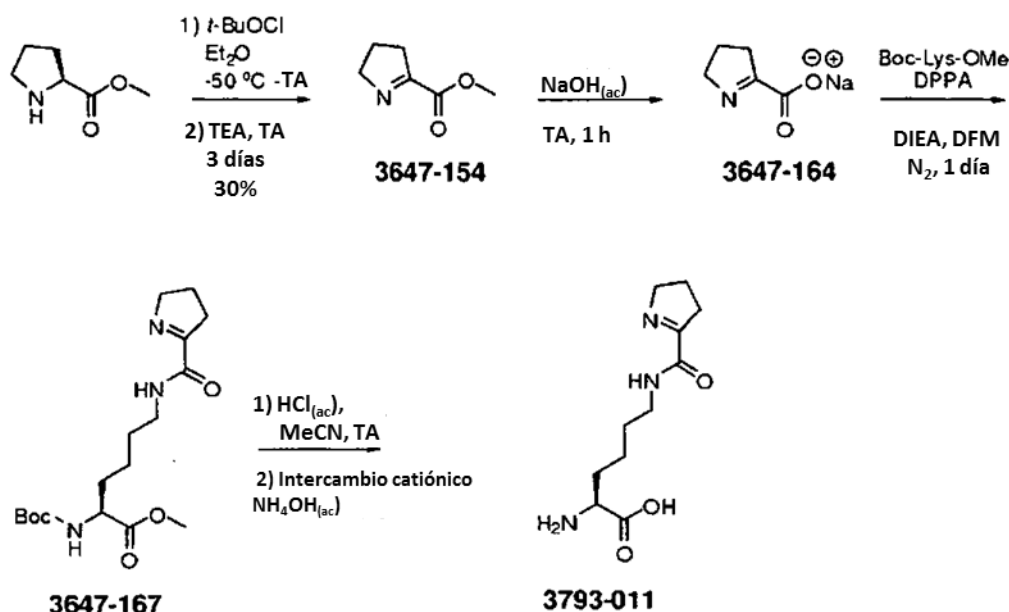


Boc-DL-8-DL-hidroxiLys(Boc) (3597-109) (886 mg) se activó mediante HATU (932mg) en 3,0 ml de DMF en presencia de DIEA (1068 ml) durante 1 hora, y se añadió a Boc-Lys-OMe (549 mg) en 3,0 ml de DMF (3,0 ml). La reacción se agitó a temperatura ambiente durante 36 horas. la mezcla de reacción se repartió entre NaCl_(aq) al 10 % (30 ml) y EtOAc (60 ml). La capa de EtOAc se lavó con ácido cítrico al 5 % (15 ml), NaHCO₃ al 10 % (15 ml), y salmuera (15 ml). La capa de EtOAc se secó sobre Na₂SO₄ se filtró y se concentró a presión reducida, dando como resultado el producto bruto. El producto bruto se purificó mediante cromatografía instantánea en columna de gel de sílice con 0-15 % de MeOH/DCM, seguido por RP-C18 SPE con la elución de 25 % y 65 % de MeCN_(aq), dando como resultado el producto (3647-099) deseado como un sólido de color blanco. ESI-MS calculado para C₂₈H₅₂N₄O₁₀ [MH]⁺: 605,4; observada: 605,4.

Para Boc-Lys(Boc-DL-8-DL-hidroxiLys(Boc))-OMe (3647-099) (84,1 mg) en 1 ml de Et₂O se añadió HCl_(aq) 4 N (3 ml) y la reacción se agitó a temperatura ambiente durante 3 horas. La eliminación del disolvente dio como resultado 61,1 mg del producto bruto (3647-120) como un sólido de color blanco. ESI-MS calculado para C₁₃H₂₈N₄O₄ [MH]⁺: 305,2; observada: 305,2.

El producto bruto (3647-120) (61,1 mg) procedente del anterior se disolvió en 1 ml de H₂O y se ajustó el pH a 12 con NaOH_(aq) 1 N. La mezcla de reacción se agitó a temperatura ambiente durante 1 hora, dando como resultado la hidrólisis del éster de metilo. se añadió NaIO₄ (29,9 mg) a la mezcla de reacción y la reacción se agitó a temperatura ambiente durante 15 minutos más. La mezcla de reacción se neutralizó con HCl_(aq) 1 N y se introdujo en un SPE de intercambio catiónico para la purificación para dar H-L-Lys-Ns-(DL-1-pirrolina-5-carbonil)-OH (3647-125). ESI-MS calculado para C₁₁H₁₉N₃O₃ [MH]⁺: 242,2; observada: 242,2.

Ejemplo 36-2: Síntesis de H-L-Lys-N^ε-(DL-1-pirrolina-2-carbonil)-OH (3793-011).



5 1-pirrolina-2-carboxilato de sodio (3647-164) mediante el siguiente método descrito por Gyorgy Szollosi y col. (Chirality, 2001, 13(10), 619-624).

10 Trietilamina (8,4 ml) se añadió a H-Pro-OMe HCl (7,512 g) en éter (27 ml) y la reacción se agitó durante 2 horas. La mezcla de reacción se filtró, y se eliminó el Et₂O mediante evaporación. El residuo se purificó mediante destilación a vacío, dando como resultado 4,119 g de H-Pro-OMe como un aceite incoloro. *t*-BuOCl (3,59 ml) se añadió gota a gota a una disolución de Et₂O (100 ml) de H-Pro-OMe (4,089 g) y la reacción se agitó a -50 °C durante 1 hora. La mezcla de reacción se calentó a temperatura ambiente y se añadió trietilamina (4,64 ml) a la mezcla de reacción, seguido por agitación durante 3 días. La mezcla de reacción se filtró, y se concentró a presión reducida. El residuo se purificó mediante destilación a vacío para obtener pirrolina-2-carboxilato de metilo (3647-154) como un aceite incoloro.

15 ESI-MS calculada para C₆H₉NO₂ [MH]⁺: 128,1; observada: 128,2. RMN ¹H (400 MHz, CDCl₃): 4,09 (tt, J = 7,6 Hz, 2,5 Hz, 2H), 3,85 (s, 3H), 2,81 (tt, J = 8,2 Hz, 2,8 Hz, 2H), 1,97 (quintuplete, J = 8,0 Hz, 2H). RMN ¹³C (100 MHz, CDCl₃): 168,2, 163,1, 62,5, 52,5, 35,2, 22,1. ESI-MS calculada para C₅H₇NO₂ [MH]⁺: 114,1; observada: 114,2. RMN ¹H (400 MHz, D₂O): 3,84 (tt, J = 7,4 Hz, 2,5 Hz, 2H), 2,73 (tt, J = 8,2 Hz, 2,6 Hz, 2H), 1,92 (quintuplete, J = 7,8 Hz, 2H). RMN ¹³C (100 MHz, D₂O): 175,5, 171,8, 60,1, 35,6, 21,8. ESI-MS calculada para C₅H₇NO₂ [MH]⁺: 114,1; observada: 114,2. RMN ¹H (400 MHz, D₂O): 3,84 (tt, J = 7,4 Hz, 2,5 Hz, 2H), 2,73 (tt, J = 8,2 Hz, 2,6 Hz, 2H), 1,92 (quintuplete, J = 7,8 Hz, 2H). RMN ¹³C (100 MHz, D₂O): 175,5, 171,8, 60,1, 35,6, 21,8.

25 Pirrolina-2-carboxilato de metilo (2,265 g) se disolvió en NaOH_(aq) 1 N (17,8 ml) y se agitó a temperatura ambiente durante 1 hora. La mezcla de reacción se liofilizó para dar pirrolina-2-carboxilato de sodio (3647-164) como un polvo de color blanco. ESI-MS calculada para C₅H₇NO₂ [MH]⁺: 114,1; observada: 114,2.

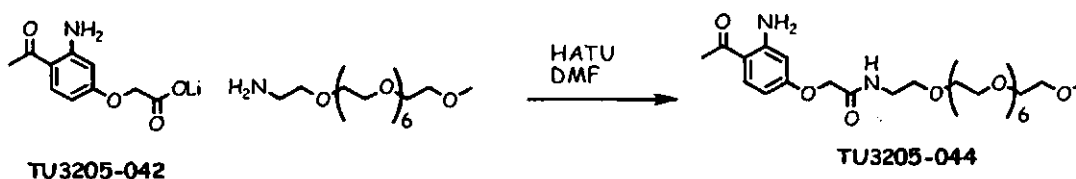
30 1-Pirrolina-2-carboxilato de sodio (459 mg) y DPPA (886 ml) se añadieron a Boc-Lys-OMe (sal de acetato, 897 mg) y DIEA (1184 ml) en 20 ml de DMF y la reacción se agitó con N₂ durante 24 horas. La mezcla de reacción se repartió entre NaCl_(aq) al 10 % (50 ml) y EtOAc (100 ml). La capa de EtOAc se secó con Na₂SO_{4(s)}, se filtró y se concentró a presión reducida, dando como resultado el producto bruto como un aceite de color naranja claro. El producto bruto se purificó mediante cromatografía instantánea en columna de gel de sílice con gradiente de elución de 0-15 % de MeOH/DCM, seguido por RP-C 18 SPE, dando como resultado el producto (3647-167) deseado como un sólido incoloro. ESI-MS calculado para C₁₇H₂₉N₃O₅ [MH]⁺: 356,2; observada: 356,0. RMN ¹H (400 MHz, CDCl₃): 7,14 (s, 1H), 5,08 (d, d=4 Hz), 4,27 (cuartete, J = 5,2 Hz, 1H), 3,99 (td, J = 7,2 Hz, 1,6 Hz, 2H), 3,72 (d, J = 1,6 Hz, 3H), 3,31 (cuartete, J = 6,9 Hz, 2H), 2,83 (td, J = 8,3 Hz, 1,3 Hz, 2H), 1,97 (quintuplete, J = 8,2 Hz, 2H), 1,88-1,72 (m, 2H), 1,69-1,51 (m, 2H), 1,48-1,33 (m, 11H). RMN ¹³C (100 MHz, CDCl₃): 171,6, 162,4, 155,4, 129,7, 79,8, 61,7, 53,2, 52,3, 38,7, 34,0, 32,2, 29,0, 28,3, 22,6, 22,5.

40 HCl (conc, 1 ml) se añadió a una disolución de CN_(aq) al 50 % (5 ml) del producto (3647-167) (49,3 mg) procedente del anterior y la mezcla de reacción se agitó a temperatura ambiente durante 2 horas. Tras la liofilización, el residuo se purificó mediante un cartucho SPE de intercambio catiónico. El éster de metilo se hidrolizó mediante NH₄OH_(aq) en el eluyente durante la noche. Tras la liofilización, se obtuvo el producto deseado (3793-011) como un polvo de color amarillo claro. ESI-MS calculada para C₁₁H₁₉N₃O₃ [MH]⁺: 242,2; observada: 242,2. RMN ¹H (400 MHz, D₂O): 3,94 (tt,

J = 7,6 Hz, 2,5 Hz, 2H), 3,65 (t, J = 6,2 Hz, 1H), 3,30 (t, J = 6,8 Hz, 2H), 2,78 (tt, J = 8,4 Hz, 2,6 Hz, 2H), 1,97 (quintuplete, J = 7,9 Hz, 2H), 1,87-1,77 (m, 2H), 1,59 (quintuplete, J = 7,2 Hz, 2H), 1,46-1,34 (m, 2H). RMN 13C (100 MHz, D₂O): 175,7, 172,2, 165,1, 60,9, 54,7, 38,9, 34,3, 30,5, 27,9, 21,8, 21,7. Ejemplo 37: Síntesis de reactivos 2-AAP-PEG

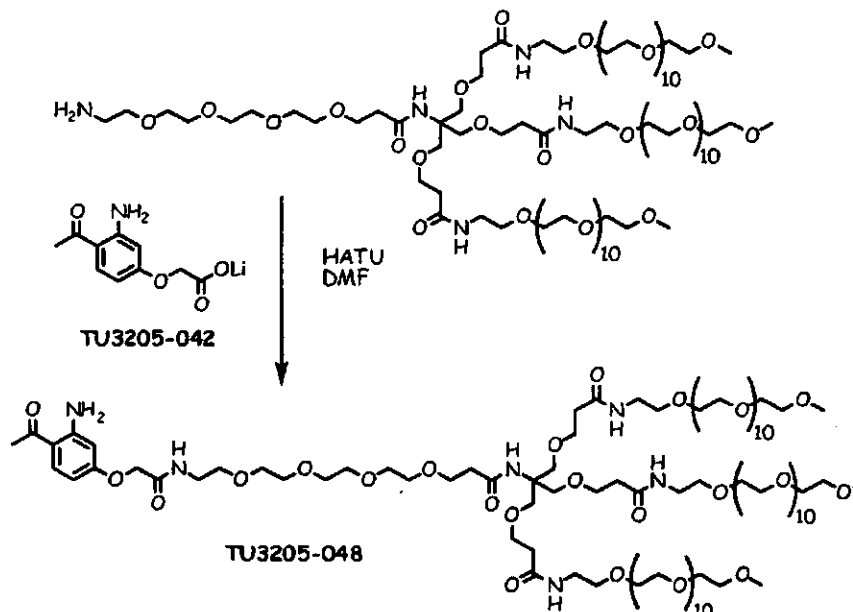
5 Se describe a continuación la síntesis de derivados de 2-AAP-PEG usados en los ejemplos proporcionados en el presente documento. derivados de 2-amino-acetofenona de PEG con pesos moleculares de aproximadamente 500, 2400, 23000 Da así como 5000 y 10000 Da se sintetizaron como sigue:

10 *Ejemplos 37-1: Síntesis de TU3205-044 (0,5 kDa 2-AAP-mPE)*



15 2-(4-Acetil-3-aminofenoxy)acetato de litio (TU3205-042) (55,9 mg) se introdujo en un matraz de fondo redondo de 10 ml, y se añadieron DMF (3 ml) y HATU (98,9 mg). La suspensión resultante se agitó a temperatura ambiente durante 40 minutos. La disolución se convirtió en una disolución de color amarillo durante este periodo. A la reacción se añadió a continuación mPEG-amina (Quanta Biodesign, PM 383,5, 100 mg) en 2 ml de DMF. Tras 5 minutos, el análisis de LC-MS mostró una reacción completa. La mezcla de reacción se agitó durante 40 minutos más y se concentró a presión reducida. el residuo se purificó mediante cromatografía instantánea en SiO₂ (MeOH al 3 % en DCM), dando como resultado TU3205-044 (0,5 kDa 2-AAP-mPE) como un aceite viscoso de color amarillo. MS (ESI+): calc. 575,31, encontrado 575,30 (MH⁺). H-RMN (400 MHz, CDCl₃): 2,210 (3H, s), 3,370 (3H, s), 3,544 (4H, m), 3,650 (28H, m), 4,499 (2H, s), 6,172 (1H, d, J = 2,4 Hz), 6,243 (1H, dd, J = 8,8 Hz, 2,8 Hz), 6,540 (2H, br.s), 7,034 (1H, br.s), 7,649 (1H, d, J = 8,8 Hz),

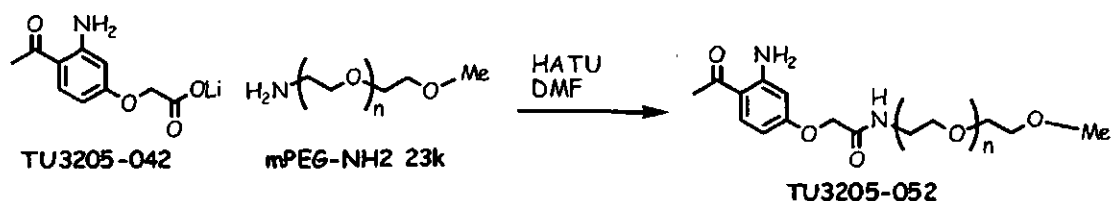
25 *Ejemplo 37-2: Síntesis de TU3205-048 (2,4 kDa 2-AAP-PE)*



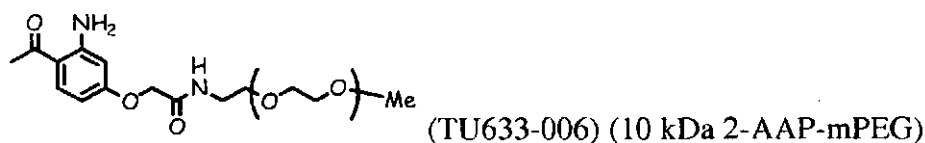
30 2-(4-Acetil-3-aminofenoxy)acetato de litio (TU3205-042) (12,2 mg) se introdujo en un matraz de fondo redondo de 10 ml, y se añadieron DMF (1 ml) y HATU (21,5 mg). La suspensión resultante se agitó a temperatura ambiente durante 35 minutos. La disolución se convirtió en una disolución de color amarillo durante este periodo. A la reacción se añadió a continuación mPEG-amina (Quanta Biodesign, PM 2209, 100 mg) en 2 ml de DMF. La mezcla de reacción se agitó durante 18 horas y se concentró a presión reducida. el residuo se purificó mediante cromatografía instantánea en SiO₂ (MeOH en DCM), dando como resultado TU3205-048 (2,4 kDa 2-AAP-mPE) como un aceite viscoso de color amarillo. MS (ESI+): calc. 800,8, encontrado 800,5 (M⁺³H⁺/3), calc. 600,9, encontrado 600,7 (M⁺⁴H⁺/4).

35

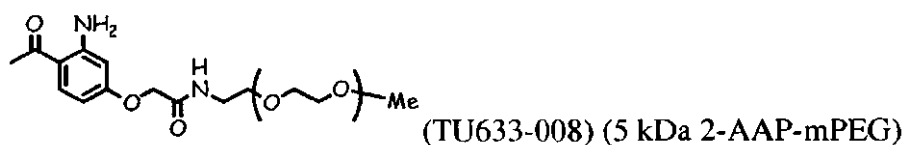
Ejemplo 37-3: Síntesis de TU3205-052 (23 kDa 2-AAP-PE)



- 5 2-(4-acetil-3-aminofenoxy)acetato de litio (TU3205-042) (21,5 mg) y HATU (38,0 mg) se cargaron en un matraz de fondo redondo de 10 ml, y se añadió DMF (0,5 ml). La suspensión resultante se agitó a temperatura ambiente durante 50 minutos. La disolución de color amarillo resultante se añadió a mPEG-NH₂ (Laysan Bio, PM promedio 23k, n promedio = 520, 0,50 g) en 5 ml de DMF es un vial de vidrio de 20 ml. La reacción se agitó a temperatura ambiente durante 2,5 horas. A continuación, la mezcla de reacción se diluyó con 10 ml de agua, y se aplicó una alícuota de 2,5 ml de cada una de las disoluciones a las columnas PD-10 (GE Healthtech) y el producto deseado se eluyó con agua de acuerdo con las instrucciones del suministrador. La disolución acuosa combinada se combinó, se congeló, y se liofilizó, dando como resultado TU3205-052 (23 kDa 2-AAP-mPE) como un sólido de color blanco. (Nota; la caracterización de reactivos PEG de alto PM se obtuvo solo después que se conjugaron con las proteínas que contenían PCL.)

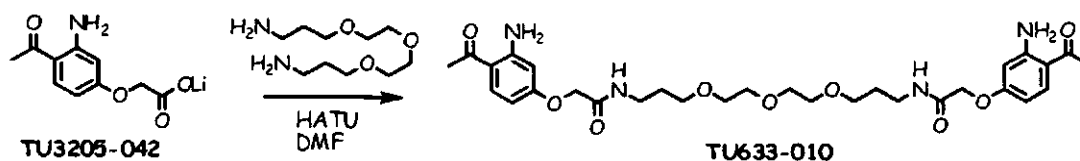


- 15 y



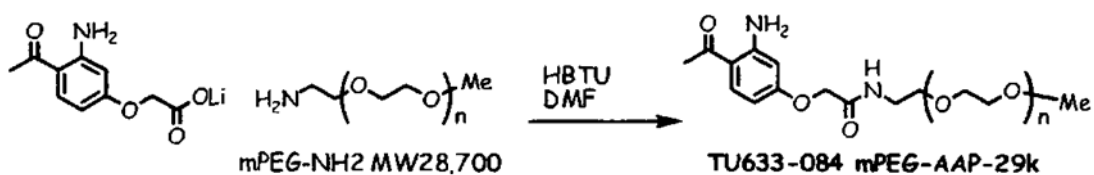
- 20 se prepararon de maneras similares como TU3205-052 usando los correspondientes PM de 10.000 (n promedio = 225) y 5.000 (n promedio = 111) de mPEG-NH₂. (Nota; la caracterización de reactivos PEG de alto PM se obtuvo solo después que se conjugaron con las proteínas que contenían PCL.)

25 Ejemplo 37-4: Síntesis de TU633-010: enlazador bifuncional



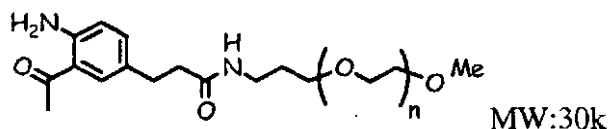
- 30 2-(4-acetil-3-aminofenoxy)acetato de litio (TU3205-042) (94,6 mg) y HATU (167 mg) se cargaron en un matraz de fondo redondo de 10 ml, y se añadió DMF (2 ml). La suspensión resultante se agitó a temperatura ambiente durante 45 minutos. a la disolución amarilla resultante se añadió 4,7,10-trioxa-1,13-tridecanodiamina (Fluka, 44 mg) en 1 ml de DMF. La mezcla de reacción se agitó a temperatura ambiente durante 18 horas, y se concentró a presión reducida. el residuo se purificó mediante cromatografía instantánea en SiO₂ (MeOH en DCM), seguido por una purificación de LC en fase inversa preparativa. El material purificado mediante LC se disolvió en EtOAc y se trató con disolución saturada acuosa de NaHCO₃ para eliminar el ácido trifluoroacético. La evaporación del disolvente dio como resultado el enlazador TU633-010 bifuncional como un aceite transparente. MS (ESI⁺) calc. 603,3, encontrado 603,3 (MH⁺).
- 35

Ejemplo 37-5: Síntesis de m-PEG-AAP 29k (TU633-084).

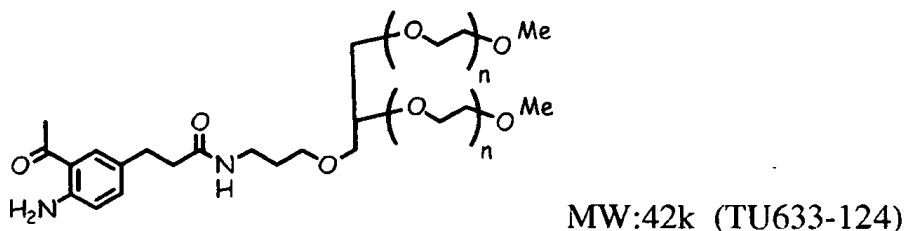
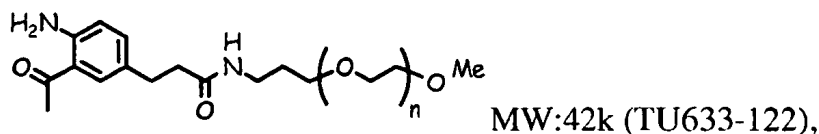


5 2-(4-acetil-3-aminofenoxi)acetato de litio (187 mg) y HBTU (330 mg) se colocaron en un vial de vidrio de 20 ml y se añadieron 10 ml de DMF seco. La suspensión resultante se agitó a temperatura ambiente. En un lapso de tiempo de 20 minutos, la reacción se volvió en una disolución de color amarillo y después de 80 minutos, se añadió una alícuota de 9,5 ml de la mezcla de reacción a mPEG-NH₂ (Laysan Bio, PM 28.700) se disolvió en 40 ml de DMF seco en un matraz de fondo redondo de 100 ml. La reacción se agitó a temperatura ambiente durante 19 horas. A continuación se aplicó la mezcla de reacción a 24 piezas de columnas PD-10 (GE Helthtech) y el producto deseado se eluyó con agua de acuerdo con las instrucciones del suministrador. Los eluyentes combinados se congelaron y liofilizaron, dando como resultado un sólido de color blanco. El sólido se disolvió en agua doblemente desionizada, y se dializó de manera exhaustiva frente al agua doblemente ionizada utilizando una membrana de diálisis MWCO 3500. La disolución dializada se congeló y liofilizó, dando como resultado TU633-084 como un sólido blanco suave y esponjoso. (Nota; la caracterización de reactivos PEG de alto PM se obtuvo solo después que se conjugaron con proteínas que contenían PCL.)

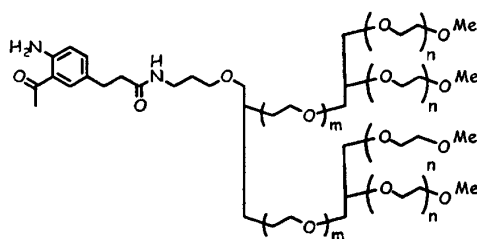
Ejemplo 37-6: Síntesis de mPEG-AAP-30k (TU633-120)



20 3-(3-Acetil-4-aminofenil)propanoato de litio (139 mg) y HBTU (247mg) se colocaron en un vial de vidrio de 20 ml, y se añadieron 13 ml de DMF seco. La suspensión resultante se agitó a temperatura ambiente durante 30 minutos y la disolución resultante se usó para la preparación de TU633-120, TU633-122, TU633-124 y TU633-126. En un vial de vidrio de 45 ml se colocó mPEGamina (NOF Corp., SUNBRIGHT MEPA-30T PM 30.298, 3,0g), y se añadieron 20 ml de DMF seco, seguido por agitación suave para disolver mPEGamina en DMF. Para la disolución de mPEG-amina resultante se añadió una alícuota de 2,2 ml de la disolución del éster activado, y el vial se agitó a temperatura ambiente durante 18 horas y a continuación a 37°C durante 24 horas. Las mezclas de reacción se transfirieron a una tubería de membrana de diálisis (Fisher Scientific, n° de cat. 21-152-9, MWCO 3500) y se dializó exhaustivamente frente a agua doblemente desionizada durante 2 días. La disolución dializada se congeló y liofilizó, dando como resultado TU633-120 como un sólido de tipo algodón de color blanco. (Nota; la caracterización de reactivos PEG de alto PM se obtuvo solo después que se conjugaron con las proteínas que contenían PCL.)



35 y

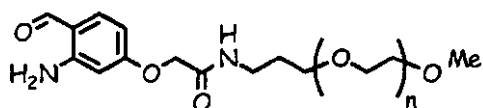


MW:40k (TU633-126)

se prepararon de una manera similar excepto para usar los PEG correspondientes activados procedentes de NOF Corp., SUNBRIGHTMEPA- 40T (PM 42036), SUNBRIGHT GL2-400PA (MW 42348), SUNBRIGHT GL4-400PA, respectivamente, en vez de SUN- BRIGHT MEPA-30T.

(Nota; la caracterización de reactivos PEG de alto PM se obtuvo solo después que se conjugaron con las proteínas que contenían PCL.)

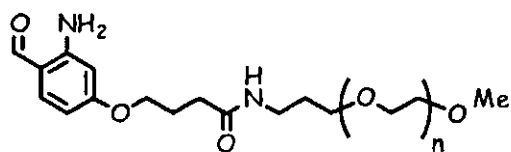
Ejemplo 37-7: Síntesis de mPEG-ABA-30kD (TU3627-024).



MW:30k

se preparó el compuesto del título de la misma manera que TU633-120 excepto por la utilización de 2-(3-amino-4-formilfenoxi)acetato de litio en vez de 3-(3-acetil-4-aminofenil)propanoato de litio. (Nota; la caracterización de reactivos PEG de alto PM se obtuvo solo después que se conjugaron con las proteínas que contenían PCL.)

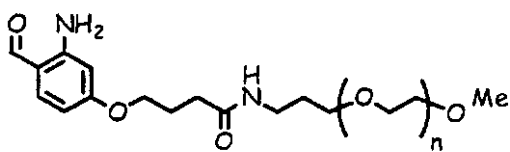
Ejemplo 37-8: Síntesis de mPEG-ABA-30kD (TU3627-084).



MW:30k

se preparó el compuesto del título de la misma manera que TU633-120 excepto por la utilización de 4-(3-amino-4-formilfenoxi)butanoato de litio en vez de 3-(3-acetil-4-aminofenil)propanoato de litio. El análisis de RMN H reveló el PEG de partida sin modificar en el producto. Actividad terminal: >95 % por RMN-H.

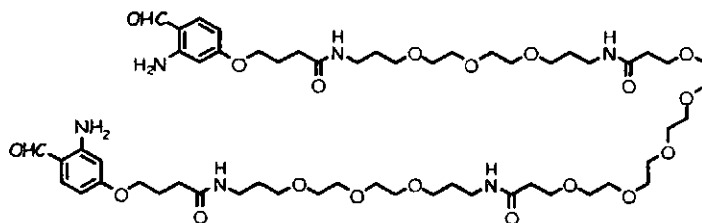
Ejemplo 37-9: Síntesis de mPEG-ABA-40k (TU3627-086).



MW:40k

El compuesto del título se preparó de la misma manera que TU3627-084 excepto por utilizar SUNBRIGHTMEPA-40T en vez de SUNBRIGHTMEPA-30T. El análisis de RMN H reveló el PEG de partida sin modificar en el producto. Actividad terminal: >95 % por RMN-H.

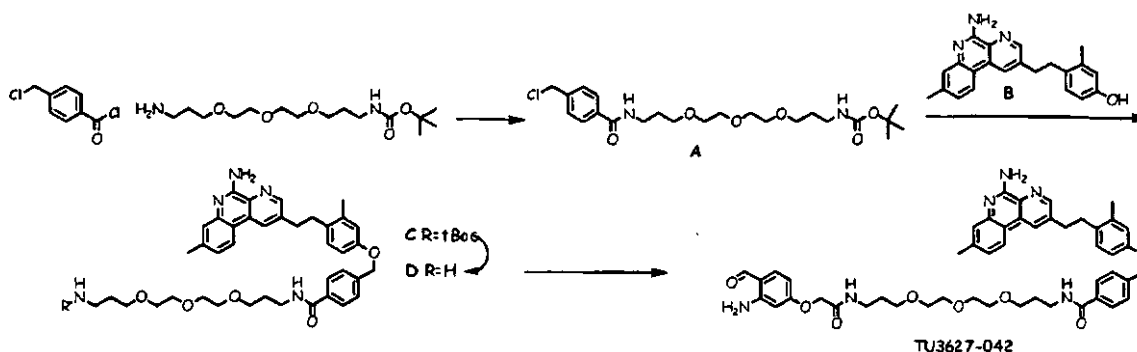
Ejemplo 37-10: Síntesis de *N*¹-(18-(3-Amino-4-formilfenoxi)-15-oxo-4,7,10-trioxa-14-azaoctadecyl)-*N*¹⁹-(18-(3-amino-4-formilfenoxi)-15-oxo-4,7,10-trioxa-9,14-diazaoctadecil)-4,7,10,13,16-pentaoxonadecano-1,19-diamida (X3678-40)



5
4-(3-amino-4-formilfenoxi)butanoato de litio (94,5mg) se activó con HBTU (151,7 mg) en 4 ml de DMF. Diamido-dPEGTM₁₁-diamina (QuantaBiodesign, n° de cat.10361, 74,3 mg) se añadió a la mezcla de reacción se agitó a temperatura durante toda la noche. DIEA (17,5 ml, 0,1 mmol) se añadió y la mezcla se calentó a 40 °C durante varias horas. Se aisló el compuesto del título mediante RP-HPLC preparativa y se eliminó TFA haciéndolo pasar a un cartucho MP SPE de PL-HCO₃ (Varian Inc.) MS (ESI+) m/z 1153,60 (MH+).

Ejemplo 38: Síntesis de reactivo para acoplamiento de inmunomoduladores con proteínas

Ejemplo 38-1: Síntesis de *N*-(1-(3-Amino-4-formilfenoxi)-2-oxo-7,10,13-trioxa-3-azahexadecan-16-il)-4-((4-(2-(5-amino-8-metilbenzo[f][1,7]naftiridin-2-il)etil)-3-metilfenoxi)metil)benzamida (TU3627-042).



20
Para el 3-(2-(2-(3-aminopropoxi)etoxi)etoxi)propilcarbamato de t-butilo (QuantaBiodesign, n° de cat. 10225, 500 mg) y DIEA (348 ml) in 10 ml y se añadió DCM a cloruro de 4-clorometilbenzoilo en 2 ml de DCM con enfriamiento en un baño de hielo. La reacción se agitó a temperatura ambiente durante 1 hora. La mezcla de reacción se diluyó con EtOAc, se lavó sucesivamente con H₂O, ácido cítrico acuoso diluido, disolución acuosa de NaHCO₃ y disolución acuosa saturada de NaCl, se secó sobre Na₂SO₄, se filtró y se concentró a presión reducida, dando como resultado el producto (A) como un aceite viscoso transparente. MS (ESI+): calc. 473,2, encontrado 473,3 (MH+). H-RMN (400MHz, CDCl₃): 1,405 (9H, s), 1,685 (2H, m), 1,879 (2H, m), 3,173 (2H, m), 3,438 (4H, m), 3,567 (4H, m), 3,633 (6H, m), 4,584 (2H, s), 4,933 (1H, br.s), 7,246 (1H, br.s), 7,411 (2H, d, J = 8,4Hz), 7,789 (2H, d, J = 8,4Hz). C-RMN (100 MHz, CDCl₃): 28,385, 28,673, 29,553, 38,430, 39,087, 45,447, 69,474, 70,042, 70,216, 70,379, 70,516, 70,798, 78,882, 127,430, 128,496, 134,745, 140,341, 155,966, 166,550.

25
El producto A, procedente de la anterior (303 mg), el compuesto B (200 mg), carbonato de cesio (208 mg) en DMSO anhidro se combinaron y la reacción se agitó a temperatura ambiente. Después de 18 horas, 60 mg más del compuesto A se disolvieron en 8 ml de DMSO y 30 mg más de carbonato de cesio se añadieron a la reacción y la reacción se agitó durante 40 horas a temperatura ambiente, en cuyo momento solo quedó una cantidad traza del compuesto B, indicada por el análisis de LCMS. La mezcla de reacción se diluyó con EtOAc, se lavó sucesivamente con H₂O, disolución saturada acuosa de NaCl, se secó con Na₂SO₄, se filtró y se concentró a presión reducida. El residuo se purificó mediante cromatografía instantánea en gel de sílice usando un gradiente lineal de 1 a 5 % de disolvente B en disolvente A (A: DMC; B: NH₃ al 2 % en MeOH), dando como resultado el compuesto deseado parcialmente purificado. recristalizado después a partir de MTBE, se obtuvo el producto como cristales ligeramente amarillos. MS (ESI+): calc. 780,43, encontrado 780,50 (MH+). H-RMN (400MHz, DMSO-d₆): 1,357 (9H, s), 1,578 (2H, quint, J = 6,8Hz), 1,751 (2H, quint, J = 6,8Hz), 2,264 (3H, s), 2,439 (3H, s), 2,94 (4H, m), 3,08 (2H, m), 3,34 (4H, m), 3,45 (4H, m), 3,50 (6H, m), 5,105 (2H, s), 6,757 (1H, dd, J = 2,4, 8,4Hz), 6,77 (1H, br.s), 6,836 (1H, d, J = 2,8Hz), 7,060 (2H, br.s), 7,091 (1H, d, J = 8,4Hz), 7,138 (1H, dd, J = 2,4, 8,4Hz), 7,349 (1H, s), 7,495 (2H, d, J = 8,4Hz), 7,837 (2H, d, J = 8,4Hz), 8,330 (1H, d, J = 8,4Hz), 8,439 (1H, t, J = 5,6Hz), 8,701 (1H, d, J = 2,0Hz), 8,827 (1H, d, J = 1,6Hz). C-RMN

(100 MHz, DMSO- d₆): 19,153, 21,206, 28,178, 29,297, 29,638, 33,166, 33,336, 36,584, 37,085, 67,989, 68,200, 68,455, 69,495, 69,678, 69,703, 77,307, 111,863, 116,456, 117,005, 122,743, 123,321, 125,292, 127,047, 127,178, 127,764, 129,725, 129,823, 131,414, 132,600, 133,883, 137,040, 138,775, 139,426, 140,303, 144,984, 149,405, 155,471, 155,663, 156,341, 165,783.

5 El grupo Boc del compuesto C se eliminó mediante tratamiento con HCl metanólico 3M a temperatura ambiente. La concentración de la mezcla de reacción a presión reducida dio como resultado el compuesto D como sal de diclorhidrato. MS (ESI+): calc. 680,4, encontrado 680,4 (MH⁺). H-RMN (400MHz, DMSO-d₆): 1,77 (4H, m), 2,289 (3H, s), 2,504 (solapándose con la señal de DMSO-d₆), 2,85 (2H, m), 2,98 (2H, m), 3,14 (2H, m), 3,31 (2H, m), 3,5 (12H, m), 5,109 (2H, s), 6,768 (1H, dd, J = 2,4, 8,4Hz), 6,852 (1H, d, J = 2,4Hz), 7,107 (1H, d, J = 2,4Hz), 7,427 (1H, d, J = 8,4Hz), 7,496 (2H, d, 8,4Hz), 7,532 (1H, s), 7,858 (2H, d, J = 8,0Hz), 7,89 (3H, br.s), 8,53 (2H, m), 8,872 (1H, s), 9,01 (1H, br.s), 9,03 (1H, s), 9,731 (1H, s). C-RMN (100 MHz, DMSO-d₆): 19,183, 21,127, 27,076, 29,325, 32,793, 33,290, 36,601, 67,251, 68,198, 68,423, 69,389, 69,480, 69,587, 69,693, 111,903, 115,654, 116,491, 117,597, 124,065, 126,488, 127,058, 127,227, 129,646, 129,818, 130,877, 131,129, 131,841, 132,733, 133,850, 137,112, 140,296, 141,702, 143,879, 151,265, 153,598, 156,424, 165,789.

20 2-(3-amino-4-formilfenoxi)acetato de litio (20 mg) y HBTU (38 mg) se colocaron en un vial de vidrio de 20 ml, y se añadieron 2 ml de DMF seco. La suspensión resultante se agitó a temperatura ambiente durante 30 minutos, y a continuación se añadieron el compuesto D (38 mg) y DIEA (52 ml). La mezcla de reacción se agitó a temperatura ambiente durante 17 horas. El análisis de LCMS no reveló que permanecieran materiales de partida sino la formación de un subproducto bis-acilado junto con el producto deseado. La mezcla de reacción se diluyó con EtOAc, se lavó sucesivamente con H₂O y disolución saturada acuosa de NaCl, se secó con Na₂SO₄, se filtró y se concentró a presión reducida. el residuo se capturó en 5 ml de THF, y se trató con una alícuota de 1 ml de una disolución acuosa de LiOH 1 M a temperatura ambiente durante 2 horas, en cuyo momento, el análisis de LC-MS mostró una hidrólisis del producto bis-acilado con el producto deseado. La mezcla de reacción se repartió entre EtOAc y agua, y se separó la capa orgánica, se lavó con disolución acuosa saturada de NaCl, se secó con Na₂SO₄, se filtró y se concentró a presión reducida. El material bruto se aplicó a una RP-HPLC preparativa para la purificación. El eluyente de HPLC que contenía el producto deseado se diluyó con EtOAc, y se lavó con disolución saturada acuosa de NaHCO₃ y disolución saturada acuosa de NaCl para eliminar el ácido trifluoroacético, se secó con Na₂SO₄, se filtró y se concentró a presión reducida, dando como resultado el compuesto del título N-(1-(3-Amino-4-formilfenoxi)-2-oxo-7,10,13-trioxa-3-azahexadecan-16-il)-4-((4-(2-(5-amino-8-metilbenzo[f][1,7]naftiridin-2-il)etil)-3-metilfenoxi)metil)benzamida (TU3627-042). MS (ESI+): m/z 857,40 (MH⁺). H-RMN (400MHz, DMSO-d₆): 1,651 (2H, quint, J = 6,6Hz), 1,747 (2H, quint, J = 6,6Hz), 2,264 (3H, s), 2,443 (3H, s), 2,96 (2H, m), 3,08 (2H, m), 3,16 (3H, m), 3,27-3,29 (m solapando con la señal H₂O), 3,45 (4H, m), 3,50 (7H, m), 4,451 (2H, s), 5,102 (2H, s), 6,198 (1H, d, J = 2,0Hz), 6,268 (1H, dd, J = 2,0, 8,8Hz), 6,755 (1H, 2,4, 8,4Hz), 6,836 (1H, d, J-2,4Hz), 7,095 (1H, d, J = 8,4Hz), 7,158 (1H, d, J = 8,0Hz), 7,208 (2H, br.s), 7,362 (1H, s), 7,435 91H, d, J = 8,8Hz), 7,493 (2H, d, J = 8,4Hz), 7,834 (2H, d, J = 8,0Hz), 8,070 (1H, t, J = 5,6Hz), 8,349 (1H, d, J = 8,4Hz), 8,440 (t, J = 5,6Hz), 8,711 (1H, d, J = 1,6Hz), 8,840 (1H, d, J = 1,6Hz), 9,636 (1H, s).

40 Síntesis de 4-(2-(5-amino-8-metilbenzo[f][1,7]naftiridin-2-il)etil)-3-metilfenol (compuesto B)

45 a un matraz de fondo redondo tapado con un septo se añadió 1-etinil-4-metoxi-2-metilbenceno (comercialmente disponible) (1,1 eq), 3,5-dicloropicolinonitrilo (1 eq.), trietilamina (5 eq), y DMF anhidro (0,2 M). Se hizo el vacío y se lavó con nitrógeno tres veces. se añadieron CuI (0,05 eq.) y bis(trifenilfosfina)dicloro-paladio(II) (0,05 eq). El septo se sustituyó con un condensador a reflujo y el matraz se calentó a 60 °C durante la noche en atmósfera de nitrógeno. Hasta la finalización de la reacción que se vigiló mediante TLC, el contenido del matraz se introdujo en una columna de gel de sílice grande pretratada con hexanos. La cromatografía instantánea (gel de sílice, hexanos:EtOAc (1:4 %)) dio como resultado el producto 3-cloro-5-((4-metoxi-2-metilfenil)etnil)picolinonitrilo.

50 A un matraz de fondo redondo con condensador a reflujo se añadieron 3-cloro-5-((4-metoxi-2-metilfenil)etnil)picolinonitrilo (procedente de la etapa anterior) (1 eq.), 2-(4,4,5,5-tetrametil-1,3,2-dioxaborolan-2-il)fenilcarbamato de terc-butilo (1,25 eq.), K₃PO₄ (2 eq.), tris(dibencilidenoacetona)dipaladio(0) (0,05 eq.), y se añadieron 2-diciclohexilfosfino- 2',6'-dimetoxibifenilo (0,1 eq.). n-Butanol y agua (5:2, 0,2 M), y el contenido se desgasificó (vacío seguido por lavado con nitrógeno) durante tres veces. La mezcla de reacción se agitó vigorosamente con nitrógeno a 100 °C durante la noche en un baño de aceite. Los contenidos se enfriaron y se capturaron en 200 ml de agua seguido por la extracción con cloruro de metileno. Las capas orgánicas combinadas se secaron (Na₂SO₄) y se concentraron. La cromatografía instantánea (gel de sílice, 0 - 50 % de EtOAc en CH₂Cl₂) dio como resultado el producto 2-((4-metoxi-2-metilfenil)etnil)-8-metilbenzo[f][1,7]naftiridin-5-amina.

60 A un matraz de fondo redondo se añadió 2-((4-metoxi-2-metilfenil)etnil)-8-metilbenzo[f][1,7]naftiridin- 5-amina (procedente de la etapa anterior) (1 eq.) con una barra de agitación. Se añadieron etanol y cloruro de metileno (1:2, 0,2 M), seguido por paladio en carbono

65 (polvo activado, húmedo, 10 % sobre carbono, 0,1 eq). El contenido se sometió a vacío seguido por lavado con hidrógeno durante tres veces. La mezcla de reacción se agitó vigorosamente con un balón de hidrógeno a temperatura

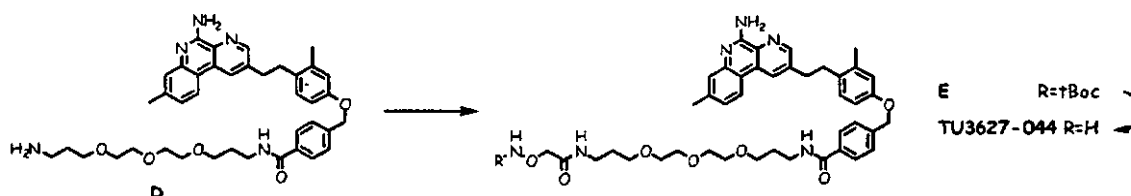
ambiente durante la noche. Después de lo cual la mezcla de reacción se filtró a través de un parche de celite, y el parche de celite se lavó posteriormente con cloruro de metileno y EtOAc hasta que el filtrado no mostró absorción UV. Los lavados orgánicos combinados se concentraron. La cromatografía instantánea (gel de sílice, 0 - 50 % EtOAc en CH₂Cl₂) dio como resultado el producto

5 2-(4-Metox-2-metilfenetil)-8-metilbenzo[f][1,7]naftiridin-5-amina.hilfenetil)-8-metilbenzo[f][1,7]naftiridin-5-amina. RMN ¹H (CDCl₃): 8,8,53 (d, 1H), 8,29 (d, 1H), 8,01 (d, 1H), 7,44 (s, 1H), 7,12 (dd, 1H), 6,93 (d, 1H), 6,67 (d, 1H), 6,60 (dd, 1H), 5,93, (bs, 2H), 3,70 (s, 3H), 3,05 - 3,00 (dd, 2H), 2,93 - 2,88 (dd, 2H), 2,44 (s, 3H), 2,19 (s, 3H). LRMS [M+H] = 358,2

10 A una disolución agitada de 2-(4-metoxi-2-metilfenetil)-8-metilbenzo[f][1,7]naftiridin-5-amina en cloruro de metileno (0,2 M) en un baño de agua helada se añadió una disolución 1 N de BBr₃ (2 eq) en CH₂Cl₂ de una manera gota a gota. En 30 minutos, la reacción se inactivó rápidamente y se concentró a vacío para obtener un residuo bruto. El material bruto se purificó mediante cromatografía instantánea en un sistema COMBIFLASH® (ISCO) utilizando 0-20 % de metanol en diclorometano para dar el 4-(2-(5-amino-8-metilbenzo[f][1,7]naftiridin-2-il)etil)-3-metilfenoxi)metil)-N-(1-(aminooxi)-2-oxo-7,10,13-trioxa-3-az

15 de color blanco. RMN ¹H (DMSO-d₆): δ 8,99 (s, 1H), 8,75 (d, 1H), 8,60 (d, 1H), 8,27 (d, 1H), 7,28 (s, 1H), 7,09 (dd, 1H), 6,99, (bs, 2H), 6,88 (d, 1H), 6,49 (d, 1H), 6,42 (dd, 1H), 3,02 - 2,96 (dd, 2H), 2,86 - 2,81 (dd, 2H), 2,38 (s, 3H), 2,13 (s, 3H). LRMS [M+H] = 344,2

20 *Ejemplo 38-2: Síntesis del clorhidrato de 4-((4-(2-(5-Amino-8-metilbenzo[f][1,7]naftiridin-2-il)etil)-3-metilfenoxi)metil)-N-(1-(aminooxi)-2-oxo-7,10,13-trioxa-3-azahexadecan-16-il)benzamida (TU3627-044).*

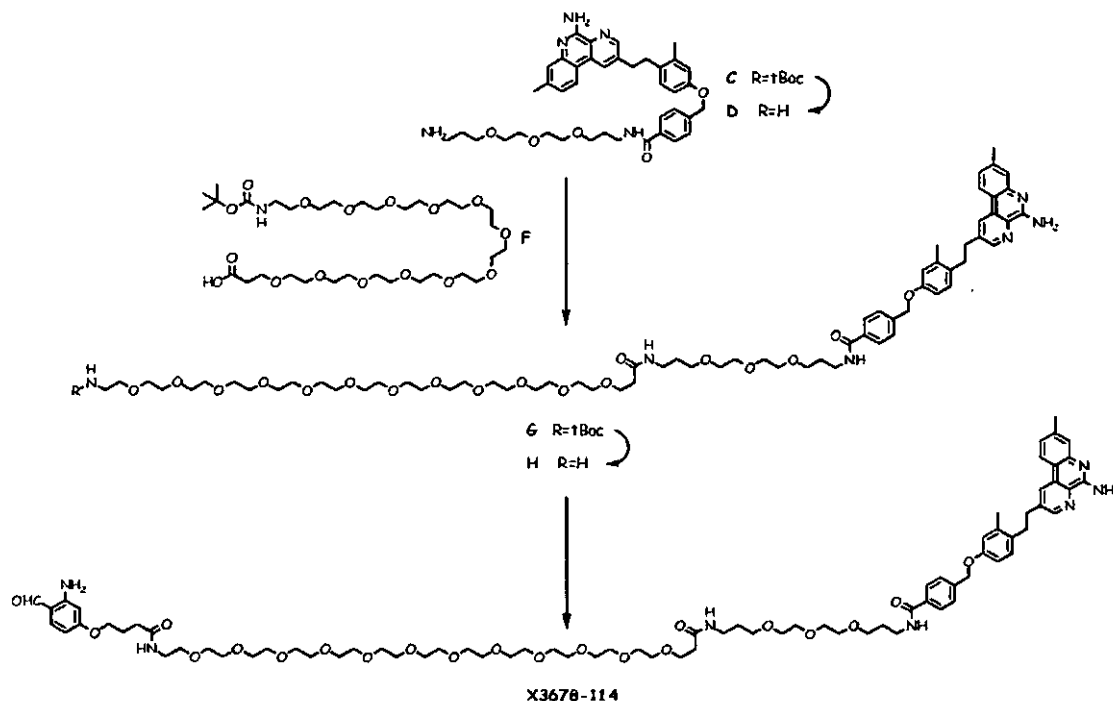


25 HBTU (38 mg), Ácido 2-(terc-Butoxicarbonilaminooxi)acético (Fluka, 25mg), DIEA (44 ml) y 5ml de DMF se combinaron en un vial de vidrio de 20 ml. La reacción se agitó a temperatura ambiente durante 30 minutos, y a continuación se añadió el compuesto D (38mg). La mezcla de reacción se agitó a temperatura ambiente durante 17 horas. El análisis de LCMS no reveló que permanecieran materiales de partida sino la formación de subproductos sobreacilados junto con el producto deseado. La mezcla de reacción se diluyó con EtOAc, se lavó sucesivamente con H₂O y disolución saturada acuosa de NaCl, se secó con Na₂SO₄, se filtró y se concentró a presión reducida. el residuo se capturó en 5 ml de THF, y se trató con una alícuota de 1 ml de una disolución acuosa de LiOH 1 M a temperatura ambiente durante 2 horas, en cuyo momento, el análisis de LC-MS mostró una hidrólisis de los productos sobreacilados con el producto deseado. La mezcla de reacción se repartió entre EtOAc y agua, y se separó la capa orgánica, se lavó con disolución acuosa saturada de NaCl, se secó con Na₂SO₄, se filtró y se concentró a presión

30 reducida. El material bruto se purificó mediante cromatografía instantánea en gel de sílice usando un 5 % de disolvente B en disolvente A (A: DMC; B: NH₃OH al 2 % en MeOH), dando como resultado el compuesto deseado, E. MS (ESI+): calc. 853,44, encontrado 853,50 (MH⁺). H-RMN (400MHz, DMSO-d₆): 1,401 (9H, s), 1,643 (2H, quint. J = 6,8Hz), 1,750 (2H, quint. J = 6,8Hz), 2,269 (3H, s), 2,453 (3H, s), 2,96 (2H, m), 3,09 (2H, m), 3,162 (2H, c, J = 6,0Hz), 3,298 (solapándose con la señal de H₂O), 3,393 (2H, t, J = 5,6Hz), 3,45 (4H, m), 3,51 (6H, m), 4,219 (2H, s), 5,106 (2H, s), 6,758 (1H, dd, J = 2,4, 8,4Hz), 6,840 (1H, d, J = 2,8Hz), 7,094 (1H, d, J = 8,4Hz), 7,102 (1H, br.d, J = 8,0Hz), 7,394 (1H, s), 7,495 (2H, d, J = 8,4Hz), 7,837 (2H, d, J = 8,4Hz), 7,994 91H, br.t, 5,2Hz), 8,379 (1H, d, 8,4Hz), 8,441 (1H, t, 5,4Hz), 8,737 (1H, s), 8,869 (1H, s), 10,307 (1H, br.s).

45 El compuesto E (28 mg) se trató con HCL metanólico 3M a temperatura ambiente durante 30 minutos, y se concentró a presión reducida. Esta operación se repitió. El compuesto del título clorhidrato de 4-((4-(2-(5-Amino-8-metilbenzo[f][1,7]naftiridin-2-il)etil)-3-metilfenoxi)metil)-N-(1-(aminooxi)-2-oxo-7,10,13-trioxa-3-aza hexadecan-16-il)benzamida (TU3627-044). MS (ESI+): m/z 753,40 (MH⁺). H-RMN (400MHz, DMSO-d₆): 1,651 (2H, quint, J = 6,6Hz), 1,752 (2H, quint, J = 6,6Hz), 2,290 (3H, s), 2,5 (un pico de metilo solapándose con la señal DMSO), 2,98 (2H, m), 3,15 (4H, m), 3,155 (2H, c, J = 7,2Hz), 3,47 (12H, m), 4,463 (2H, s), 5,109 (2H, s), 6,767 (1H, dd, J = 2,8, 8,4Hz), 6,854 (1H, d, J = 2,8Hz), 7,108 (1H, d, 8,4Hz), 7,431 (1H, d, J = 8,4Hz), 7,498 (2H, d, J = 8,4Hz), 7,545 (1H, s), 7,848 (2H, d, J = 8,4Hz), 8,240 (1H, t, 5,6Hz), 8,467 (1H, t, J = 5,6Hz), 8,542 (1H, d, J = 8,4Hz), 8,875 (1H, 2d, J = 2,0Hz), 8,993 (1H, br.s), 9,027 (1H, d, J = 2,0Hz), 9,729 (1H, br.s). C-RMN (100 MHz, DMSO-d₆): 19,177, 21,141, 29,034, 29,322, 32,851, 33,345, 35,615, 36,581, 67,843, 68,199, 68,426, 69,470, 69,503, 69,685, 69,713, 71,237, 111,909, 115,691, 116,495, 117,691, 124,022, 126,495, 127,069, 127,212, 129,652, 129,824, 130,850, 131,127, 131,872, 132,820, 133,883, 137,116, 140,285, 141,689, 143,895, 151,303, 153,491, 156,438, 165,784, 166,659.

Ejemplo 38-3: - Síntesis de *N*-(58-(3-amino-4-formilfenoxi)-15,55-dioxo4,7,10,18,21,24,27,30,33,36,39,42,45,48,51-pentadecaoca-14,54-diazaoctapentacontil)-4-((4-(2-(5-amino-8-metil-benzo[*f*][1,7]naftiridin-2-il)etil)-3-metilfenoxi)metil)benzamida (X3678-114).

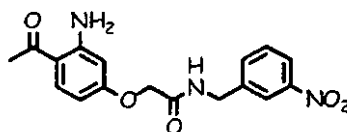


5

Una mezcla de dPEG-ácido **F** (100 mg), HATU (52,8 mg) y DIEA (97 ml) en 2,0 ml de DMF se agitó a temperatura ambiente durante 30 minutos. Se añadió amina **D** (preparada a partir de 109 mg de compuesto **C**) y la reacción se agitó a temperatura ambiente hasta el agotamiento tal como se vigiló mediante HPLC. El producto se aisló mediante RP-HPLC preparativa. El grupo Boc del producto **G** (138 mg) procedente de la etapa anterior se eliminó mediante el tratamiento de HCl 3 N en metanol, seguido por concentración hasta sequedad. Una mezcla de 4-(3-amino-4-formilfenoxi)butanoato de litio (25,2 mg), HATU (38,0 mg) y DIEA (69,7 ml) en 2,0 ml de DMF se agitó a temperatura ambiente durante 30 minutos. Se añadió amina **H** y la reacción se agitó a temperatura ambiente hasta el agotamiento vigilada mediante HPLC. El producto (**X3678-114**) se aisló mediante HPLC preparativa. MS (ESI+): m/z 743,00 ($MH_2^{2+}/2$).

15

Ejemplo 38-4: Síntesis of 2-(4-Acetil-3-aminofenoxi)-N-(3-nitrobenzil)acetamida (TU633-068).

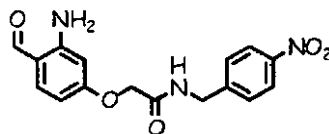


20

2-(4-acetil-3-aminofenoxi)acetato de litio (215 mg), HBTU (379mg), y 10 ml de DMF se combinaron en un vial de vidrio de 20 ml y la suspensión resultante se agitó a temperatura ambiente durante 1 hora momento en el cual la mezcla de reacción se volvió casi homogénea. A continuación se añadieron al vial HCl de 3'-nitrobenzilamina (192 mg) y DIEA (191 ml), y la reacción se agitó a temperatura ambiente durante 30 minutos. La mezcla de reacción se diluyó con EtOAc, se lavó sucesivamente con agua, disolución saturada acuosa NaCl, se secó con Na_2SO_4 , se filtró y se concentró a presión reducida. El material bruto se purificó mediante cromatografía instantánea en gel de sílice (hexanos/EtOAc), para dar el compuesto del título como un sólido de color amarillo. MS (ESI+): m/z 344,10 (MH^+). H-RMN (400MHz, $CDCl_3$): 2,526 (3H, s), 4,587 (2H, s), 4,641 (2H, d, $J = 6,4Hz$), 6,091 (1H, d, $J = 2,8Hz$), 6,239 (1H, dd, $J = 2,8Hz, 8,8Hz$), 6,969 (1H, br.s), 7,511 (1H, t, $J = 7,6Hz$), 7,624 (1H, d, $J = 7,6Hz$), 7,678 (1H, d, $J = 8,8Hz$), 7,849 (1H, m), 8,131 (1H, s), 8,141 (1H, d, $J = 7,6Hz$).

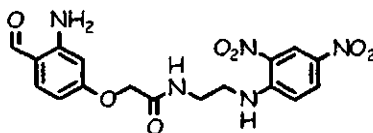
30

Ejemplo 38-5: Síntesis de 2-(3-Amino-4-formilfenoxi)-N-(4-nitrobenzil)acetamida (TU3627-020).



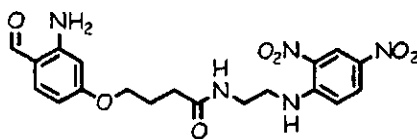
5 2-(3-amino-4-formilfenoxi)acetato de litio (44,2mg) y HBTU (79,6 mg) se colocaron en un vial de vidrio de 20 ml y se
añadió DMF seco (5 ml). La suspensión resultante se agitó a temperatura ambiente, que se volvió homogénea en un
lapso de 10 minutos. Después 15 min, DIEA (50 ml) y HCl de 4'-nitrobenzilamina HCl (37,7mg) se añadieron a la
mezcla de reacción, y la reacción se agitó a temperatura ambiente. Después de 30 minutos, el análisis de LC-MS
10 reveló una reacción completa. La mezcla de reacción se repartió entre EtOAc y agua. Se separó la capa orgánica, se
lavó sucesivamente con ácido cítrico ac. dil., disolución saturada acuosa de NaHCO₃, y disolución saturada acuosa
NaCl, se secó con Na₂SO₄, se filtró y se concentró a presión reducida, para dar el compuesto del título como un sólido
de color crema. MS (ESI+): m/z 330,10 (MH⁺). H-RMN (400MHz, DMSO-d₆): 4,463 (2H, d, J = 6,0Hz), 4,597 (2H, s),
6,246 (1H, d, J= 2,4hz), 6,308 (1H, dd, J = 2,4Hz, 8,8Hz), 7,235 (1H, br.s), 7,465 (1H, d, 8,8Hz), 7,518 (2H, d, J =
8,8Hz), 8,173 (2H, d, J = 8,8Hz), 8,821 (1H, t, J = 6,2Hz), 9,657 (1H, s). C-RMN (400MHz, DMSO-d₆): 41,415, 66,627,
15 98,724, 104,297, 113,217, 123,267, 128,056, 137,581, 146,303, 147,287, 152,698, 162,872, 167,526, 191,685.

Ejemplo 38-6: Síntesis de 2-(3-Amino-4-formilfenoxi)-N-(2-(2,4-dinitrofenilamino)etil)acetamida (TU3627-022).

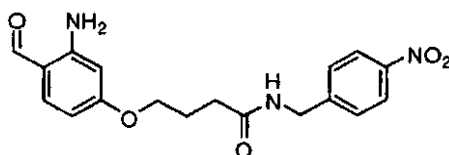


20 El compuesto del título se preparó de la misma manera que **TU3627-020** excepto que
N¹-(2,4-dinitrofenil)etano-1,2-diamina (Oakwood, n° de cat. 015083, 45,2 mg) se usó en vez del HCl de
4'-nitrobenzilamina en ausencia de DIEA. **TU3627-022** se obtuvo como un sólido de color amarillo. Rf. 0,15 (SiO₂, 5 %
de MeOH en DCM). MS (ESI+): m/z 404,10 (MH⁺). H-RMN (400MHz, DMSO-d₆): 3,420 (2H, m), 3,593 (2H, m), 4,467
25 (2H, s), 6,181 (1H, d, 2,0Hz). 9,248 (1H, dd, J = 2,4, 8,8Hz), 7,173 (2H, br.s), 7,302 (1H, d, J = 9,6Hz), 7,406 (1H, d, J=
8,4hz), 8,230 (1H, dd, J = 2,4, 9,6Hz), 8,370 (1H, t, J = 6,0Hz), 8,841 (1H, d, J = 2,8Hz), 8,924 (1H, t, J = 5,6Hz), 9,618
(1H, s).

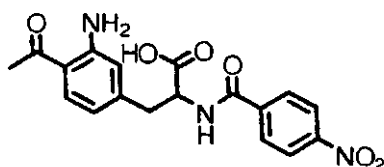
Ejemplo 38-7: Síntesis de 4-(3-Amino-4-formilfenoxi)-N-(2-(2,4-dinitrofenilamino)etil)butanamida (TU3627-088).



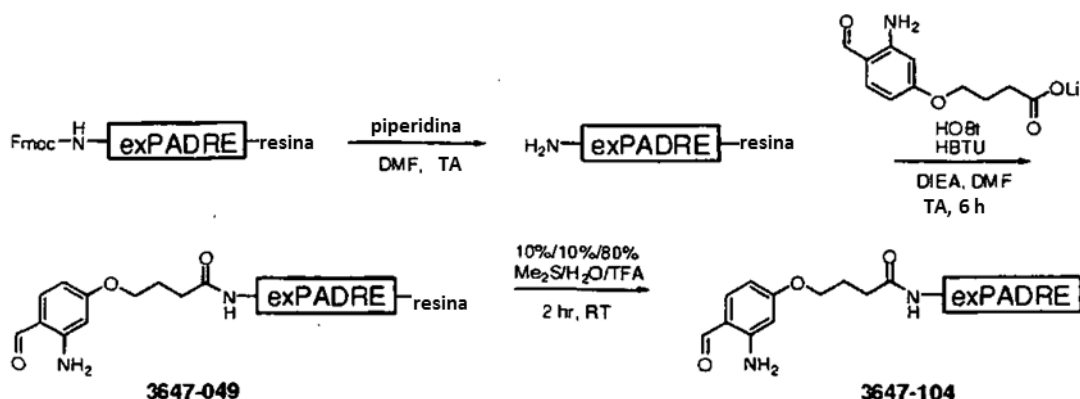
35 4-(3-amino-4-formilfenoxi)butanoato de litio (50 mg), HBTU (80 mg) y DMF (2 ml) se combinaron en un vial de vidrio de
20 ml y se agitaron a temperatura ambiente. Después de 30 minutos, N¹-(2,4-dinitrofenil)etano-1,2-diamina (Oakwood,
45 mg) se añadió en una porción y la reacción se agitó a temperatura ambiente durante la noche. La mezcla de
reacción se diluyó con EtOAc, se lavó sucesivamente con ácido cítrico ac. dil., agua, disolución saturada acuosa de
NaHCO₃, agua, y disolución saturada acuosa de NaCl, se secó con Na₂SO₄, se filtró y se concentró a presión reducida,
para dar el compuesto del título como un sólido de color amarillo. Rf: 0,18 (SiO₂, EtOAc). MS (ESI+): m/z 432,20
(MH⁺). H-RMN (400MHz, DMSO-d₆): 1,914 (2H, quint, J = 7,0Hz), 2,224 (2H, t, J = 7,0Hz), 3,349 (2H, c, J = 6,0Hz),
40 3,538 (2H, c, J = 6,0Hz), 3,928 (2H, t, J = 6,4Hz), 6,165 (1H, s), 6,176 (1H, dd, J = 2,4, 10,0Hz), 7,133 (1H, br.s), 7,268
(1H, d, J = 10,0Hz), 7,371 (1H, d, J = 8,4Hz), 8,167 (1H, t, J = 5,8Hz), 8,250 (1H, dd, J = 2,8, 8,8Hz), 8,835 (1H, d, J =
2,8Hz), 8,913 (1H, t, J = 5,6Hz), 9,599 (1H,s). C-RMN (100 MHz, DMSO- d₆): 24,489, 31,436, 37,331, 42,745, 66,782,
98,040, 104,339, 112,751, 115,130, 123,472, 123,564, 129,725, 134,731, 137,484, 148,305, 152,824, 163,873,
45 172,251, 191,400.

Ejemplo 38-8: Síntesis de 4-(3-Amino-4-formilfenoxi)-N-(4-nitrobenzil)butanamida (3793-001).

5 HATU (114,1 mg) se añadió a 1 ml de disolución DMF de 4-(3-amino-4-formilfenoxi)-butanoato de litio (68,7 mg), y la reacción se agitó a temperatura ambiente durante 1 hora. La disolución resultante se añadió a continuación a 1 ml de disolución de DMF de sal de HCl de 4-nitrobenzilamina (56,6 mg) y trietilamina (84 ml), con otro ml de DMF para ayudar a la transferencia. La reacción se agitó a temperatura ambiente durante 2 horas. Hasta la finalización, la mezcla de reacción se repartió entre 4 ml de NaCl(aq) 10 % y 8 ml de EtOAc. Las fases se separaron y la capa acuosa se extrajo de nuevo con 8 ml de EtOAc. Las capas orgánicas combinadas se secaron con Na₂SO₄, se filtraron y se concentraron a presión reducida, dando como resultado un aceite bruto de color naranja. El aceite bruto se purificó mediante cromatografía instantánea en columna de gel de sílice con gradiente de elución de 0-10 % de MeOH/DCM para dar el producto deseado como un sólido de color amarillo claro. MS (ESI⁺): m/z 358,2. RMN ¹H (400 MHz, CDCl₃): 9,70 (s, 1H), 8,14 (d, J = 8,8 Hz, 2H), 7,40 (d, J = 8,8 Hz, 2H), 7,35 (d, J = 8,4 Hz, 1H), 6,26 (dd, J = 2,2 Hz, 8,8 Hz, 1H), 6,20 (br, 2H), 6,03 (d con br, J = 2,0 Hz, 2H), 4,54 (d, J = 6,0 Hz, 2H), 4,02 (t, J = 5,8 Hz, 2H), 2,47 (t, J = 7,0 Hz, 2H), 2,17 (quintuplete, J = 6,5 Hz, 2H). RMN ¹³C (100 MHz, CDCl₃): 192,0, 172,2, 164,4, 152,2, 147,2, 145,8, 137,8, 128,2, 123,9, 113,9, 105,3, 98,9, 66,7, 42,8, 32,5, 24,8.

Ejemplo 38-9: Síntesis de ácido 3-(4-acetil-3-aminofenil)-2-(4-nitrobenzamido)propanoico (X3471-116).

20 A ácido 4-nitrobenzoico (50,1 mg), HATU (114 mg) y DIEA (105 ml) se añadió 1 ml de DMF y la mezcla se agitó durante 30 minutos. A continuación, la disolución resultante se añadió al ácido 3-(4-acetil-3-aminofenil)-2-aminopropanoico (66,7 mg, 1,0 equiv) en 1,0 ml de DMF, y la reacción se agitó a temperatura ambiente, se vigiló mediante LC-MS. A continuación, el producto se aisló mediante HPLC preparativa. RMN ¹H (400 MHz, MeOD) : 8,28 (d, J = 8,8 Hz, 2H), 7,93 (d, J = 8,8 Hz, 2H), 7,72 (d, J = 8,4 Hz, 1H), 6,69 (d, J = 1,6 Hz, 1H), 6,61 (dd, J = 8,4 Hz, J'' = 1,6 Hz, 1H), 3,28 (dd, J = 13,8 Hz, J'' = 10,0 Hz, 1H), 3,02 (dd, J = 13,8 Hz, J'' = 10,0 Hz, 1H), 2,51 (s, 3H); ESI-MS (m/z) 372,34 (MH⁺).

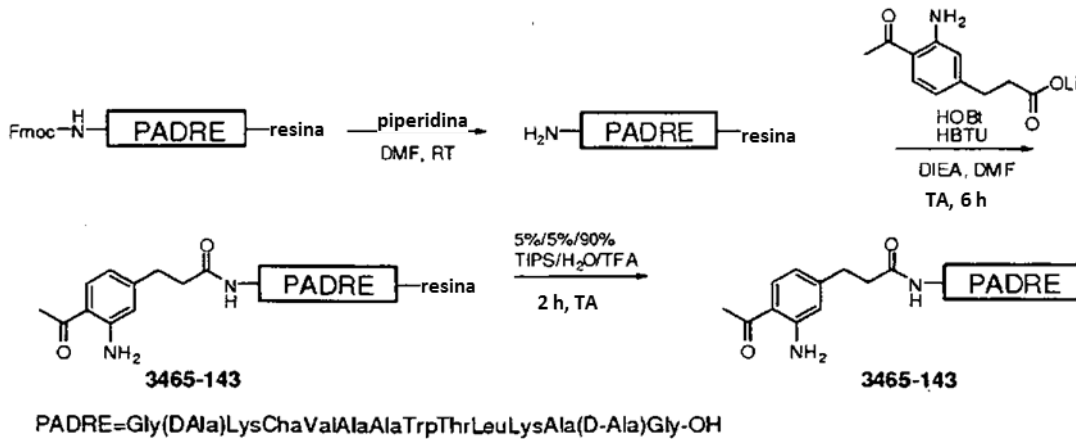
Ejemplo 38-10: Síntesis de 4-(3-Amino-4-Formilfenoxi)butanoil-AlaGlySerArgSerGly(D-Ala)LysChaValAlaAlaTrp-ThrLeuLysAla(D-Ala)Gly-OH (3647-104).

exPADRE=AlaGlySerArgSerGly(DAla)LysChaValAlaAlaTrpThrLeuLysAla(D-Ala)Gly-OH

35 El grupo Fmoc de la resina Fmoc-exPADRE CLEAR (549,3 mg, 0,1 mmol, adquirida de Peptide International Inc.) se eliminó mediante piperidina al 20 %/DMF (8 ml x3). La resina se lavó con DMF (1,5 ml x 5). A continuación, la resina se trató con 3 ml de una disolución de DMF de 4-(3-amino-4-formilfenoxi)-butanoato (34,4 mg), HBTU (45,5 mg), HOBT (16,2 mg), y DIEA (52 ml) durante 6 horas a temperatura ambiente. La resina se lavó a continuación con DMF y el péptido se escindió a partir de la resina usando H₂O/Me₂S/TFA (10/10/80 % en v/v, 10 ml) durante 2 horas a

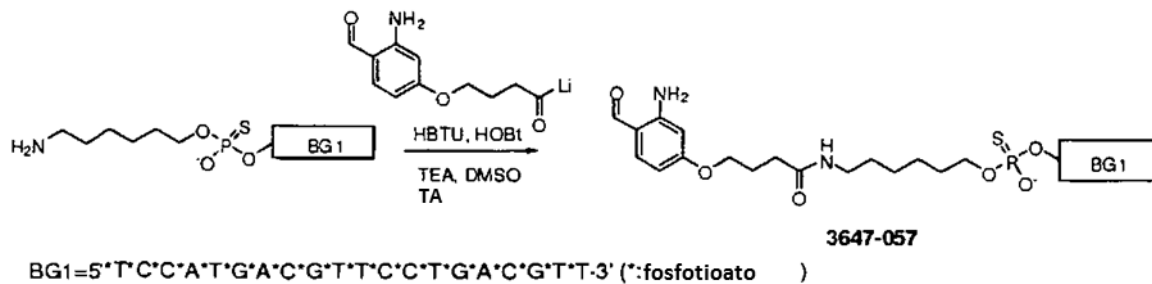
temperatura ambiente. La suspensión de la escisión se filtró a través de lana de vidrio, y la mayor parte del TFA se eliminó del filtrado mediante evaporación. El residuo se neutralizó con NaOH^{aq} 1 N y se diluyó con acetonitrilo, seguido por diálisis utilizando la membrana de diálisis SpectraPro™ 7 (MWCO de 1000) en 50 % de MeCN_{aq} (4l x 10). La disolución que quedó en la membrana de diálisis se liofilizó, dando como resultado el producto deseado como un sólido de color blanco. ESI-MS calculada para C₉₄H₁₅₀N₂₆O₂₆ [M+2H]²⁺/2: 1030,6, [M+3H]³⁺/3: 687,4; observada: 1030,7, 687,6.

Ejemplos 38-11: Síntesis de 3-(4-Acetil-3-aminofenil)propanoil-Gly(D-Ala)LysChaValAlaAlaTrpThrLeuLysAla(D-Ala)Gly-OH (3465-143).



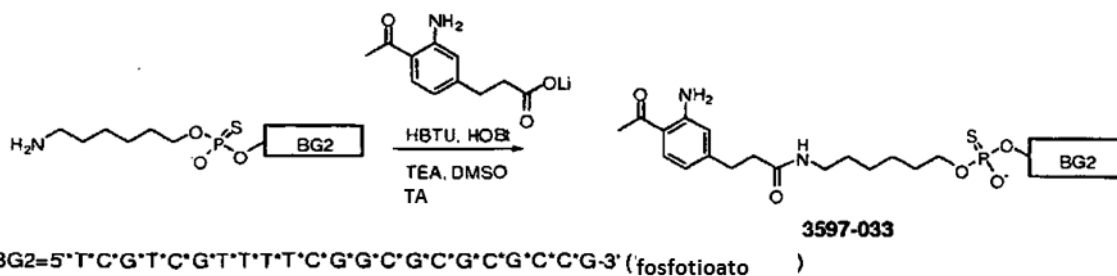
El grupo Fmoc de la resina Fmoc-PADRE CLEAR (105,3 mg, 0,02 mmol, adquirida de Peptide International Inc.) se eliminó mediante piperidina al 20 %/DMF (8 ml x3). La resina se lavó con DMF (1,5 ml x 5). A continuación, la resina se acopló con 0,5 ml de una disolución de DMF de 3-(4-acetil-3-aminofenil)propanoato de litio (5,1 mg), HBTU (9,1 mg), y HOBt (3,3 mg) durante 7 horas a temperatura ambiente. La resina se lavó con DMF y el péptido se escindió a partir de la resina mediante TIPS/H₂O/TFA (5/55/90 % en v/v, 3 ml) durante 2 horas. La suspensión de la escisión se filtró a través de lana de vidrio y la mayor parte del TFA se eliminó del filtrado mediante evaporación. El residuo se lavó mediante hexanos (3 ml x3), y se disolvió en 50 % de MeCN_(aq). Tras la liofilización, se obtuvo el producto bruto que se purificó a continuación mediante HPLC preparativa, dando como resultado el producto deseado como un polvo de color amarillo claro. ESI-MS calculada para C₇₇H₁₂₀N₁₈O₁₈ [M+2H]²⁺/2: 793,5, [M+3H]³⁺/3: 529,3; observada: 793,6, 529,4.

Ejemplo 38-12: Síntesis de 6-(4-(3-Amino-4-formilfenoxi)butanamido)hexil -5'- *T*C*C*A*T*G*A*C*G*T*T*C*C*T*G*A*C*G*T*T-3' (*: fosfotioato) (3647-057).



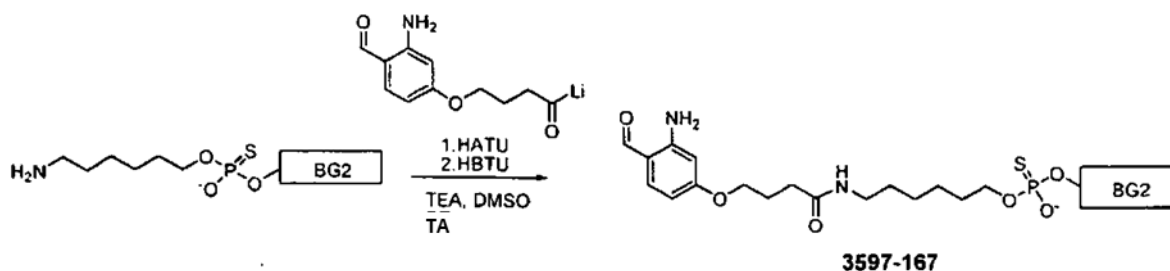
HBTU (6,1mg), HOBt (2,2mg), y trietilamina (7 ml) se añadieron a 1 ml de disolución de DMSO de 4-(3-amino-4-formilfenoxi)-butanoato de litio (4,6 mg), y la reacción se agitó a temperatura ambiente durante 1 hora. Una alícuota de 276 ml de la disolución resultante se añadió a continuación a 1,2 ml de disolución de DMSO de oligo BG1 modificado con amino (12,1 mg, 1,8 mmol, adquirido de Integrated DNA Technologies, Inc.) y trietilamina (15 ml), y la reacción se agitó a temperatura ambiente durante 2 días. La mezcla de reacción se diluyó con agua y se dializó frente a agua (4l x10) utilizando Slide-A-Lyzer™ (MWCO de 3500). La disolución dializada se liofilizó, dando como resultado el producto deseado como un sólido ESI-Q-TOF de color blanco calculado para C₂₁₁H₂₇₃N₆₉O₁₀₈P₂₀S₂₀: 67 59,7; observada: 6759,1.

Ejemplo 38-13: Síntesis de 6-(3-(4-acetil-3-aminofenil)propanamido)hexil-5'-*T*C*G*T*C*G*T*T*T*C*G*G*C*G*C*G*C*G*C*G-3' (*:fosfotioato) (3597-033).



5 HBTU (5,9 mg) y HOBt (2,1 mg) se añadió a 1 ml de disolución de DMSO 3-(4-acetil-3-aminofenil)propanoato de litio (3,3 mg), y la reacción se agitó a temperatura ambiente durante 1 hora. Una alícuota de 20 ml de la disolución resultante se añadió a continuación a 0,2 ml de una disolución de DMSO de un oligo BG2 modificado con amino (1,88 mg, 0,26 mmol, adquirido de Integrated DNA Technologies, Inc.) y trietilamina (2 ml), y la reacción se agitó a temperatura ambiente durante 20 horas. La mezcla de reacción se aplicó a una columna NAP-25™ equilibrada con H₂O y eluida con H₂O. Cada ml de fracción se vigiló mediante LC-MS y las fracciones que contenían el producto deseado se combinaron y liofilizaron, dando como resultado el producto deseado como un sólido de color blanco. MS (ESI+): m/z 1858,8 ([M+4H]⁴⁺/4).

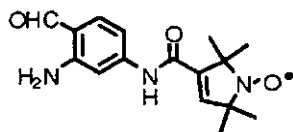
15 **Ejemplo 38-14:** Síntesis de 6-(4-(3-amino-4-formilfenoxi)butanamido)hexil-5'-*T*C*G*T*C*G*T*T*T*C*G*G*C*G*C*G*C*G*C*G-3' (*:fosfotioato) (3597-167).



20 HATU (7,4 mg) se añadió a 1 ml de disolución de DMSO de 4-(3-amino-4-formilfenoxi)-butanoato de litio (5,5 mg), y la reacción se agitó a temperatura ambiente durante 1 hora. Una alícuota de 60 ml de la disolución resultante se añadió a continuación a 0,6 ml de una disolución de DMSO de un oligo BG2 modificado con amino (6,9 mg, 1,0 mmol) y trietilamina (7,5 ml), y la reacción se agitó a temperatura ambiente durante 20 horas. A continuación, otra alícuota de 60 ml de la disolución del éster activado que se preparó de manera reciente de la misma manera excepto por utilizar HBTU en vez de HATU se añadió a la mezcla de reacción, y la reacción se agitó durante 2 días más. La mezcla de reacción se separó en 3 porciones. Cada porción se aplicó a una columna NAP-25™ equilibrada con H₂O y eluida con H₂O. Cada ml de fracción se vigiló mediante LC-MS y las fracciones que contenían el producto deseado se combinaron y liofilizaron, dando como resultado el producto deseado como un sólido de color blanco. ESI-Q-TOF calculada para C₂₂₉H₂₉₆N₇₈O₁₂₀P₂₂S₂₂: 7442,8; observada: 7442,7.

Ejemplo 39: Síntesis de reactivos ABA de marca de espín

Ejemplos 39-1: Síntesis de N-(3-Amino-4-formilfenil)- 1-hidroxi-2,2,5,5-tetrametil-3-pirrolin-1 -oxil-3-carboxamida (X3626-112b).

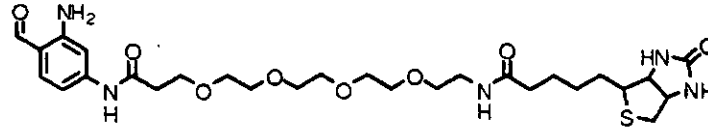


40 Al ácido 2,2,5,5-tetrametil-3-pirrolin-1-oxil-3-carboxílico (55,3 mg), se añadieron HBTU (102 mg) y DIEA (105 ml) 2,0 ml de DMF. Tras agitarse a temperatura ambiente durante 30 minutos, se añadió 2,4-diaminobenzaldehído (40,8 mg). La mezcla se agitó a temperatura ambiente hasta que la reacción se completó, se vigiló mediante HPLC. El producto del título se aisló mediante HPLC preparativa. MS (ESI+): m/z 303,15 (MH⁺).

Ejemplo 40: Síntesis de reactivos de biotina-ABA

Ejemplo 40-1: Síntesis de N-(3-amino-4-formilfenil)-1-(biotinamido)-3,6,9,12-tetraoxapentadecan-15-amida (X3626-140).

5

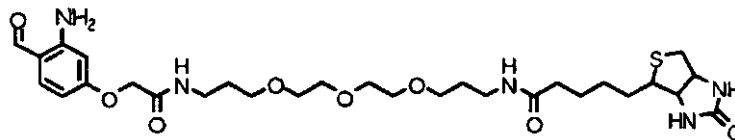


A 2,4-diaminobenzaldehído (11,5 mg), NHS-dPEG™4 biotina (QuantaBioscience, nº de cat. 10200, 50mg) y DMAP (10,4 mg) se añadió 1,0 ml de DMF. La mezcla de reacción se agitó a temperatura ambiente hasta que se completó la reacción, se vigiló mediante HPLC. El producto se aisló mediante cromatografía instantánea en gel de sílice. MS (ESI+): 610,28 (MH+).

10

Ejemplo 40-2: Síntesis de N-(1-(3-Amino-4-formilfenoxi)-2-oxo-7,10,13-trioxa-3-azahexadecan-16-il)-biotinamida (X3626-142).

15



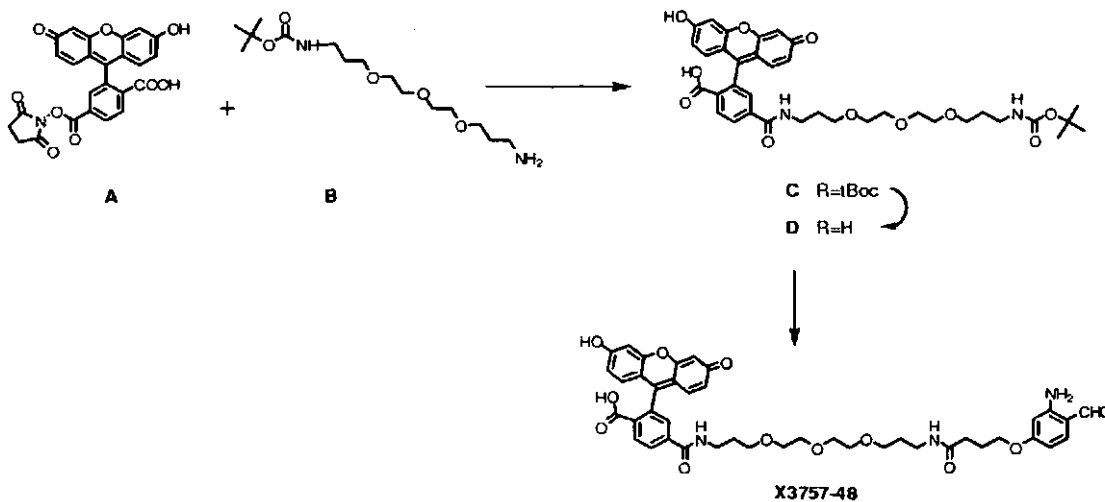
2-(3-amino-4-formilfenoxi)butanoato de litio (45,3 mg), HBTU (77,8 mg) y DIEA(31 ml) se añadió 1 ml de DMF, y la mezcla se agitó durante 30 minutos. La disolución resultante se añadió a biotina-dPEG™3-NH3+TFA (QuantaBioscience, nº de cat.10193, 100 mg), DIEA (47 ml) en 1,0 ml de DMF, y la reacción se agitó a temperatura ambiente, se vigiló mediante LCMS. El producto se aisló mediante HPLC preparativa. MS (ESI+): 624,30 (MH+).

20

Ejemplo 41: Síntesis de reactivos de PEG-ABA fluorescentes

Ejemplo 41-1: Síntesis de Fluoresceína-PEG-ABA (X3757-48).

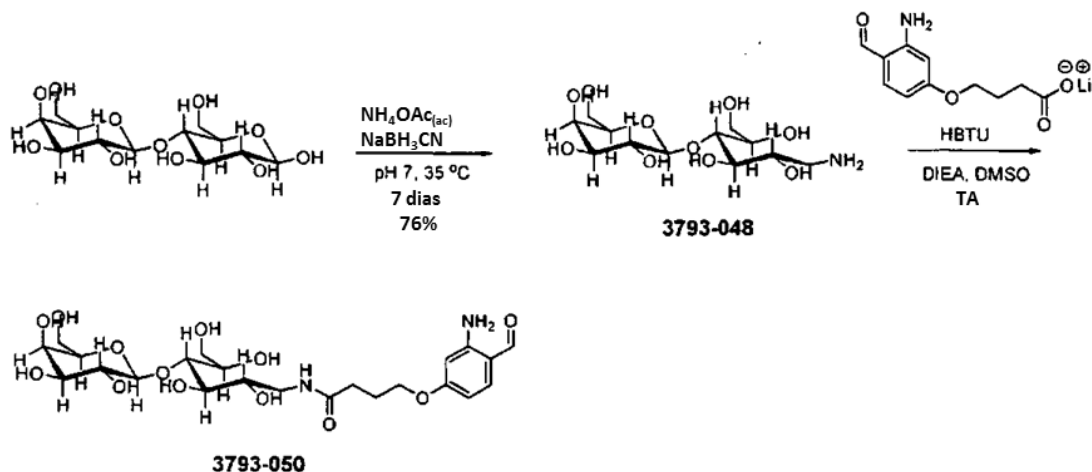
25



DMF (2,0 ml) se añadió a NHS-Fluoresceína A (23,6 mg), amina B (17,6 mg) y DIEA (8,7 ml). La mezcla se agitó a temperatura ambiente hasta que se consumió A, se vigiló mediante HPLC. El producto C se aisló mediante HPLC preparativa. ESI-MS (m/z) 679,72 (MH+). El producto (C, 7,4 mg) se disolvió a continuación en HCl 3 M (1,0 ml), y se eliminó el metanol mediante evaporación tras agitar 5 minutos a temperatura ambiente. Esta operación se repitió, dando como resultado la eliminación del grupo Boc que proporciona la amina D. A la amina D se añadieron 4-(3-amino-4-formilfenoxi)butanoato de litio (2,3 mg), HBTU (3,8 mg), DIEA (7,0 ml) y DMF (2 ml) a temperatura ambiente. El producto del título se aisló a continuación mediante HPLC preparativa. MS (ESI+): 784,30 (MH+).

30

35

Ejemplo 42: Síntesis de reactivos de oligosacáridos-ABA**Ejemplo 42-1:** Síntesis de Gal-Glu-1-amida de 3-amino-4-formilfenoxibutirato (3793-050).

5

se añadió NaBH₃CN (94,3 mg) a una disolución de H₂O (10 ml) de NH₄OAc (771 mg) y lactosa (180 mg) a pH 7 y la reacción se agitó a 35 °C durante 7 días. La mezcla de reacción se liofilizó, y el residuo se disolvió en H₂O, seguido por un paso a través de una resina de intercambio aniónico Dowex 1X8-400 (forma OH, 45 g) para eliminar el exceso de BH₃CN⁻, sus subproductos y el acetato. El eluyente se liofilizó, dando como resultado el producto bruto, que se purificó mediante cromatografía de intercambio catiónico utilizando la resina Dowex 50WX8-400 (forma H⁺, 30 g), dando como resultado el producto deseado (3793-048) como un polvo de color amarillo claro. ESI-MS calculada para C₁₂H₂₅NO₁₀ [MH]⁺: 344,1; observada: 344,2.

10

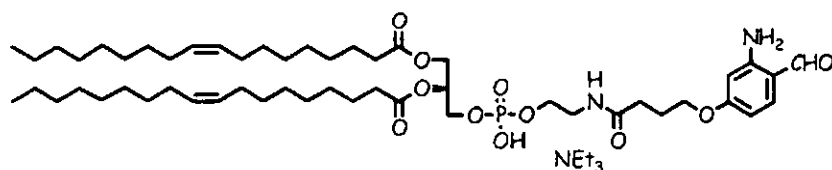
15

3-Amino-4-formilfenoxibutirato de litio (6,0 mg) se trató con HBTU (9,9 mg) y DIEA (7,7 ml) en DMSO anhidro (200 ml) durante 1 hora. Una disolución de DMSO (250 ml) de Gal-Glu-1-amina (3793-048) (7,7 mg) a partir del anterior se añadió a continuación a la mezcla de reacción a temperatura ambiente, seguido por agitación durante 1 día. La mezcla de reacción se liofilizó y purificó mediante HPLC-MS prep con una elución de NH₄OAc, dando como resultado el producto deseado (3793-050) como un polvo de color amarillo. ESI-MS calculado para C₂₃H₃₆N₂O₁₃ [MH]⁺: 549,2; observada: 549,3

20

Ejemplo 43: Síntesis de reactivos de fosfolípidos-ABA**Ejemplo 43-1:** Síntesis de DOPE-ABA (TU3627-092).

25



4-(3-amino-4-formilfenoxi)butanoato de litio (34 mg) y HBTU (57 mg) se colocaron en un vial de vidrio de 20 ml, y se añadieron 2 ml de DMF. La mezcla de reacción se agitó a temperatura ambiente durante 30 minutos para la activación. En un vial de 20 ml separado se colocó DOPE (76 mg, 1,2,-dioleoil-sn-glicero-fosfoetanolamina, NOF Corp.), seguido por DIEA (35 ml) y 3 ml de DCM. La disolución de color amarillo del éster activado del primer vial se transfirió al segundo vial, y la reacción se agitó a temperatura ambiente. Después de 24 horas, la mezcla de reacción completa se aplicó a una columna de SiO₂ preempacutada de 12 g equilibrada con el disolvente A (disolvente A: 5 % de NEt₃ en DMC, disolvente B: NEt₃ al 5 % en MeOH), y la columna se eluyó con un gradiente lineal de 0 a 15 % de B en A durante 15 minutos, dando como resultado el producto parcialmente purificado como un aceite muy viscoso amarillo claro. Este producto se purificó de nuevo mediante cromatografía instantánea utilizando una columna de SiO₂ de 12 g (disolvente A: 5 % de NEt₃ en DMC, disolvente B: NEt₃ al 5 % en MeOH), y se eluyó con un gradiente lineal de 2 a 10 % de B en A durante 15 minutos. La fracción que contiene el producto bruto se concentró a presión reducida, dando como resultado la sal de trietilamonio del compuesto del título. Rf: 0,43 (SiO₂, 10 % de MeOH en DCM). MS (ESI⁺): m/z 949,60 (MH⁺). H-RMN (400MHz, CDCl₃): 0,875 (6H, t, J = 6,8Hz), 1,28 (40H, picos amplios múltiples), 1,318 (9H, t, J = 7,4Hz), 1,582 (4H, m), 2,00 (m, 8H), 2,124 (2H, quint, J = 6,8Hz), 2,28 (4H, m), 2,373 (2H, t, J = 7,0Hz), 3,042 (6H, c, J = 7,2Hz), 3,476 (2H, m), 4,00 (6H, m), 4,154 (1H, dd, J = 6,8, 12,0Hz), 4,373 (1H, dd, J = 3,2, 12,0Hz), 5,233 (1H, m), 5,337 (4H, m), 6,217 (1H, d, J = 2,0Hz), 6,255 (1H, dd, J = 2,0, 8,4Hz), 6,56 (2H, pico muy amplio), 7,313 (1H, d, J = 8,8Hz), 7,432

30

35

40

(1H, br.t, J = 8,4Hz), 9,674 (1H, s), 11,97 (1H, br.s)

Ejemplo 44: Reducción del enlace de acoplamiento con PCL

5 Se encontró que la reducción del enlace de acoplamiento con PCL evita la disociación de los conjugados de proteína basados en PCL. La Figura 54A muestra el análisis por espectrometría de masas ESI de hFGF21-Lys150PCL acoplada con 2-ABA y a continuación reducida con 20 mM de NaCNBH₃ durante 1 hora. La Figura 54B muestra el análisis por espectrometría de masas ESI del conjugado de hFGF21- 150PCL 2-ABA reducido tras haberse dializado en tampón fosfato 10 mM (pH 7,5) e incubado a 50°C durante 1 día.

10 La Figura 55 demuestra la estabilidad de la unión de PCL para la FGF21 PEGilada con y sin reducción usando NaCNBH₃. Se hizo reaccionar el Arg154PCL mutante de FGF21 con 30,3 kDa-2-ABA-PEG (véase TU3627-024; Ejemplo 37-7) y se purificó. El FGF21Ara154PCL purificado de 30,3 kDa-2-ABA-PEG se redujo con 20 mM de NaBH₃CN durante 16 horas (temperatura ambiente, pH 7,5, 100 mM de proteína). Las muestras se incubaron durante 16 horas a 4 °C, temperatura ambiente 37°C y 50°C, y a 95°C durante 5 minutos. Se muestran geles SDS-PAGE de muestras reducidas y muestras no reducidas en la Figura 55A. Además, La Figura 55B muestra un gel SDS-PAGE de las muestras no reducidas incubadas durante 60 horas a 4°C, temperatura ambiente 37°C y 50°C, y 95°C.

20 Ejemplo 45: Estudios de RMN de PCL-A (Lys-P5C) y PCL-B (Lys-P2C) modificado covalentemente por 2-ABA.

La reacción de PCL-A (véase 3647-125, Ejemplo 36-1) y PCL-B (véase 3793-011, Ejemplo 36-2) con 2-aminobenzaldehído (2-ABA) y la estructura del aducto de PCL-ABA resultante se estudiaron mediante espectroscopía de resonancia magnética nuclear (RMN) normalizada 1D y 2D.

25 Para la reacción de PCL-A y PCL-B con 2-ABA y la posterior caracterización de los productos, se adquirieron los datos de RMN a 300 K en un instrumento de RMN a 400 MHz Bruker Avance (Bruker Biospin, Billerica, MA) equipado con una criosonda ¹H/¹³C/¹⁹F/³¹P-QNP. los espectros de ID ¹H se registraron normalmente con 16 barridos, un retraso de la relajación de 5 s, puntos de datos del complejo 16384 con una anchura de recorrido de 12 ppm, los espectros COSY ¹H-¹H se registraron normalmente con 4 barridos, 256 experimentos t₁ y los espectros ROESY ¹H-¹H con 8 barridos, 512 experimentos t₁. los espectros HMBC ¹H-¹³C HMBC se registraron normalmente con 32 barridos, 256 experimentos t₁ y los espectros HMQC ¹H-¹³C se registraron con 4 barridos, 128 experimento t₁ utilizando una anchura espectral de 222 ppm en la dimensión del carbono y 12 ppm o 7,5 ppm en la dimensión del protón.

35 Para la caracterización del aducto reducido, todos los espectros se registraron a 300 K en un instrumento Bruker Avance a 600 MHz equipado con una criosonda ¹H/¹³C/¹⁵N-TXI. los espectros de ¹H se registraron con 64 barridos, un retraso de la relajación de 2 s, puntos de datos del complejo 16384 con una anchura de recorrido de 14 ppm con esculpido de excitación para la supresión de agua. los espectros COSY ¹H-¹H se registraron normalmente con 16 barridos, 1024 experimentos t₁ usando una anchura espectral de 10 ppm. los espectros HMQC ¹H-¹³C se registraron normalmente con 8 barridos, 256 experimentos t₁ utilizando una anchura espectral de 160 ppm en la dimensión del carbono y 10 ppm en la dimensión del protón. los espectros HMBC ¹H-¹³C se registraron normalmente con 88 barridos, 256 experimentos t₁ a 300K utilizando una anchura espectral de 180 ppm en la dimensión del carbono y 10 ppm en la dimensión del protón.

45 La reacción del PCL-A con 2-ABA se vigiló como sigue: 1,0 mg de PCL-A sintetizada tal como se describe en el Ejemplo 36-1 se disolvió en 0.5 ml de 1X PBS en D₂O. 10 ml de 10 mM de ácido 3-(trimetilsilil)propiónico (TSP) en D₂O se añadieron como patrón interno y para la determinación de la concentración mediante RMN. 3,7 mg de 2-aminobenzaldehído (2-ABA; adquiridos de Sigma) se disolvieron en 0,5 ml de 1X PBS en D₂O y 10 ml de TSP 10 mM. La concentración de ambas muestras se determinó mediante RMN. Las señales de RMN de los materiales de partida (Tabla 7) se asignaron utilizando métodos de RMN normalizados que incluían 1D ¹H, COSY ¹H-¹H, ROESY ¹H-¹H, experimentos de HMBC ¹H-¹³C y ¹H-¹³C HMQC.

Tabla 7: Asignaciones de señal de RMN de PCL-A y 2-ABA sin reaccionar

Número atómico	Cambio (pp2m)	H	Tipo	J (Hz)
19	1,31	2	m	-
18	1,47	2	quin	7,34
4	1,68	1	m	-
20	1,74	1	m	-
4	2,16	1	m	-
3	2,57	1	m	-
17	3,14	1	t	6,97
21	3,59	1	t	6,24
5	4,51	1	m	-

ES 2 477 543 T3

12,10	6,76	1	m	-
11	7,34	1	ddd	8,50, 7,03, 1,59
13	7,53	1	dd	8,19, 1,59
2	7,76	1	m	-
7	9,68	1	s	-

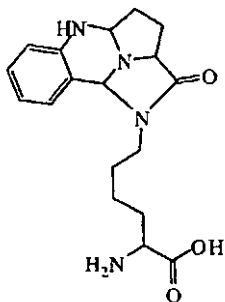
5 Para iniciar la reacción, 325 ml de una disolución PCL-A se mezclaron con 175 ml de la disolución de 2-ABA. La mezcla de reacción resultante se transfirió a un tubo de RMN y contenía PCL-A y 2-ABA a una relación molar 1:1 adecuada. Se dejó que la reacción continuara a temperatura ambiente y los espectros de RMN se adquirieron periódicamente en los momentos indicados (Figura 56). Durante las señales de reacción para el inicio del material de PCL-A (puntos) y para el reactivo 2-ABA (estrellas) desapareció y la reacción continuó hasta la finalización convirtiéndose todo el PCL-A (Existió un ligero exceso de 2-ABA en la muestra). En el primer punto temporal adquirido 0,5 h después de la mezcla, se detectaron dos nuevas especies a una relación de aproximadamente 2:1 (las resonancias representativas están marcadas por flechas). Las especies menores en el curso de algunos días
10 convirtieron completamente a las especies mayores.

El producto de reacción final de PCL-A con 2-ABA se caracterizó por 1D ^1H y ^{13}C , COSY 2D ^1H - ^1H , ROESY ^1H - ^1H , la espectroscopía de RNM de HMBC ^1H - ^{13}C y HMQC ^1H - ^{13}C con las muestras preparadas en D_2O y d_6 -DMSO. La muestra en d_6 -DMSO se purificó mediante HPLC. En la Tabla 8 se resumen las asignaciones de la señal para las resonancias de protones y carbono y las correlaciones observadas en los espectros de HMBC ^1H - ^{13}C en d_6 -DMSO.
15

Tabla 8:

correlaciones de HMBC ^{13}C observadas para la forma mayor del aducto de PCL-A/2-ABA (en d6-DMSO)											
Correlaciones entre ^1H - ^{13}C						correlaciones entre ^1H - ^{13}C					
Nº	Átomo F2	Átomo F1	F2 (ppm)	F1 (ppm)		Nº	Átomo F2	Átomo F1	F2 (ppm)	F1 (ppm)	
1	1	2	5,77	72,69		27	1	9	5,77	148,47	
2	4	2	2,00	72,73		28	2	9	4,06	148,42	
3	4	2	1,84	72,83		29	7	9	5,42	148,50	
4	5	2	3,78	72,59		30	11	9	7,09	148,43	
5	7	2	5,42	72,72		31	12	9	6,70	148,55	
6	1	3	5,76	32,36		32	13	9	7,21	148,39	
7	2	3	4,06	32,20		33	1	10	5,77	115,57	
8	4	3	1,82	26,41		34	7	10	5,43	115,78	
9	4	3	1,82	32,22		35	11	10	7,09	115,53	
10	4	3	2,01	32,42		36	12	10	6,70	115,47	
11	5	3	3,78	32,41		37	13	10	7,21	115,50	
12	2	4	4,06	26,55		38	13	11	7,21	129,27	
13	5	4	3,78	26,55		39	10	12	6,66	118,14	
14	7	4	5,43	26,38		40	1	13	5,77	128,39	
15	2	5	4,06	64,16		41	7	13	5,43	128,48	
16	4	5	1,85	63,96		42	11	13	7,09	128,49	
17	4	5	2,01	64,04		43	12	13	6,70	128,11	
18	10	7	6,66	71,92		44	4	14	2,00	173,00	
19	13	7	7,21	71,95		45	4	14	1,88	173,08	
20	17	7	2,87	71,90		46	4	14	1,83	173,09	
21	17	7	2,38	71,99		47	5	14	3,78	173,11	
22	1	8	5,76	120,55		48	17	14	2,87	173,28	
23	7	8	5,42	120,43		49	17	14	2,37	173,29	
24	10	8	6,67	120,29		50	18	17	1,04	39,08	
25	11	8	7,09	120,35		51	18	17	0,73	39,40	
26	12	8	6,70	120,32		52	19	17	1,04	39,08	

En la Figura 57A se muestran la numeración de los átomos y las correlaciones de enlace a su través seleccionadas en el HMBC. En contraste con las observaciones en el material de partida de PCL-A, los dos protones del metileno en el carbono 17 se observan en dos diferentes valores de desplazamiento químico. Estos protones muestran también una correlación de enlace a su través heteronuclear en los espectros de HMBC para el carbono 7 sugiriendo la formación de un enlace covalente entre el nitrógeno 16 y el carbono 7. El protón en el carbono 7 resuena a 5,6 ppm (el protón 7 es la señal del aducto de PCL-A/2-ABA resaltado en la Figura 201) y muestra correlaciones con el carbono 2, 8, 9, 5 y 13. De forma similar, el protón del carbono 2 presenta una correlación de HMBC con el carbono 9. Estas observaciones y las otras observaciones de RMN de las dos muestras caracterizadas en D₂O y en D6-DMSO son consistentes con la estructura representada en la Figura 57A para el producto mayor de PCL-A/2-ABA. Por tanto, en su conjunto, los puntos de evidencia de RMN para la siguiente estructura para el producto de la reacción entre PCL-A y 2-ABA.



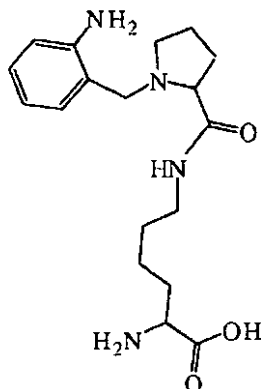
Se ha intentado también caracterización estructural de la forma menor observada en la mezcla de reacción (Figura 56). El análisis se complicó debido a la baja concentración de la forma menor y al hecho de que se convierte lentamente a la forma mayor. El análisis no es concluyente. Las formas menores y mayores tienen un protón en el carbono 7 y los cambios químicos de ambos son muy similares (flechas a 5,6 ppm en la Figura 56). Los contactos ROESY en el espacio a su través son también similares en ambas formas. Los dos protones del metileno en el carbono 17 están degenerados, lo que significa que presentan el mismo desplazamiento químico. Para la forma menor del aducto PCL-A/2-ABA así como para el PCL-A sin reaccionar. Estos protones, sin embargo, tienen distintos cambios químicos en la forma mayor. Estos protones muestran también una correlación HMBC para el carbono 7 en las formas menor y mayor sugiriendo un enlace covalente entre el carbono 7 y el nitrógeno 16. La observación de los cambios químicos degenerados de los protones del metileno en el carbono 17 en la forma menor pueden implicar que el enlace covalente entre el carbono 7 y el nitrógeno 16 podría ser semiestable, y que las observaciones de RMN para la forma menor son el resultado del intercambio químico entre dos o más especies. Se puede pensar que la forma menor sea un estereoisómero semiestable alternativo de la forma mayor en el interdesplazamiento químico con la forma protonada del aducto de PCL-A/2-ABA (Figura 57B).

Además, se mezclaron disoluciones de PCL-A sintético y 2-ABA en D₂O tal como se describe anteriormente y se dejó continuar la reacción hasta la finalización que se vigiló mediante RMN. La adición de alícuotas de una disolución de cianoborohidruro de sodio (NaCNBH₃) en D₂O al tubo de RMN dio como resultado la reducción del aducto de PCL-A/2-ABA a una nueva especie. Se añadieron alícuotas de NaCNBH₃ adicionales hasta la resonancia a 5,6 ppm que es característica del aducto de PCL-A/2-ABA (Figura 56, flechas) desapareció completamente. La muestra de RMN del aducto de PCL-A/2-ABA completamente reducida se liofilizó a continuación de forma repetida y se volvió a disolver en D₂O seco a fin de volver a concentrar la muestra y eliminar el agua. La muestra final se reconstituyó en 0,5 ml de D₂O seco. La forma reducida se caracterizó mediante COSY 2D ¹H-¹H normalizado, espectroscopía de RMN de HMBC ¹H-¹³C y HMQC ¹H-¹³C. En la Tabla 9 se resumen las asignaciones de la señal para las resonancias de protones y carbono y las correlaciones del enlace a su través heteronuclear procedentes de los espectros de HMBC ¹H-¹³C.

Tabla 9:

correlaciones de HMBC ¹ H- ¹³ C observadas para la forma reducida del aducto de PCL-A/2-ABA (en D2O)													
Nº	Correlaciones entre ¹ H				Nº	Correlaciones entre ¹³ C							
	Átomo F2	Átomo F1	F2 (ppm)	F1 (ppm)		Átomo F2	Átomo F1	F2 (ppm)	F1 (ppm)				
1	3	2	1,85	54,62	21	7	3,89	145,27	41	19	18	1,30	27,87
2	7	2	3,58	54,79	22	7	3,58	145,30	42	17	19	2,97	21,71
3	7	2	3,89	54,86	23	11	7,24	145,30	43	18	19	1,35	21,70
4	2	3	2,60	23,39	24	13	7,21	145,25	44	18	19	1,42	21,72
5	4	3	1,76	23,31	25	12	6,87	116,70	45	20	19	1,86	21,71
6	4	3	2,29	23,38	26	13	7,22	129,05	46	21	19	3,72	21,70
7	2	4	3,27	30,38	27	10	6,90	119,25	47	18	20	1,31	29,99
8	5	4	3,21	30,39	28	7	3,58	130,66	48	18	20	1,42	30,15
9	2	5	3,27	67,08	29	7	3,89	130,71	49	18	20	1,36	30,26
10	4	5	2,29	67,06	30	11	7,24	130,59	50	19	20	1,31	29,99
11	4	5	1,77	67,09	31	4	1,76	177,53	51	19	20	1,36	30,26
12	7	5	3,89	66,98	32	4	2,29	177,58	52	21	20	3,72	30,16
13	7	5	3,58	67,05	33	5	3,21	177,50	53	19	21	1,30	54,62
14	2	7	2,61	56,85	34	17	2,97	177,56	54	20	21	1,85	54,62
15	5	7	3,21	56,79	35	18	1,37	38,49	55	20	23	1,87	175,22
16	13	7	7,22	56,83	36	18	1,42	38,56	56	21	23	3,72	175,17
17	7	8	3,89	123,94	37	18	1,34	38,59					
18	7	8	3,58	124,02	38	19	1,37	38,49					
19	10	8	6,90	124,03	39	19	1,34	38,59					
20	12	8	6,87	124,06	40	17	2,97	27,90					

En la Figura 57C se muestra la numeración de los átomos. Las diferencias clave de las correlaciones observadas en la forma mayor del aducto de PCL-A/2-ABA son la ausencia de una correlación de enlace a su través heteronuclear entre los protones del metileno en el carbono 17 y el carbono 7 y la ausencia de una correlación de enlace a su través entre los protones en el carbono 2 y el carbono 9. Estas y las demás observaciones de RMN son consistentes con la siguiente estructura (también en la Figura 57C) para el aducto de PCL-A/2-ABA reducido.



Se estudió también la reacción de PCL-B con 2-ABA en condiciones idénticas que aquellas para la reacción de PCL-A con 2-ABA. PCL-b sintetizada tal como se describe en el Ejemplo 36-2 se disolvió en D₂O (como anteriormente), y una mezcla de reacción con PCL-B y 2-ABA a una relación molar aproximada de 1.1 se transfirió en un tubo de RMN. En la Tabla 10 se relacionan las asignaciones de la señal de los materiales de partida.

Tabla 10: Asignaciones de señal de RMN de PCL-B y 2-ABA sin reaccionar

Número atómico	Desplazamiento (ppm)	H	Tipo	J (Hz)
19	1,33	1	m	-
18	1,50	1	q	7,02
20	1,77	1	m	-
3	1,88	1	m	-
4	2,69	1	m	-
17	3,21	1	t	6,94
21	3,61	1	m	-
2	3,85	1	m	-
12, 10	6,76		m	-
11	7,34	1	ddd	8,44, 7,05, 1,61
13	7,53	1	dd	8,18, 1,61
7	9,67	1	S	-

Se dejó que la reacción continuara a temperatura ambiente y los espectros de RMN se adquirieron periódicamente en los momentos indicados (Figura 58). En contraste con la muestra de PCL-A, PCL-B no reacciona fácilmente con 2-ABA. Incluso después de 17 días, mucho del material de partida (estrellas para 2-ABA; puntos para PCL-B) está aún presente aunque una pequeña cantidad se convirtió en una nueva especie (flechas). Esta especie podría no caracterizarse adicionalmente. Sin embargo, El análisis de RMN de las dos reacciones indica claramente que la reactividad de PCL-A con 2-ABA es mucho mayor que la de PCL-B.

Ejemplo 46: Derivatización de Pirrolisina (Pyl) y PCL incorporadas en mEGF

La incorporación de pirrolisina (Pyl) y PCL en mEGF se llevó a cabo tal como se describe en el Ejemplo 14, excepto que las células de E. coli BL21(DE3) se transformaron simultáneamente con pAra-pylSTBCD y el gen mutante mEGF Tyr10TAG en un vector pET22b. El mEGF mutante resultante se denomina a partir de ahora mEGF-Tyr10PCL/Pyl. El análisis de ESI-MS reveló que PCL y PYL se incorporaron en la posición 10 de mEGF (Figura 59A; masa esperada de mEGF Tyr10PCL = 7296 Da; masa esperada de mEGF Tyr10Pyl = 7310 Da). De esta manera, se obtuvo una proteína mEGF tanto con PCL como PYL incorporadas. En esta muestra concreta, la especie predominante fue mEGF con PCL incorporada en el presente documento.

Para demostrar la derivatización de esta mezcla y para demostrar que PYL está modificada usando los métodos proporcionados en el presente documento, el acoplamiento del reactivo ABA TU3627-014 (véase el Ejemplo 34-2) con

5 mEGF Tyr10PCL/PYL se llevó a cabo en 10 x PBS (pH 7,0) y 1 % (v/v) de DMSO a 25°C durante 16 horas. La reacción de conjugación se inició mediante la adición de 10 mM de mEGF-Tyr10PCL/PYL y 1 mM de TU3627-014. Se confirmó la formación de los conjugados de proteína mediante espectrometría de masas con ionización por electropulverización (ESI-MS) (Figura 59B). La relación entre el aducto de PCL y PYL asemeja la relación de PCL y PYL en la proteína sin reaccionar (Figura 58A), indicando por tanto una reactividad similar para PCL y Pyl.

Se entiende que los ejemplos y realizaciones descritos en el presente documento son solo a fines ilustrativos.

10 LISTADO DE SECUENCIAS

<110> IRM Geierstanger, Bernhard

15 <120> PIRROLINA-CARBOXI-LISINA GENERADA BIOSINTÉTICAMENTE Y MODIFICACIONES DE PROTEÍNAS ESPECÍFICAS DE SITIO MEDIANTE DERIVATIZACIÓN QUÍMICA DE RESTOS DE PIRROLINA-CARBOXI-LISINA Y PIRROLISINA

<130> PAT053160-WO
<160> 35

20 <170> PatentIn version 3.5

<210> 1
<211> 1365
<212> ADN

25 <213> Methanosarcina mazeii

<400> 1

ES 2 477 543 T3

atggataaaa aaccactaaa cactctgata tctgcaaccg ggctctggat gtccaggacc 60
ggaacaattc ataaaataaa acaccacgaa gtctctcgaa gcaaaatcta tattgaaatg 120
gcatgcggag accaccttgt tgtaaacacac tccaggagca gcaggactgc aagagcgctc 180
aggcaccaca aatacaggaa gacctgcaaa cgctgcaggg tttcggatga ggatctcaat 240
aagttcctca caaaggcaaa cgaagaccag acaagcgtaa aagtcaaggt cgtttctgcc 300
cctaccagaa cgaaaaaggc aatgccaanaa tccgttgcca gagccccgaa acctcttgag 360
aatacagaag cggcacaggc tcaaccttct ggatctaaat tttcacctgc gataccggtt 420
tccaccaag agtcagtttc tgtcccgga tctgtttcaa catcaatate aagcatttct 480
acaggagcaa ctgcatccgc actggtaaaa ggaatacga accccattac atccatgtct 540
gccctgttc aggcaagtgc ccccgcaactt acgaagagcc agactgacag gcttgaagtc 600
ctgttaaacc caaagatga gatttccttg aattccggca agcctttcag ggagcttgag 660
tccgaattgc tctctcgag aaaaaaagac ctgcagcaga tctacgcgga agaaagggag 720
aattatctgg gaaactcga gcgtgaaatt accaggttct ttgtggacag gggttttctg 780
gaaataaaat ccccgatcct gatccctctt gagtatatcg aaaggatggg cattgataat 840
gataccgaac tttcaaaaca gatcttcagg gttgacaaga acttctgcct gagaccatg 900
cttgctcaa acctttaca ctacctgccc aagcttgaca gggccctgcc tgatccaata 960
aaaatTTTTG aataggccc atgctacaga aaagagtcg acggcaaaga acacctgaa 1020
gagtttacca tgctgaactt ctgccagatg ggatcgggat gcacacggga aaatcttgaa 1080
agcataatta cggacttct gaaccacctg ggaattgatt tcaagatcgt aggcgatcc 1140
tgcattgtct atggggatac ccttgatgta atgcacggag acctggaact ttctctgca 1200

gtagtcggac ccataccgct tgaccgggaa tggggatttg ataaacctg gataggggca 1260
ggtttcgggc tcgaacgcct tctaaagggt aaacacgact ttaaaaatat caagagagct 1320
gcaaggtccg agtcttacta taacgggatt tctaccaacc tgtaa 1365

<210> 2
<211> 1050
<212> ADN
<213> Methanosarcina mazeii

5

<400> 2

ES 2 477 543 T3

atgatccaga aaatggcaac cgaagaactt gacaggttcg gggagaaaat tattgaaggt 60
 tttaaattgt ctgatgatga cctcagggct cttctttctc ttgaattcga agaagagctg 120
 gaaaagcttt actatgtagc tagaaaggtc agaaactatt atttcggcaa caggggtgttt 180
 cttaaactggt ttattttattt ctcaacttat tgtaaaaacc agtgctcttt ttgctactat 240
 aactgtaaaa acgaaattaa cgcctaccgc ctgaccgggtg aagagggttaa agagatgtgc 300
 aaagccctga aagggtgcagg ctttcacatg atcgacctga caatgggaga ggatccctat 360
 tactatgatg accctgaccg cttcgttgaa cttgtcagga cagtaaaaga agaactcggg 420
 cttccaataa tgattttctcc gggagttatg gatgacagca ccctcctgaa agccagggaa 480
 gaaggagcaa atttctttgc cctttatcag gagacttatg accgcgaact ttatggaaag 540
 ctaagggtag gtcagtcctt cgaaggaagg tttaatgcc gcaggtttgc aaaagaacag 600
 gggacttgta tagaagacgg cattcttacc ggcgtaggaa atgatatcga atcaactctt 660
 atatccctga aggggatgaa agcaaacaat cctgatatgg taagggtaat gacttttctg 720
 cctcaggaag gaactccgct tgaaggtttc agcgatagtt caaagctttc ggagctgaaa 780
 atcatagcga ttctcaggct catgtttcct gaatgcctga taccggcttc tcttgacctt 840
 gaaggcatag acggcatggt gcaccgttta aatgccggag caaatattgt aacctccatc 900
 ctcccagatt cacgcctgga aggggttgcc aattacgacc gccgcatgga agagagggac 960
 agggacgtta caagcgttgt caaaaggctg aaggttatgg gaatggaacc tgcgccgcag 1020
 gctgagtttg agagagtcct ggggtgctaa 1050

<210> 3
 <211> 1116
 <212> ADN
 <213> Methanosarcina mazeii

5

<400> 3

atgagagagt cctgggggtgc tagcctgaaa acaatatgcc ttataggcgg gaagctgcag 60
 ggcttcgagg ctgcatacct atctaagaaa gccggaatga aagtgcttgt aatagacaaa 120
 aaccgcgagg cgcttataag gaattatgcg gatgagttcc agtgttttaa cataacggaa 180

10

ES 2 477 543 T3

gagccggaaa aactcgtcgc gatatcaaaa aatggtgatg ccatactgcc ggtaaatagaa 240
 aaccttgaat gtatagaatt tctgaattct ataaaagaaa aattctcctg cccggtactt 300
 ttcgat,tttg aagcttacag gatcagcagg gataagagaa aatcaaaaaga atacttcgca 360
 tccataggaa ccccgacccc tcaggacaaa ccgctcggaac caccttattt tgtaaagcct 420
 ccctgcgaaa gcagcagtgt gggagcgaga ataatccatg acaggaaaga gcttaaagag 480
 cttgagcccg ggatgctcat agaagaatac gttgaagggg aagtggcttc acttgaggtc 540
 ataggggatg gaaataattt tgctgtggta aaggaaaccc ttgtacatat cgatgacacc 600
 tatgactgcc atatggtgac ccctctccct ctagaccctt ccttcaggga actatcctac 660
 tccttgacag caaacctgcc cttaaaagga attatggacg tggaagcgat ttccggcccc 720
 ctggggttaa aagttattga gatagatgcc cgtttcccga gccagactcc gactgcggtc 780
 tattattctt ccgggatcaa cctcatagaa ctctgttcc gggcttttaa tggaggcata 840
 gaagagatca aaactctccc tgaagacagg tactgcattt acgaacatct catgcttgca 900
 gaaaatggag tacttatccc tgtgggagaa caggtcctgt ccatgggaaa tgattacggc 960
 aattattatg aagaacctgg aatagagatt ttctgtgca aaggagagaa ccctgtatc 1020
 accctggttt tctggggcag agacagggaa gaagctgaag ctagaaaaaa caaagggctt 1080
 tcgattctaa aaagccgttt cggagctgct gcataa 1116

<210> 4

<211> 795

5 <212> ADN

<213> Methanosarcina mazeii

<400> 4

ES 2 477 543 T3

atggcacttt taaccccaga agacctggaa aatattaaca aacagcttca agaagctgat 60
 tctactgtcc gcagagttac agggcttgat ataaaaggta tctgtaaaga tttctacggc 120
 acaactccat gctgtgaaaa agtaggtatc gtgcctgtga cctcagggaa cgggatcata 180
 gggagctttt ccgaatccct gaatgcaatt gccgggtatt tcgggtttga cagttttatt 240
 actgatatgc ctgacgtcag cggatattat gaggcagtaa agaacggagc ccggatcata 300
 cttatggcag atgataatac cttccttgcc cacaacctga aaaatggaaa aatcgccaat 360
 aaccagccgt gtacaggcat aatttatgct gaaatagctt caagatacct gaaagccgat 420
 tccaaagaag tgcttgccgt gggctctggg aagggtggat ttccgggagc agcccatctc 480
 gtacagaaag gcttcaaggt ttacggatat gatgctgaca gaacccttct agaaaaaagc 540
 gtttccagcc tcggaattat acctttcaat cccgtcagcc ccgaaggcga caggcaaagg 600
 aagttttcca ttattttcga agcaaccccc tgtgcagaca cgattccgga atccgtaatt 660

 tcggaaaact gtgtgatttc taccctggg ataccctgtg caatctcaaa ggagctgcaa 720
 aaaaagtgtg gagttgaact tgtaatggaa ccaactgggga taggtacagc atcaatgctg 780
 tattctgtac tctaa 795

<210> 5
 <211> 72
 5 <212> ADN
 <213> Methanosarcina mazeii

<400> 5

ggaaacctga tcatgtagat cgaatggact ctaaatecgt tcagccgggt tagattcccc 60
 10 gggtttccgc ca 72

<210> 6
 <211> 223
 <212> PRT
 15 <213> Homo sapiens

<400> 6

ES 2 477 543 T3

Met Lys Thr Phe Ile Leu Leu Leu Trp Val Leu Leu Leu Trp Val Ile
 1 5 10 15

Phe Leu Leu Pro Gly Ala Thr Ala Gln Pro Glu Arg Asp Cys Arg Val
 20 25 30

Ser Ser Phe Arg Val Lys Glu Asn Phe Asp Lys Ala Arg Phe Ser Gly
 35 40 45

Thr Trp Tyr Ala Met Ala Lys Lys Asp Pro Glu Gly Leu Phe Leu Gln
 50 55 60

Asp Asn Ile Val Ala Glu Phe Ser Val Asp Glu Thr Gly Gln Met Ser
 65 70 75 80

Ala Thr Ala Lys Gly Arg Val Arg Leu Leu Asn Asn Trp Asp Val Cys
 85 90 95

Ala Asp Met Val Gly Thr Phe Thr Asp Thr Glu Asp Pro Ala Lys Phe
 100 105 110

Lys Met Lys Tyr Trp Gly Val Ala Ser Phe Leu Gln Lys Gly Asn Asp
 115 120 125

Asp His Trp Ile Val Asp Thr Asp Tyr Asp Thr Tyr Ala Val Gln Tyr
 130 135 140

Ser Cys Arg Leu Leu Asn Leu Asp Gly Thr Cys Ala Asp Ser Tyr Ser
 145 150 155 160

Phe Val Phe Ser Arg Asp Pro Asn Gly Leu Pro Pro Glu Ala Gln Lys
 165 170 175

Ile Val Arg Gln Arg Gln Glu Glu Leu Cys Leu Ala Arg Gln Tyr Arg
 180 185 190

Leu Ile Val His Asn Gly Tyr Cys Asp Gly Arg Ser Glu Arg Asn Leu
 195 200 205

Leu Asp Tyr Lys Asp Asp Asp Asp Lys His His His His His His
 210 215 220

<210> 7
 <211> 672
 <212> ADN
 <213> Homo sapiens

5

ES 2 477 543 T3

<400> 7

```

atgaaaacat tcatactcct gctctgggta ctgctgctct gggttatctt cctgcttccc      60
ggtgccactg ctcagcctga gcgcgactgc cgagtgagca gcttccgagt caaggagAAC      120
ttcgacaagg ctcgcttctc tgggacctgg tacgccatgg ccaagaagga ccccgagggc      180
ctctttctgc aggacaacat cgtcgcggag ttctccgtgg acgagaccgg ccagatgagc      240
gccacagcca agggccgagt ccgtcttttg aataactggg acgtgtgcgc agacatggtg      300
ggcaccttca cagacaccga ggaccctgcc aagttcaaga tgaagtactg gggcgtagcc      360
tcctttctcc agaaaggaaa tgatgaccac tggatcgtcg acacagacta cgacacgtat      420
gccgtgcagt actcctgccg cctcctgaac ctcgatggca cctgtgctga cagctactcc      480
ttcgtgtttt cccgggaccc caacggcctg ccccagaag cgcagaagat tgtaaggcag      540
cggcaggagg agctgtgcct ggccaggcag tacaggctga tcgtccacaa cggttactgc      600
gatggcagat cagaaagaaa ccttttggac tataaagacg atgacgataa gcatcaccat      660
caccatcact aa                                                                672

```

- 5 <210> 8
- <211> 672
- <212> ADN
- <213> Homo sapiens

10 <400> 8

```

atgaaaacat tcatactcct gctctgggta ctgctgctct gggttatctt cctgcttccc      60
ggtgccactg ctcagcctga gcgcgactgc cgagtgagca gcttccgagt caaggagAAC      120

ttcgacaagg ctcgcttctc tgggacctgg taggccatgg ccaagaagga ccccgagggc      180
ctctttctgc aggacaacat cgtcgcggag ttctccgtgg acgagaccgg ccagatgagc      240
gccacagcca agggccgagt ccgtcttttg aataactggg acgtgtgcgc agacatggtg      300
ggcaccttca cagacaccga ggaccctgcc aagttcaaga tgaagtactg gggcgtagcc      360
tcctttctcc agaaaggaaa tgatgaccac tggatcgtcg acacagacta cgacacgtat      420
gccgtgcagt actcctgccg cctcctgaac ctcgatggca cctgtgctga cagctactcc      480
ttcgtgtttt cccgggaccc caacggcctg ccccagaag cgcagaagat tgtaaggcag      540
cggcaggagg agctgtgcct ggccaggcag tacaggctga tcgtccacaa cggttactgc      600
gatggcagat cagaaagaaa ccttttggac tataaagacg atgacgataa gcatcaccat      660
caccatcact aa                                                                672

```

- 15 <210> 9
- <211> 672

ES 2 477 543 T3

<212> ADN
<213> Homo sapiens

<400> 9

5

```

atgaaaacat tcatactcct gctctgggta ctgctgctct gggttatctt cctgcttccc      60
gggtgccactg ctcagcctga gcgcgactgc cgagtgagca gcttccgagt caaggagaac      120
ttcgacaagg ctcgcttctc tgggacctgg tacgccatgg ccaagaagga ccccgagggc      180
ctctagctgc aggacaacat cgtcgaggag ttctccgtgg acgagaccgg ccagatgagc      240
gccacagcca agggccgagt ccgtcttttg aataactggg acgtgtgctc agacatggtg      300
ggcaccttca cagacaccga ggacctgcc aagttcaaga tgaagtactg gggcgtagcc      360
tcctttctcc agaaaggaaa tgatgaccac tggatcgtcg acacagacta cgacacgtat      420
gccgtgcagt actcctgccg cctcctgaac ctcgatggca cctgtgctga cagctactcc      480
ttcgtgtttt cccgggacct caacggcctg cccccagaag cgcagaagat tgtaaggcag      540
cggcaggagg agctgtgcct ggccaggcag tacaggctga tcgtccacaa cggttactgc      600
gatggcagat cagaaagaaa ccttttggac tataaagacg atgacgataa gcataccat      660
caccatcact aa                                                                672

```

<210> 10
<211> 672
<212> ADN
<213> Homo sapiens

<400> 10

10

```

atgaaaacat tcatactcct gctctgggta ctgctgctct gggttatctt cctgcttccc      60
gggtgccactg ctcagcctga gcgcgactgc cgagtgagca gcttccgagt caaggagaac      120
ttcgacaagg ctcgcttctc tgggacctgg tacgccatgg ccaagaagga ccccgagggc      180
ctctttctgc aggacaacat cgtcgaggag ttctccgtgg acgagaccgg ccagatgagc      240
gccacagcca agggccgagt ccgtcttttg aataactagg acgtgtgctc agacatggtg      300
ggcaccttca cagacaccga ggacctgcc aagttcaaga tgaagtactg gggcgtagcc      360
tcctttctcc agaaaggaaa tgatgaccac tggatcgtcg acacagacta cgacacgtat      420
gccgtgcagt actcctgccg cctcctgaac ctcgatggca cctgtgctga cagctactcc      480
ttcgtgtttt cccgggacct caacggcctg cccccagaag cgcagaagat tgtaaggcag      540
cggcaggagg agctgtgcct ggccaggcag tacaggctga tcgtccacaa cggttactgc      600
gatggcagat cagaaagaaa ccttttggac tataaagacg atgacgataa gcataccat      660
caccatcact aa                                                                672

```

15

<210> 11

ES 2 477 543 T3

<211> 672
 <212> ADN
 <213> Homo sapiens

5 <400> 11

```

atgaaaacat tcatactcct gctctgggta ctgctgctct gggttatctt cctgcttccc      60
ggtgccactg ctcagcctga gcgcgactgc cgagtgagca gcttccgagt caaggagaac      120
ttcgacaagg ctcgcttctc tgggacctgg tacgccatgg ccaagaagga ccccgagggc      180
ctctttctgc aggacaacat cgtcgcggag ttctccgtgg acgagaccgg ccagatgagc      240
gccacagcca agggccgagt ccgtcttttg aataactggg acgtgtgctc agacatggtg      300
ggcaccttca cagacaccga ggaccctgcc aagttcaaga tgaagttagt gggcgtagcc      360
tcctttctcc agaaaggaaa tgatgaccac tggatcgctc acacagacta cgacacgtat      420
gccgtgcagt actcctgccg cctcctgaac ctcgatggca cctgtgctga cagctactcc      480
ttcgtgtttt cccgggacct caacggcctg cccccagaag cgcagaagat tgtaaggcag      540
cggcaggagg agctgtgcct ggccaggcag tacaggctga tcgtccacaa cggttactgc      600
gatggcagat cagaaagaaa ccttttggac tataaagacg atgacgataa gcatcaccat      660
caccatcact aa                                                                672
    
```

10 <210> 12
 <211> 672
 <212> ADN
 <213> Homo sapiens

15 <400> 12

```

atgaaaacat tcatactcct gctctgggta ctgctgctct gggttatctt cctgcttccc      60
ggtgccactg ctcagcctga gcgcgactgc cgagtgagca gcttccgagt caaggagaac      120
ttcgacaagg ctcgcttctc tgggacctgg tacgccatgg ccaagaagga ccccgagggc      180
ctctttctgc aggacaacat cgtcgcggag ttctccgtgg acgagaccgg ccagatgagc      240
gccacagcca agggccgagt ccgtcttttg aataactggg acgtgtgctc agacatggtg      300
ggcaccttca cagacaccga ggaccctgcc aagttcaaga tgaagtacta gggcgtagcc      360
tcctttctcc agaaaggaaa tgatgaccac tggatcgctc acacagacta cgacacgtat      420
gccgtgcagt actcctgccg cctcctgaac ctcgatggca cctgtgctga cagctactcc      480
ttcgtgtttt cccgggacct caacggcctg cccccagaag cgcagaagat tgtaaggcag      540
cggcaggagg agctgtgcct ggccaggcag tacaggctga tcgtccacaa cggttactgc      600
gatggcagat cagaaagaaa ccttttggac tataaagacg atgacgataa gcatcaccat      660
caccatcact aa                                                                672
    
```


ES 2 477 543 T3

<210> 13
 <211> 672
 <212> ADN
 <213> Homo sapiens

5

<400> 13

```

atgaaaacat tcatactcct gctctgggta ctgctgctct gggttatctt cctgcttccc      60
ggtgccactg ctcagcctga gcgcgactgc cgagtgagca gcttccgagt caaggagaac      120
ttcgacaagg ctcgcttctc tgggacctgg tacgccatgg ccaagaagga ccccagagggc      180
ctctttctgc aggacaacat cgtcgcggag ttctccgtgg acgagaccgg ccagatgagc      240
gccacagcca agggccgagt ccgtcttttg aataactggg acgtgtgctc agacatggtg      300
ggcaccttca cagacaccga ggaccctgcc aagttcaaga tgaagtactg gggcgtagcc      360
tcctagctcc agaaaggaaa tgatgaccac tggatcgtcg acacagacta cgacacgtat      420
gccgtgcagt actcctgccg cctcctgaac ctcgatggca cctgtgctga cagctactcc      480
ttcgtgtttt cccgggaccc caacggcctg cccccagaag cgcagaagat tgtaaggcag      540
cggcaggagg agctgtgcct ggccaggcag tacaggctga tcgtccacaa cggttactgc      600
gatggcagat cagaaagaaa ccttttggac tataaagacg atgacgataa gcatcaccat      660
caccatcact aa                                                                672
    
```

10 <210> 14
 <211> 672
 <212> ADN
 <213> Homo sapiens

15 <400> 14

```

atgaaaacat tcatactcct gctctgggta ctgctgctct gggttatctt cctgcttccc      60
ggtgccactg ctcagcctga gcgcgactgc cgagtgagca gcttccgagt caaggagaac      120
ttcgacaagg ctcgcttctc tgggacctgg tacgccatgg ccaagaagga ccccagagggc      180
ctctttctgc aggacaacat cgtcgcggag ttctccgtgg acgagaccgg ccagatgagc      240
gccacagcca agggccgagt ccgtcttttg aataactggg acgtgtgctc agacatggtg      300
ggcaccttca cagacaccga ggaccctgcc aagttcaaga tgaagtactg gggcgtagcc      360
tcctttctcc agaaaggaaa tgatgaccac tggatcgtcg acacagacta cgacacgtag      420
gccgtgcagt actcctgccg cctcctgaac ctcgatggca cctgtgctga cagctactcc      480
ttcgtgtttt cccgggaccc caacggcctg cccccagaag cgcagaagat tgtaaggcag      540
cggcaggagg agctgtgcct ggccaggcag tacaggctga tcgtccacaa cggttactgc      600
gatggcagat cagaaagaaa ccttttggac tataaagacg atgacgataa gcatcaccat      660
caccatcact aa                                                                672
    
```

ES 2 477 543 T3

<210> 15
 <211> 672
 <212> ADN
 <213> Homo sapiens

5

<400> 15

```

atgaaaacat tcatactcct gctctgggta ctgctgctct gggttatctt cctgcttccc      60
ggtgccactg ctcagcctga ggcgcactgc cgagtgagca gcttccgagt caaggagAAC      120
ttcgacaagg ctgccttctc tgggacctgg tacgccatgg ccaagaagga ccccgagggc      180
ctctttctgc aggacaacat cgtcgcggag ttctccgtgg acgagaccgg ccagatgagc      240
gccacagcca agggccgagt ccgtcttttg aataactggg acgtgtgcbc agacatggtg      300
ggcaccttca cagacaccga ggacctgcc aagttcaaga tgaagtactg gggcgtagcc      360
tcctttctcc agaaaggaaa tgatgaccac tggatcgtcg acacagacta cgacacgtat      420
gccgtgcagt actcctgccg cctcctgaac ctcgatggca cctgtgctga cagctactcc      480
ttcgtgtttt cccgggaccc caacggcctg ccccagaag cgcagaagat tgtaaggcag      540
cggcaggagg agctgtgcct ggccaggcag tagaggctga tcgtccacaa cggttactgc      600
gatggcagat cagaaagaaa ccttttggac tataaagacg atgacgataa gcatcaccat      660
caccatcact aa                                                                672
    
```

10 <210> 16
 <211> 672
 <212> ADN
 <213> Homo sapiens

15 <400> 16

```

atgaaaacat tcatactcct gctctgggta ctgctgctct gggttatctt cctgcttccc      60
ggtgccactg ctcagcctga ggcgcactgc cgagtgagca gcttccgagt caaggagAAC      120
ttcgacaagg ctgccttctc tgggacctgg tacgccatgg ccaagaagga ccccgagggc      180
ctctttctgc aggacaacat cgtcgcggag ttctccgtgg acgagaccgg ccagatgagc      240
gccacagcca agggccgagt ccgtcttttg aataactggg acgtgtgcbc agacatggtg      300
ggcaccttca cagacaccga ggacctgcc aagttcaaga tgaagtactg gggcgtagcc      360
tcctttctcc agaaaggaaa tgatgaccac tggatcgtcg acacagacta cgacacgtat      420
gccgtgcagt actcctgccg cctcctgaac ctcgatggca cctgtgctga cagctactcc      480
ttcgtgtttt cccgggaccc caacggcctg ccccagaag cgcagaagat tgtaaggcag      540
cggcaggagg agctgtgcct ggccaggcag tacaggctga tcgtccacaa cggttagtgc      600
gatggcagat cagaaagaaa ccttttggac tataaagacg atgacgataa gcatcaccat      660
caccatcact aa                                                                672
    
```

ES 2 477 543 T3

<210> 17
 <211> 10
 <212> PRT
 <213> Homo sapiens
 5
 <220>
 <221> MOD_RES
 <222> (7) .. (7)
 <223> Xaa es un resto PCL
 10
 <400> 17

 Tyr Trp Gly Val Ala Ser Xaa Leu Gln Lys
 1 5 10

 15 <210> 18
 <211> 29
 <212> PRT
 <213> Homo sapiens
 20 <220>
 <221> MOD_RES
 <222> (7) .. (7)
 <223> Xaa es un resto Cyc
 25 <400> 18

 Lys Asp Pro Glu Gly Leu Xaa Leu Gln Asp Asn Ile Val Ala Glu Phe
 1 5 10 15

 Ser Val Asp Glu Thr Gly Gln Met Ser Ala Thr Ala Lys

 20 25

 30 <210> 19
 <211> 166
 <212> PRT
 <213> Homo sapiens

 35 <400> 19

ES 2 477 543 T3

Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu
 1 5 10 15

Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His
 20 25 30

Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
 35 40 45

Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp
 50 55 60

Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu
 65 70 75 80

Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp
 85 90 95

Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu
 100 105 110

Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
 115 120 125

Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val
 130 135 140

Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
 145 150 155 160

Cys Arg Thr Gly Asp Arg
 165

<210> 20
 <211> 166
 5 <212> PRT
 <213> ratón
 <400> 20

ES 2 477 543 T3

Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu
1 5 10 15

Leu Glu Ala Lys Glu Ala Glu Asn Val Thr Met Gly Cys Ala Glu Gly
20 25 30

Pro Arg Leu Ser Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
35 40 45

Tyr Ala Trp Lys Arg Met Glu Val Glu Glu Gln Ala Ile Glu Val Trp
50 55 60

Gln Gly Leu Ser Leu Leu Ser Glu Ala Ile Leu Gln Ala Gln Ala Leu
65 70 75 80

Leu Ala Asn Ser Ser Gln Pro Pro Glu Thr Leu Gln Leu His Ile Asp
85 90 95

Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Ser Leu Leu Arg Val Leu
100 105 110

Gly Ala Gln Lys Glu Leu Met Ser Pro Pro Asp Thr Thr Pro Pro Ala
115 120 125

Pro Leu Arg Thr Leu Thr Val Asp Thr Phe Cys Lys Leu Phe Arg Val
130 135 140

Tyr Ala Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Val
145 150 155 160

Cys Arg Arg Gly Asp Arg
165

<210> 21

<211> 19

5 <212> PRT

<213> Secuencia artificial

<220>

<223> construcción sintética

10

<400> 21

Met Gly Asp Ser Lys Ile His His His His His His Glu Asn Leu Tyr
1 5 10 15

Phe Gln Gly

15

<210> 22
<211> 301
<212> PRT
<213> Homo sapiens

5

<400> 22

ES 2 477 543 T3

Met Gly Ser Asp Lys Ile His His His His His His Glu Asn Leu Tyr
 1 5 10 15

Phe Gln Gly Ser Leu Leu Val Asn Pro Glu Gly Pro Thr Leu Met Arg
 20 25 30

Leu Asn Ser Val Gln Ser Ser Glu Arg Pro Leu Phe Leu Val His Pro
 35 40 45

Ile Glu Gly Ser Thr Thr Val Phe His Ser Leu Ala Ser Arg Leu Ser
 50 55 60

Ile Pro Thr Tyr Gly Leu Gln Cys Thr Arg Ala Ala Pro Leu Asp Ser
 65 70 75 80

Ile His Ser Leu Ala Ala Tyr Tyr Ile Asp Cys Ile Arg Gln Val Gln
 85 90 95

Pro Glu Gly Pro Tyr Arg Val Ala Gly Tyr Ser Tyr Gly Ala Cys Val
 100 105 110

Ala Phe Glu Met Cys Ser Gln Leu Gln Ala Gln Gln Ser Pro Ala Pro
 115 120 125

Thr His Asn Ser Leu Phe Leu Phe Asp Gly Ser Pro Thr Tyr Val Leu
 130 135 140

Ala Tyr Thr Gln Ser Tyr Arg Ala Lys Leu Thr Pro Gly Ser Glu Ala
 145 150 155 160

Glu Ala Glu Thr Glu Ala Ile Cys Phe Phe Val Gln Gln Phe Thr Asp
 165 170 175

Met Glu His Asn Arg Val Leu Glu Ala Leu Leu Pro Leu Lys Gly Leu
 180 185 190

Glu Glu Arg Val Ala Ala Ala Val Asp Leu Ile Ile Lys Ser His Gln
 195 200 205

Gly Leu Asp Arg Gln Glu Leu Ser Phe Ala Ala Arg Ser Phe Tyr Tyr

ES 2 477 543 T3

210

215

220

Lys Leu Arg Ala Ala Glu Gln Tyr Thr Pro Lys Ala Lys Tyr His Gly
 225 230 235 240

Asn Val Met Leu Leu Arg Ala Lys Thr Gly Gly Ala Tyr Gly Glu Asp
 245 250 255

Leu Gly Ala Asp Tyr Asn Leu Ser Gln Val Cys Asp Gly Lys Val Ser
 260 265 270

Val His Val Ile Glu Gly Asp His Arg Thr Leu Leu Glu Gly Ser Gly
 275 280 285

Leu Glu Ser Ile Ile Ser Ile Ile His Ser Ser Leu Ala
 290 295 300

<210> 23

<211> 18

5 <212> PRT

<213> Secuencia artificial

<220>

<223> construcción sintética

10

<400> 23

Met Gly Ser Ser His His His His His His Leu Glu Val Leu Phe Gln
 1 5 10 15

Gly Pro

15 <210> 24

<211> 125

<212> PRT

<213> Homo sapiens

20 <400> 24

Met Gly Ser Ser His His His His His His Leu Glu Val Leu Phe Gln
 1 5 10 15

Gly Pro Gly Val Gln Val Glu Thr Ile Ser Pro Gly Asp Gly Arg Thr
 20 25 30

Phe Pro Lys Arg Gly Gln Thr Cys Val Val His Tyr Thr Gly Met Leu
 35 40 45

Glu Asp Gly Lys Lys Phe Asp Ser Ser Arg Asp Arg Asn Lys Pro Phe

ES 2 477 543 T3

50

55

60

Lys Phe Met Leu Gly Lys Gln Glu Val Ile Arg Gly Trp Glu Glu Gly
65 70 75 80

Val Ala Gln Met Ser Val Gly Gln Arg Ala Lys Leu Thr Ile Ser Pro
85 90 95

Asp Tyr Ala Tyr Gly Ala Thr Gly His Pro Gly Ile Ile Pro Pro His
100 105 110

Ala Thr Leu Val Phe Asp Val Glu Leu Leu Lys Leu Glu
115 120 125

<210> 25

<211> 196

5 <212> PRT

<213> Homo sapiens

<400> 25

Met Gly Ser Ser His His His His His His Ser Ser Gly Glu Asn Leu
1 5 10 15

Tyr Phe Gln Gly Asp Ser Ser Pro Leu Leu Gln Phe Gly Gly Val Arg
20 25 30

Gln Arg Tyr Leu Tyr Thr Asp Asp Ala Gln Gln Thr Glu Ala His Leu
35 40 45

Glu Ile Arg Glu Asp Gly Thr Val Gly Gly Ala Ala Asp Gln Ser Pro
50 55 60

Glu Ser Leu Leu Gln Leu Lys Ala Leu Lys Pro Gly Val Ile Gln Ile
65 70 75 80

Leu Gly Val Lys Thr Ser Arg Phe Leu Cys Gln Arg Pro Asp Gly Ala
85 90 95

Leu Tyr Gly Ser Leu His Phe Asp Pro Glu Ala Cys Ser Phe Arg Glu
100 105 110

Leu Leu Leu Glu Asp Gly Tyr Asn Val Tyr Gln Ser Glu Ala His Gly
115 120 125

Leu Pro Leu His Leu Pro Gly Asn Lys Ser Pro His Arg Asp Pro Ala
130 135 140

10

ES 2 477 543 T3

Pro Arg Gly Pro Ala Arg Phe Leu Pro Leu Pro Gly Leu Pro Pro Ala
145 150 155 160

Leu Pro Glu Pro Pro Gly Ile Leu Ala Pro Gln Pro Pro Asp Val Gly
165 170 175

Ser Ser Asp Pro Leu Ser Met Val Gly Pro Ser Gln Gly Arg Ser Pro
180 185 190

Ser Tyr Ala Ser
195

<210> 26
<211> 173
<212> PRT
<213> ratón

5

<400> 26

Met Arg Gly Ser His His His His His His Gly Ser Gly Ile Glu Gly
1 5 10 15

Arg Leu Arg Ser Ser Ser Gln Asn Ser Ser Asp Lys Pro Val Ala His
20 25 30

Val Val Ala Asn His Gln Val Glu Glu Gln Leu Trp Leu Ser Gln Arg
35 40 45

Ala Asn Ala Leu Leu Ala Asn Gly Met Asp Leu Lys Asp Asn Gln Leu
50 55 60

Val Val Pro Ala Asp Gly Leu Tyr Leu Val Tyr Ser Gln Val Leu Phe
65 70 75 80

Lys Gly Gln Gly Cys Pro Asp Tyr Val Leu Leu Thr His Thr Val Ser
85 90 95

Arg Phe Ala Ile Ser Tyr Gln Glu Lys Val Asn Leu Leu Ser Ala Val
100 105 110

Lys Ser Pro Cys Pro Lys Asp Thr Pro Glu Gly Ala Glu Leu Lys Pro
115 120 125

Trp Tyr Glu Pro Ile Tyr Leu Gly Gly Val Phe Gln Leu Glu Lys Gly
130 135 140

Asp Gln Leu Ser Ala Glu Val Asn Leu Pro Lys Tyr Leu Asp Phe Ala

10

ES 2 477 543 T3

145 150 155 160

Glu Ala Ser Gly Gln Val Tyr Phe Gly Val Ile Ala Leu
 165 170

5 <210> 27
 <211> 62
 <212> PRT
 <213> ratón
 <400> 27

Met Asn Ser Tyr Pro Gly Cys Pro Ser Ser Tyr Asp Gly Tyr Cys Leu
 1 5 10 15

Asn Gly Gly Val Cys Met His Ile Glu Ser Leu Asp Ser Tyr Thr Cys
 20 25 30

Asn Cys Val Ile Gly Tyr Ser Gly Asp Arg Cys Gln Thr Arg Asp Leu
 35 40 45

10 Arg Trp Trp Glu Leu Arg Leu Glu His His His His His His
 50 55 60

<210> 28
 <211> 14
 <212> PRT
 15 <213> Secuencia artificial
 <220>
 <223> construcción sintética

20 <220>
 <221 > MOD_RES
 <222> (2) .. (2)
 <223> D-alanina

25 <220>
 <221 > MOD_RES
 <222> (4) .. (4)
 <223> ciclohexil-alanina

30 <220>
 <221 > MOD_RES
 <222> (4) .. (4)
 <223> ciclohexil-alanina

35 <220>
 <221 > MOD_RES
 <222> (13) .. (13)
 <223> D-alanina

40 <400> 28

Gly Xaa Lys Xaa Val Ala Ala Trp Thr Leu Lys Ala Xaa Gly
 1 5 10

<210> 29
 <211> 19
 <212> PRT
 <213> Secuencia artificial
 5
 <220>
 <223> construcción sintética
 <220>
 10 <221 > MOD_RES
 <222> (7)..(7)
 <223> D-alanina
 <220>
 15 <221 > MOD_RES
 <222> (9) .. (9)
 <223> ciclohexil-alanina
 <220>
 20 <221 > MOD_RES
 <222> (9) .. (9)
 <223> ciclohexil-alanina
 <220>
 25 <221 > MOD_RES
 <222> (18)..(18)
 <223> D-alanina
 <400> 29
 30
 Ala Gly Ser Arg Ser Gly Xaa Lys Xaa Val Ala Ala Trp Thr Leu Lys
 1 5 10 15
 Ala Xaa Gly
 <210> 30
 <211> 20
 35 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> construcción sintética
 40
 <220>
 <221> misc_feature
 <222> (1) .. (20)
 <223> estructura de fosfotioato
 45
 <400> 30
 tccatgacgt tctgacgt 20
 <210> 31
 <211> 22
 50 <212> ADN
 <213> Secuencia artificial
 <220>
 55 <223> construcción sintética
 <220>
 <221> misc_feature
 <222> (1) .. (22)
 60 <223> estructura de fosfotioato

ES 2 477 543 T3

<400> 31
 tcgtcgtttt cggcgcgcgc cg 22

5 <210> 32
 <211> 22
 <212> PRT
 <213> ratón

10 <220>
 <221 > MOD_RES
 <222> (11)..(11)
 <223> PCL

15 <400> 32

Met	Asn	Ser	Tyr	Pro	Gly	Cys	Pro	Ser	Ser	Xaa	Asp	Gly	Tyr	Cys	Leu
1				5					10					15	

Asn Gly Gly Val Cys Met
 20

20 <210> 33
 <211> 22
 <212> PRT
 <213> ratón

25 <220>
 <221 > MISC_FEATURE
 <222> (11) .. (11)
 <223> pirrolisina

<400> 33

Met	Asn	Ser	Tyr	Pro	Gly	Cys	Pro	Ser	Ser	Xaa	Asp	Gly	Tyr	Cys	Leu
1				5					10					15	

Asn Gly Gly Val Cys Met
 20

30

35 <210> 34
 <211> 14
 <212> PRT
 <213> ratón

40 <220>
 <221 > MISC_FEATURE
 <222> (3) .. (3)
 <223> pirrolisina

<400> 34

Asn	His	Xaa	Val	Glu	Glu	Gln	Leu	Glu	Trp	Leu	Ser	Gln	Arg
1				5					10				

45

50 <210> 35
 <211> 14
 <212> PRT
 <213> ratón

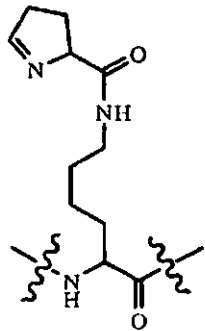
REIVINDICACIONES

1. Un compuesto que tiene la estructura de la Fórmula (II):

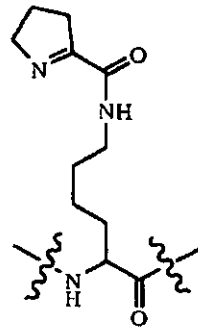


donde:

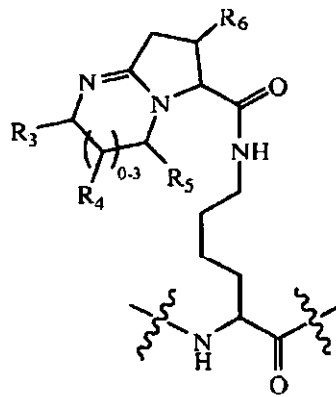
10 R_1 es H o un grupo de modificación del amino terminal; R_2 es OH o un grupo de modificación del carboxi terminal;
 n es un número entero de 1 a 5000;
cada BB se selecciona de forma independiente entre un resto de aminoácido, un resto de aminoácido análogo de
pirrolisina que tiene la estructura de la Fórmula (A-2), un resto de aminoácido análogo de pirrolisina que tiene la
estructura de la Fórmula (B-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula
15 (C-1) un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (D-1), un resto de
aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (E-1), un resto de aminoácido análogo de
pirrolisina que tiene la estructura de la Fórmula (F-1), un resto de aminoácido análogo de pirrolisina que tiene la
estructura de la Fórmula (G-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la
Fórmula (H-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (I-1), un resto
de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (J-1), un resto de aminoácido análogo
20 de pirrolisina que tiene la estructura de la Fórmula (K-1) y un resto de aminoácido análogo de pirrolisina que tiene
la estructura de la Fórmula (L-1);



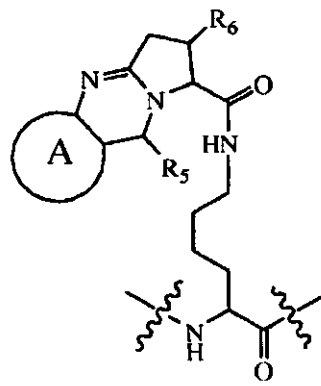
(A-2)



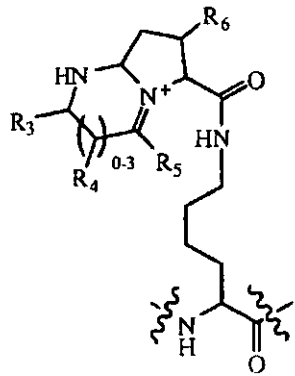
(B-2)



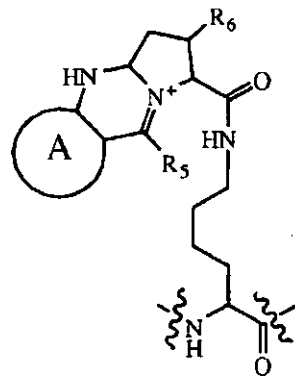
(C-1)



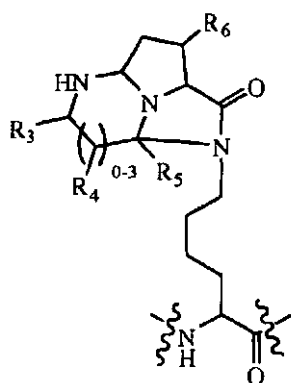
(D-1)



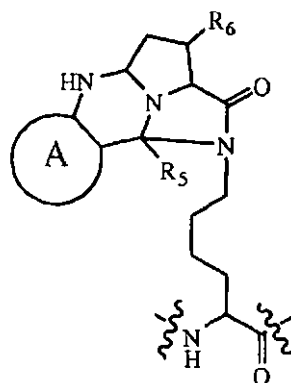
(E-1)



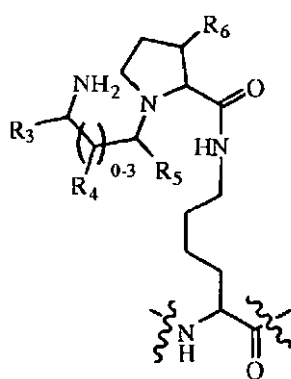
(F-1)



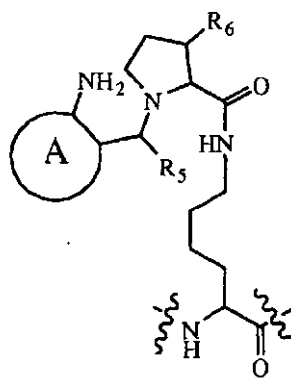
(G-1)



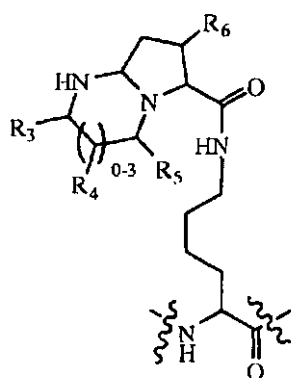
(H-1)



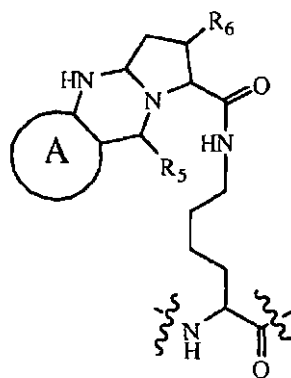
(I-1)



(J-1)



(K-1)



(L-1)

donde:

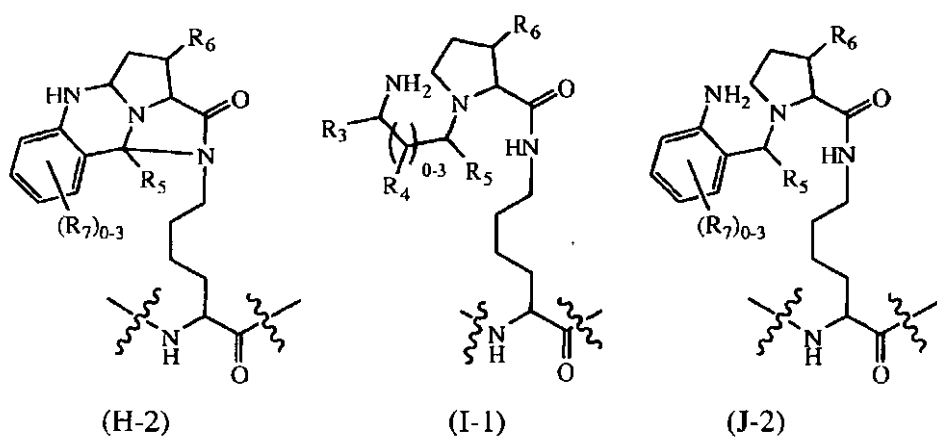
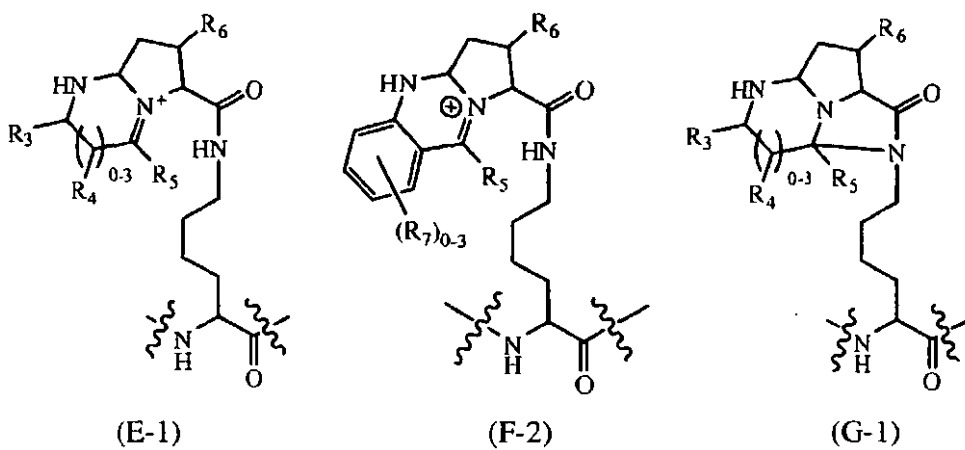
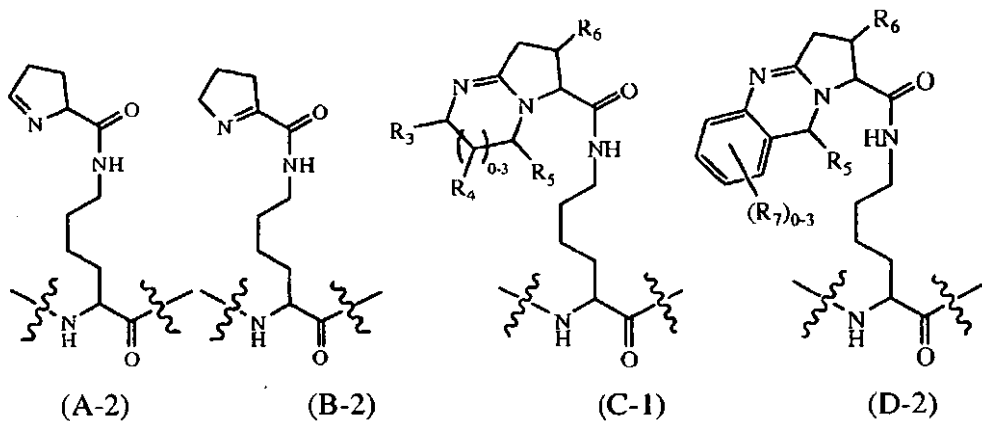
- 5 R₃, R₅ y cada R₄ se seleccionan independientemente entre H, -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidrox-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹;
 R₆ es H o alquilo C₁;
 A es un cicloalquilo C₃₋₈, heterocicloalquilo C₃₋₈, un arilo monocíclico de 5-6 miembros, un heteroarilo monocíclico
 10 de 5-6 miembros, un anillo bicíclico condensado de 9-10 miembros o un anillo tricíclico condensado de 13-14 miembros, en donde A está opcionalmente sustituido con 1 a 5 sustituyentes seleccionados independientemente entre -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidrox-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹;
 L se selecciona entre un enlace, alquileno C₁₋₈, alquileno C₁₋₈ halo-sustituido, alquileno C₁₋₈ hidrox-sustituido,

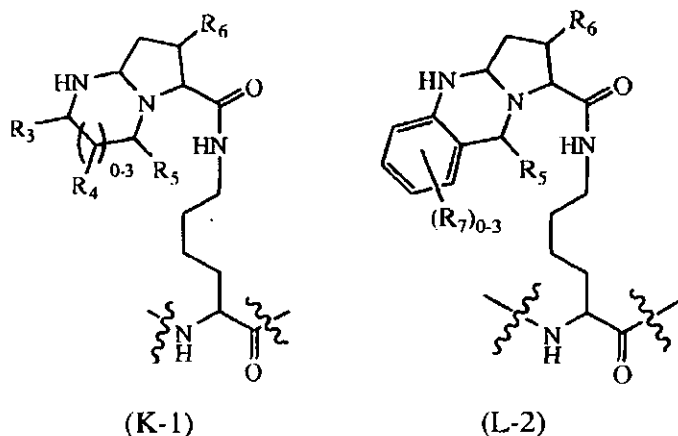
alquilenilo C_{2-8} , alquilenilo C_{2-8} halo-sustituido, alquilenilo $-C_{2-8}$ halo-sustituido, $-O(CR^{11}R^{12})_k-$, $-S(CR^{11}R^{12})_k-$,
 $-S(O)_k(CR^{11}R^{12})_k-$, $-O(CR^{11}R^{12})_k-NR^{11}C(O)-$, $-O(CR^{11}R^{12})_kC(O)NR^{11}-$, $-C(O)-$, $-C(O)(CR^{11}R^{12})_k-$, $-C(S)-$,
 $-C(S)(CR^{11}R^{12})_k-$, $-C(O)NR^{11}-$, $-NR^{11}C(O)-$, $-NR^{11}(CR^{11}R^{12})_k-$, $CONR^{11}(CR^{11}R^{12})_k-$, $-N(R^{11})CO(CR^{11}R^{12})_k-$,
 $-C(O)NR^{11}(CR^{11}R^{12})_k-$, $-NR^{11}C(O)(CR^{11}R^{12})_k-$, en donde cada R^{11} y R^{12} son independientemente H, alquilo C_{1-8} ,
 5 alquilo C_{1-8} halo-sustituido o alquilo C_{1-8} hidroxio-sustituido, y k es un número entero de 1 a 12, y
 X^1 se selecciona de una marca, un colorante, un polímero, un polímero soluble en agua, un polialquilenglicol, un
 poli(etilenglicol), un derivado de poli(etilenglicol), un azúcar, un lípido, un fotorreticulador, un compuesto
 citotóxico, un fármaco, una marca de afinidad, una marca de fotoafinidad, un compuesto reactivo; una resina, un
 péptido, una segunda proteína o un polipéptido o un análogo de polipéptido, un anticuerpo o un fragmento de
 10 anticuerpo, un quelante metálico, un cofactor, un ácido graso, un hidrato de carbono, un polinucleótido, un ADN,
 un ARN, una sonda de la PCR, un polinucleótido antisentido, un ribooligonucleótido, un desoxirribonucleótido,
 ADN modificado con fosfortioato, ADN y ARN modificados, ácido nucleico peptídico, un sacárido, un disacárido,
 un oligosacárido, un polisacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial, una
 15 nanopartícula, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un grupo
 funcional novedoso, un grupo que interactúa de forma covalente o no covalente con otras moléculas, un resto
 fotoestimulado, un resto excitable mediante radiación actínica, un ligando un resto fotoisomerizable, biotina, un
 análogo de biotina, un resto que incorpora un átomo pesado, un grupo escindible químicamente, un grupo
 fotoescindible, una cadena lateral alargada, un azúcar unido a carbono, un agente redox activo, un
 20 aminotioácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un
 grupo cromóforo, un grupo quimioluminiscente, un resto fluorescente, un grupo denso en electrones, un grupo
 magnético, un grupo intercalante, un grupo quelante, un cromóforo, un agente de transferencia de energía, un
 agente biológicamente activo, una marca detectable, una molécula pequeña, un ácido ribonucleico inhibidor, un
 ARNip, un radionucleótido, un agente de captura de neutrones, un derivado de biotina, punto(s) cuántico(s), un
 25 nanotransmisor, un radiotransmisor, una abzima, una enzima, un activador complejo activado, un virus, una
 toxina, un adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un
 agonista de TLR8, un epítipo de linfocito T, un fosfolípido, una molécula de tipo LPS, hemocianina de lapa
 californiana (KLH), un hapteno inmunógeno, un aglicano, un alérgeno, una angioestatina, una antihormona, un
 antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimotopo,
 30 un receptor, una micela inversa, detergentes, potenciadores de la respuesta inmune, colorantes fluorescentes,
 reactivos FRET, sondas de radioimagen, otra sonda de espectroscopía, profármacos, toxinas para
 inmunoterapia, un soporte sólido, $-CH_2CH_2-(OCH_2CH_2O)_p-OX^2$, $-O-(CH_2CH_2O)_pCH_2CH_2-X^2$, y cualquier
 combinación de los mismos, donde p es de 1 a 10.000 y X^2 es H, un alquilo C_{1-8} , un grupo protector o un grupo
 funcional terminal, y

35 donde al menos un BB es un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (C-1) o la
 Fórmula (D-1) o la Fórmula (E-1) o la Fórmula (F-1) o la Fórmula (G-1) o la Fórmula (H-1) o la Fórmula (I-1) o la
 Fórmula (J-1) o la Fórmula (K-1) o la Fórmula (L-1).

40 2. El compuesto de la reivindicación 1, donde el anillo A se selecciona entre fenilo, naftaleno y piridina.

3. El compuesto de la reivindicación 1, donde cada BB se selecciona de forma independiente entre un resto de
 aminoácido, un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (A-2), un resto de
 aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (B-2), un resto de aminoácido análogo de
 45 pirrolisina que tiene la estructura de la Fórmula (C-1) un resto de aminoácido análogo de pirrolisina que tiene la
 estructura de la Fórmula (D-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula
 (E-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (F-2), un resto de aminoácido
 análogo de pirrolisina que tiene la estructura de la Fórmula (G-1), un resto de aminoácido análogo de pirrolisina que
 tiene la estructura de la Fórmula (H-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la
 Fórmula (I-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (J-2), un resto de
 50 aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (K-1) y un resto de aminoácido análogo de
 pirrolisina que tiene la estructura de la Fórmula (L-2);





donde,

- 5 R_3 , R_5 y cada R_4 se seleccionan independientemente entre H, -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituído, alquilo C₁₋₈ hidroxi-sustituído, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹;
 R_6 es H o alquilo C₁;
cuando está presente cada R_7 se selecciona independientemente entre -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈
halo-sustituído, alquilo C₁₋₈ hidroxi-sustituído, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹;
- 10 L se selecciona entre un enlace, alquilenilo C₁₋₈, alquilenilo C₁₋₈ halo-sustituído, alquilenilo C₁₋₈ hidroxi-sustituído,
alquenileno C₂₋₈, alquenileno C₂₋₈ halo-sustituído, alquenileno C₂₋₈ hidroxi-sustituído, un polialquilenglicol, un
poli(etilenglicol), -O(CR¹¹R¹²)_k-, -S(CR¹¹R¹²)_k-, -S(O)_k(CR¹¹R¹²)_k-, -O(CR¹¹R¹²)_k-NR¹¹C(O)-,
-O(CR¹¹R¹²)_kC(O)NR¹¹-, -C(O)-, -C(O)(CR¹¹R¹²)_k-, -C(S)-, -C(S)(CR¹¹R¹²)_k-, -C(O)NR¹¹-, -NR¹¹C(O)-,
- NR¹¹(CR¹¹R¹²)_k-, -CONR¹¹(CR¹¹R¹²)_k-, -N(R¹¹)CO(CR¹¹R¹²)_k-, -C(O)NR¹¹(CR¹¹R¹²)_k-, -NR¹¹C(O)(CR¹¹R¹²)_k-, en
15 donde cada R¹¹ y R¹² son independientemente H, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituído, o alquilo C₁₋₈
hidroxi-sustituído, y k es un número entero de 1 a 12, y
X¹ se selecciona de una marca, un colorante, un polímero, un polímero soluble en agua, un polialquilenglicol, un
poli(etilenglicol), un derivado de poli(etilenglicol), un azúcar, un lípido, un fotorreticulador, un compuesto
20 citotóxico, un fármaco, una marca de afinidad, una marca de fotoafinidad, un compuesto reactivo; una resina, un
péptido, una segunda proteína o un polipéptido o un análogo de polipéptido o un anticuerpo o un fragmento de
anticuerpo, un quelante metálico, un cofactor, un ácido graso, un hidrato de carbono, un polinucleótido, un ADN,
un ARN, una sonda de la PCR, un polinucleótido antisentido, un ribooligonucleótido, un desoxirribonucleótido,
ADN modificado con fosforotioato, ADN y ARN modificados, un ácido nucleico peptídico, un sacárido, un
25 disacárido, un oligosacárido, un polisacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial,
una nanopartícula, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un grupo
funcional novedoso, un grupo que interactúa de forma covalente o no covalente con otras moléculas, un resto
fotoestimulado, un resto excitable mediante radiación actínica, un ligando, un resto fotoisomerizable, biotina, un
análogo de biotina, un resto que incorpora un átomo pesado, un grupo escindible químicamente, un grupo
30 fotoescindible, una cadena lateral alargada, un azúcar unido a carbono, un agente redox activo, un
aminioácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un
grupo cromóforo, un grupo quimioluminiscente, un resto fluorescente, un grupo denso en electrones, un grupo
magnético, un grupo intercalante, un grupo quelante, un cromóforo, un agente de transferencia de energía, un
agente biológicamente activo, una marca detectable, una molécula pequeña, un ácido ribonucleico inhibidor, un
ARNip, un radionucleótido, un agente de captura de neutrones, un derivado de biotina, punto(s) cuántico(s), un
35 nanotransmisor, un radiotransmisor, una abzima, una enzima, un activador complejo activado, un virus, una
toxina, un adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9,
un agonista de TLR8, un epítipo de linfocito T, un fosfolípido, una molécula de tipo LPS, hemocianina de lapa
californiana (KLH), un hapteno inmunógeno, un aglicano, un alérgeno, una angioestatina, una antihormona, un
antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimotopo,
40 un receptor, una micela inversa, detergentes, potenciadores de la respuesta inmune, colorantes fluorescentes,
reactivos FRET, sondas de radioimagen, otra sonda de espectroscopía, profármacos, toxinas para
inmunoterapia, un soporte sólido, -CH₂CH₂-(OCH₂CH₂O)_p-OX², -O-(CH₂CH₂O)_pCH₂CH₂-X², y cualquier
combinación de los mismos, donde p es de 1 a 10.000 y X² es H, un alquilo C₁₋₈, un grupo protector o un grupo
funcional terminal.
- 45 4. El compuesto de una cualquiera de las reivindicaciones 1-3, donde R₇ es LX¹.
- 50 5. El compuesto de la reivindicación 4, donde X¹ es un azúcar, un polietilenglicol, un resto fluorescente, un
inmunomodulador, un ácido ribonucleico, un ácido desoxirribonucleico, una proteína, un péptido, una biotina, un
fosfolípido, un agonista de TLR7, un hapteno inmunógeno o un soporte sólido.

6. El compuesto de la reivindicación 5, donde L es un poli(alquilenglicol), un poli(etilenglicol), alquileo C₁₋₈, un alquileo C₁₋₈ halo-sustituido o un alquileo C₁₋₈ hidroxí-sustituido.

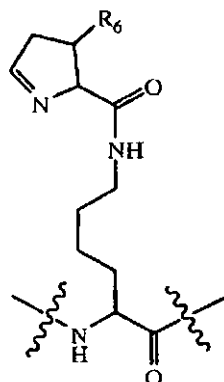
7. Un método para derivatizar una proteína, donde la proteína tiene la estructura de acuerdo con la Fórmula (I), comprendiendo el método poner en contacto la proteína con un reactivo de Fórmula (III) o de Fórmula (IV); donde la Fórmula (I) corresponde a:



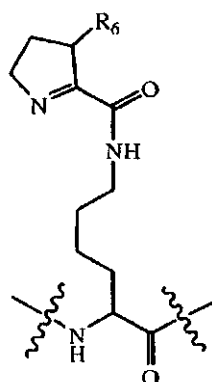
donde:

R₁ es H o un grupo de modificación del amino terminal;
R₂ es OH o un grupo de modificación del carboxi terminal;

n es un número entero de 1 a 5000;
cada AA se selecciona de forma independiente entre un resto de aminoácido, un resto del aminoácido pirrolisina, un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (A-1) y un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (B-1);

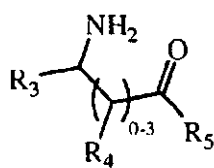


(A-1)

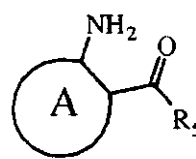


(B-1)

R₆ es H o alquilo C₁, y al menos un AA es un resto del aminoácido pirrolisina o un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (A-1) o de la Fórmula (B-1); y donde la Fórmula (III) y la Fórmula (IV) corresponden a:



(III)



(IV)

donde:

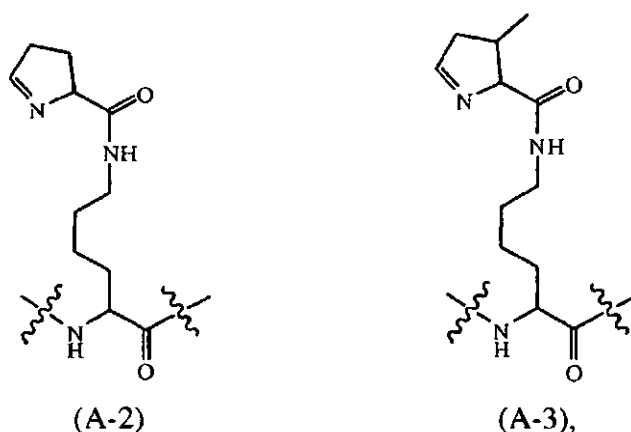
R₃, R₅ y cada R₄ se seleccionan independientemente entre H, -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidroxí-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹;

A es un cicloalquilo C₃₋₈, heterocicloalquilo C₃₋₈, un arilo monocíclico de 5-6 miembros, un heteroarilo monocíclico de 5-6 miembros, un anillo bicíclico condensado de 9-10 miembros o un anillo tricíclico condensado de 13-14 miembros, en donde A está opcionalmente sustituido con 1 a 5 sustituyentes seleccionados independientemente entre -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidroxí-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹;

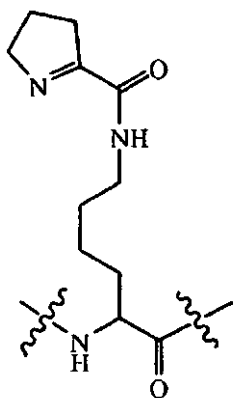
L se selecciona entre un enlace, alquileo C₁₋₈, alquileo C₁₋₈ halo-sustituido, alquileo C₁₋₈ hidroxí-sustituido, alquilenilo C₂₋₈, alquilenilo C₂₋₈ halo-sustituido, alquilenilo C₂₋₈ hidroxí-sustituido, poli(alquilenglicol), un poli(etilenglicol), -O(CR¹¹R¹²)_k-, -S(CR¹¹R¹²)_k-, -S(O)_k(CR¹¹R¹²)_k-, -O(CR¹¹R¹²)_k-NR¹¹C(O)-, -O(CR¹¹R¹²)_kC(O)NR¹¹-, -C(O)-, -C(O)(CR¹¹R¹²)_k-, -C(S)-, -C(S)(CR¹¹R¹²)_k-, -C(O)NR¹¹-, -NR¹¹C(O)-, -NR¹¹(CR¹¹R¹²)_k-, -CONR¹¹(CR¹¹R¹²)_k-, -N(R¹¹)CO(CR¹¹R¹²)_k-, -C(O)NR¹¹(CR¹¹R¹²)_k-, -NR¹¹C(O)(CR¹¹R¹²)_k-, en donde cada R¹¹ y R¹² son independientemente H, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido o alquilo C₁₋₈ hidroxí-sustituido, y k es un número entero de 1 a 12, y

X¹ se selecciona de una marca, un colorante, un polímero, un polímero soluble en agua, un polialquilenglicol, un poli(etilenglicol), un derivado de poli(etilenglicol), un azúcar, un lípido, un fotorreticulador, un compuesto citotóxico, un fármaco, una marca de afinidad, una marca de fotoafinidad, un compuesto reactivo; una resina, un péptido, una segunda proteína o un polipéptido o un análogo de polipéptido, un anticuerpo o un fragmento de anticuerpo, un quelante metálico, un cofactor, un ácido graso, un hidrato de carbono, un polinucleótido, un ADN, un ARN, una sonda de la PCR, un polinucleótido antisentido, un ribooligonucleótido, un desoxirribonucleótido, ADN modificado con fosforotioato, ADN y ARN modificados, ácido nucleico peptídico, un sacárido, un disacárido, un oligosacárido, un polisacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial, una nanopartícula, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un grupo funcional novedoso, un grupo que interactúa de forma covalente o no covalente con otras moléculas, un resto fotoestimulado, un resto excitable mediante radiación actínica, un ligando un resto fotoisomerizable, biotina, un análogo de biotina, un resto que incorpora un átomo pesado, un grupo escindible químicamente, un grupo fotoescindible, una cadena lateral alargada, un azúcar unido a carbono, un agente redox activo, un aminotioácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo cromóforo, un grupo quimioluminiscente, un resto fluorescente, un grupo denso en electrones, un grupo magnético, un grupo intercalante, un grupo quelante, un cromóforo, un agente de transferencia de energía, un agente biológicamente activo, una marca detectable, una molécula pequeña, un ácido ribonucleico inhibidor, un ARNip, un radionucleótido, un agente de captura de neutrones, un derivado de biotina, punto(s) cuántico(s), un nanotransmisor, un radiotransmisor, una abzima, una enzima, un activador complejo activado, un virus, una toxina, un adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, un epítipo de linfocito T, un fosfolípido, una molécula de tipo LPS, hemocianina de lapa californiana (KLH), un hapteno inmunógeno, un aglicano, un alérgeno, una angioestatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimotopo, un receptor, una micela inversa, detergentes, potenciadores de la respuesta inmune, colorantes fluorescentes, reactivos FRET, sondas de radioimagen, otra sonda de espectroscopía, profármacos, toxinas para inmunoterapia, un soporte sólido, $-\text{CH}_2\text{CH}_2-(\text{OCH}_2\text{CH}_2)_p-\text{OX}^2$,
 $-\text{O}-(\text{CH}_2\text{CH}_2\text{O})_p\text{CH}_2\text{CH}_2-\text{X}^2$, y cualquier combinación de los mismos, donde p es de 1 a 10.000 y X² es H, un alquilo C₁₋₈, un grupo protector o un grupo funcional terminal.

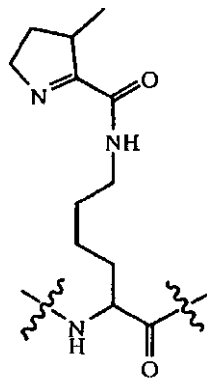
8. El método de la reivindicación 7, donde el resto de aminoácido de la Fórmula (A-1) es un resto de aminoácido que tiene la estructura de la Fórmula (A-2) o de la Fórmula (A-3):



y donde el resto de aminoácido de la Fórmula (B-1) es un resto de aminoácido que tiene la estructura de la Fórmula (B-2) o de la Fórmula (B-3):



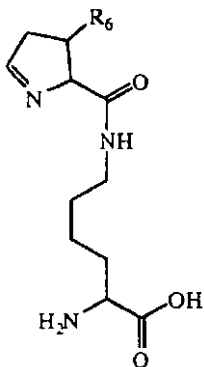
(B-2)



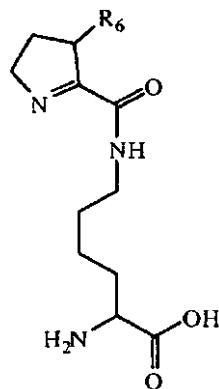
(B-3).

9. El método de reivindicación 8, donde el resto de aminoácido de la Fórmula (A-1) es un resto de un aminoácido de Fórmula (V) y el resto de aminoácido de la Fórmula (B-1) es un resto de un aminoácido de Fórmula (VI):

5



(V)



(VI),

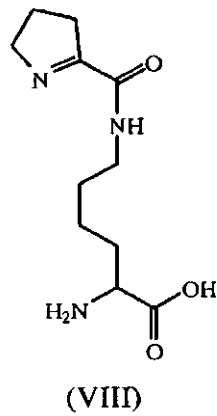
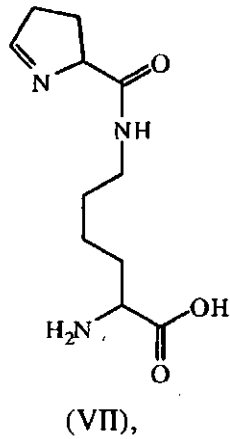
donde R₆ es H o alquilo C₁.

10. El método de reivindicación 9, donde el aminoácido de Fórmula (V) o de Fórmula (VI) se genera biosintéticamente en una célula que comprende:

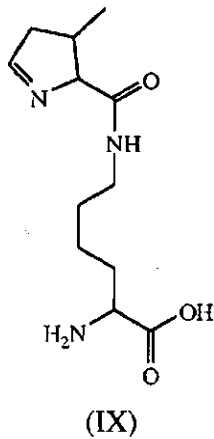
- i) un gen *pylC*, y un gen *pylD*, y la célula está en contacto con un medio de crecimiento que comprende un precursor o
- ii) un gen *pylB*, un gen *pylC* y un gen *pylD*, y la célula está en contacto con un medio de crecimiento que comprende un precursor.

11. El método de reivindicación 10, donde:

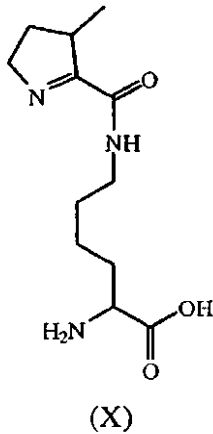
- i) el aminoácido de la Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (VII), y el aminoácido de la Fórmula (VI) es un aminoácido que tiene la estructura de la Fórmula (VIII)



5 y el precursor es D-ornitina, D-arginina, ácido (2S)-2-amino-6-(2,5-diaminopentanamido)hexanoico o ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico o ii) el aminoácido de la Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (IX)



10 y el precursor es D-ornitina, D-arginina, ácido 2,5-diamino-3-metilpentanoico, ácido (2S)-2-amino-6-(2,5-diaminopentanamido)hexanoico o ácido (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico. y el aminoácido de la Fórmula (V) es un aminoácido que tiene la estructura de la Fórmula (X)



15 y el precursor es D-ornitina, D-arginina, ácido 2,5-diamino-3-metilpentanoico, ácido (2R,3S)-2,5-diamino-3-metil-pentanoico, ácido (2R,3R)-2,5-diamino-3-metilpentanoico, ácido

(2S)-2-amino-6-(2,5-diaminopentanamido)hexanoico o ácido
 (2S)-2-amino-6-((R)-2,5-diaminopentanamido)hexanoico.

5 12. El método de una cualquiera de las reivindicaciones 10-11, donde el aminoácido de Fórmula (V), Fórmula (VI), Fórmula (VII), Fórmula (IX) o Fórmula (X) se incorpora a una proteína en la célula mediante un ARNt ortogonal (ARNt-O) y una aminoacil ARNt sintetasa ortogonal (O-RS), donde los aminoacilatos O-RS del ARNt-O con el aminoácido de Fórmula (V) o Fórmula (VI) y el ARNt-O reconocen al menos un codón selector de un ARNm en la célula y el codón selector es un codón ámbar TAG.

10 13. El método de una cualquiera de las reivindicaciones 10-11, donde la célula comprende además un gen pylS y un gen pylT y el aminoácido de Fórmula (V), Fórmula (VI), Fórmula (VII), Fórmula (IX) o Fórmula (X) se incorpora a una proteína en la célula mediante una aminoacil ARNt sintetasa y un ARNt que reconoce al menos un codón selector de un ARNm en la célula, donde la aminoacil ARNt sintetasa es un producto génico del gen pylS y el ARNt es un producto génico del gen pylT.

15 14. El método de cualquiera de las reivindicaciones 10-13, donde la célula es una célula procariota o una célula eucariota.

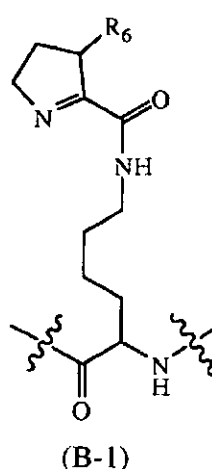
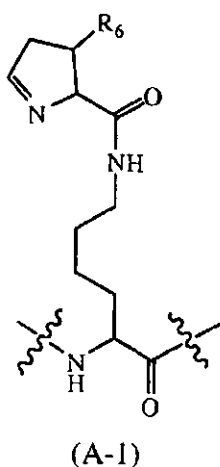
20 15. El método de reivindicación 14, donde la célula es una célula de Escherichia coli, una célula de mamífero, una célula de levadura, una célula de insecto, una célula CHO, una célula HeLa, una célula HEK293F o una célula sf9.

25 16. El método de cualquiera de las reivindicaciones 7-15, donde la proteína derivatizada tiene la estructura de acuerdo con la Fórmula (I):

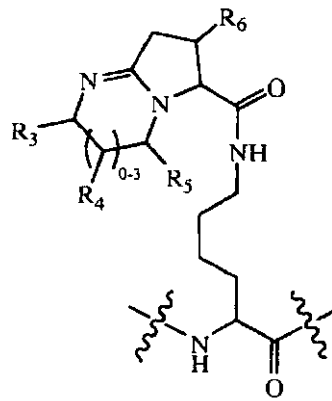
25
$$R_1(BB)_n-R_2 \quad (II)$$

donde:

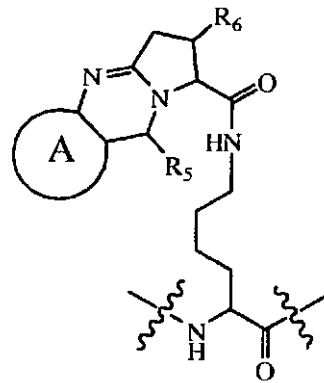
30 R_1 es H o un grupo de modificación del amino terminal;
 R_2 es OH o un grupo de modificación del carboxi terminal;
 n es un número entero de 1 a 5000;
 cada BB se selecciona de forma independiente entre un resto de aminoácido, un resto de aminoácido análogo de
 35 pirrolisina que tiene la estructura de la Fórmula (A-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (B-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (C-1) un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (D-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (E-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (F-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (G-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (H-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (I-1), un resto
 40 de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (J-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (K-1) y un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (L-1);



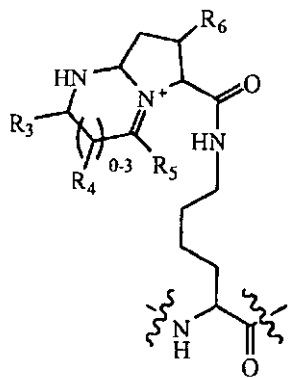
45



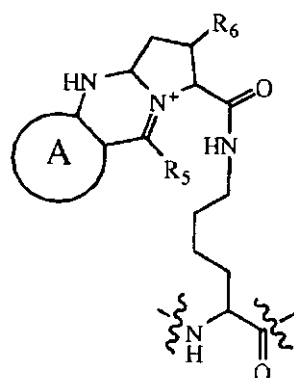
(C-1)



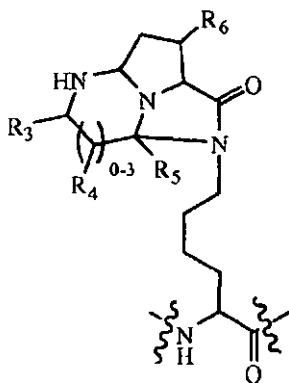
(D-1)



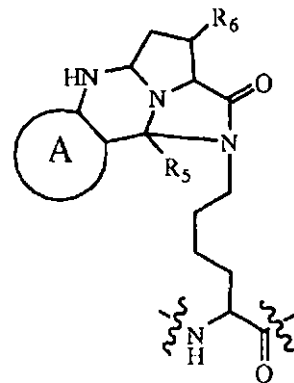
(E-1)



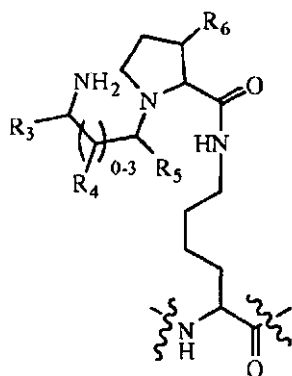
(F-1)



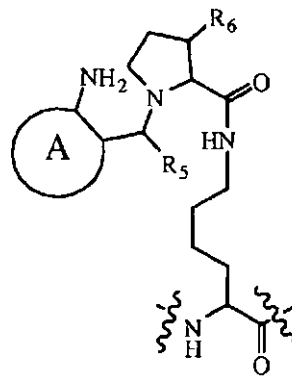
(G-1)



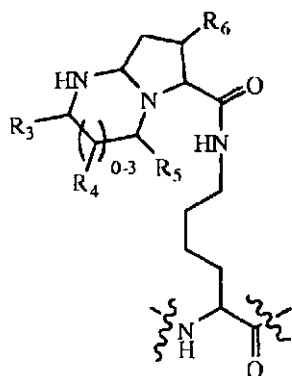
(H-1)



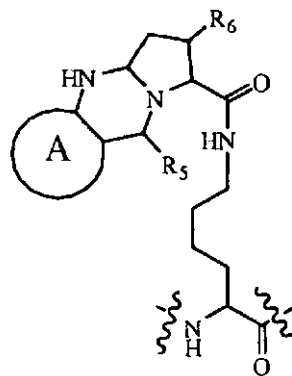
(I-1)



(J-1)



(K-1)



(L-1)

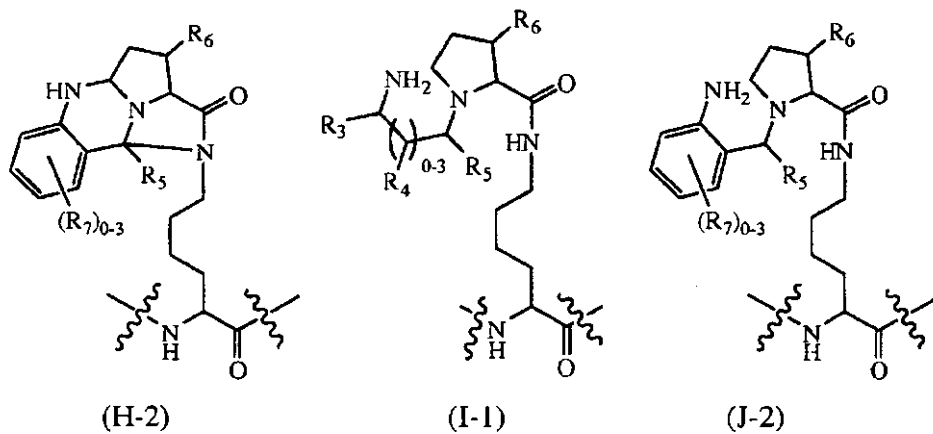
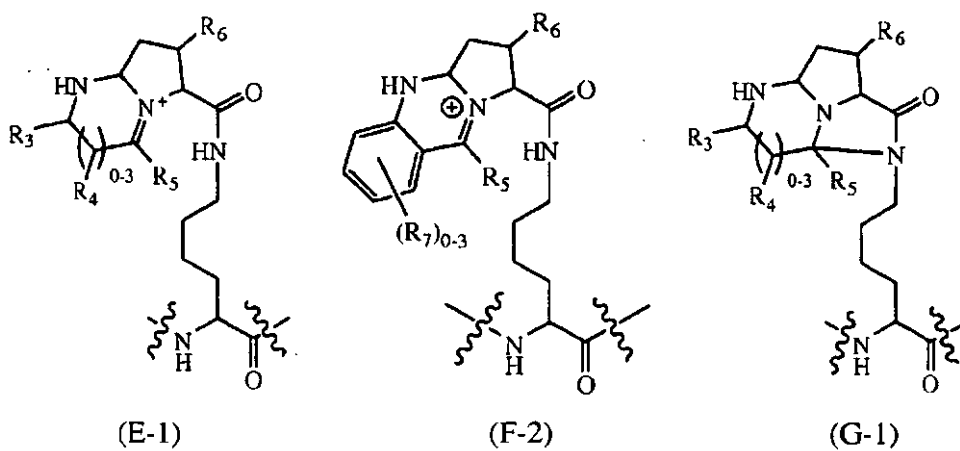
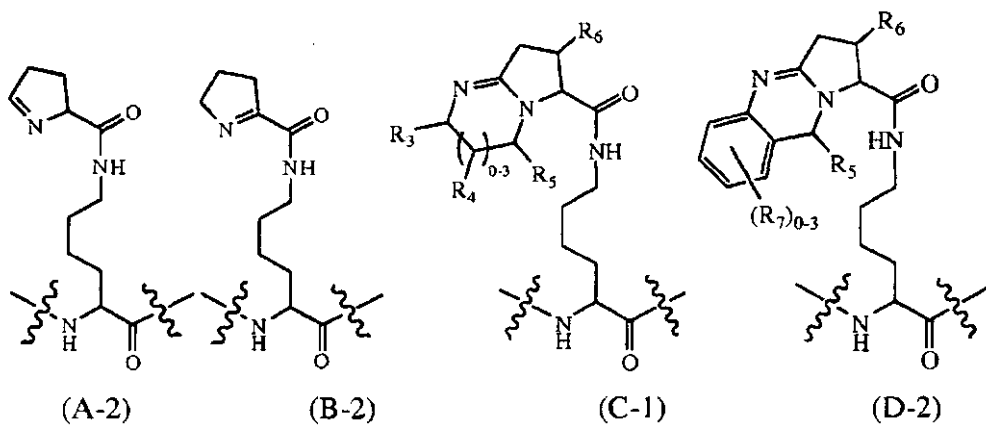
5 donde al menos un BB es un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (C-1) o la Fórmula (D-1) o la Fórmula (E-1) o la Fórmula (F-1) o la Fórmula (G-1) o la Fórmula (H-1) o la Fórmula (I-1) o la Fórmula (J-1) o la Fórmula (K-1) o la Fórmula (L-1).

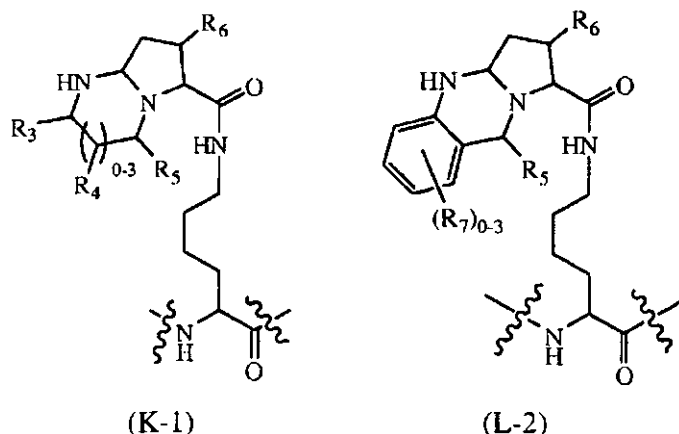
17. El método de reivindicación 16, donde el anillo A se selecciona entre fenilo, naftalilo y piridilo.

10 18. El método de reivindicación 16, donde cada BB se selecciona de forma independiente entre un resto de aminoácido, un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (A-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (B-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (C-1) un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (D-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (E-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (F-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (G-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (H-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (I-1), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (J-2), un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (K-1) y un resto de aminoácido análogo de pirrolisina que tiene la estructura de la Fórmula (L-2);

15

20





donde,

- 5 R_3 , R_5 y cada R_4 se seleccionan independientemente entre H, -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidroxi-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹; R_6 es H o alquilo C₁;
- cuando está presente cada R_7 se selecciona independientemente entre -OH, -NO₂, halo, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, alquilo C₁₋₈ hidroxi-sustituido, arilo, heteroarilo, heterocicloalquilo o cicloalquilo y -LX¹;
- 10 L se selecciona entre un enlace, alquilenilo C₁₋₈, alquilenilo C₁₋₈ halo-sustituido, alquilenilo C₁₋₈ hidroxi-sustituido, alquilenilo C₂₋₈, alquilenilo C₂₋₈ halo-sustituido, alquilenilo C₂₋₈ hidroxi-sustituido, un polialquilenglicol, un poli(etilenglicol), -O(CR¹¹R¹²)_k-, -S(CR¹¹R¹²)_k-, -S(O)_k(CR¹¹R¹²)_k-, -O(CR¹¹R¹²)_k-NR¹¹C(O)-, -O(CR¹¹R¹²)_kC(O)NR¹¹-, -C(O)-, -C(O)(CR¹¹R¹²)_k-, -C(S)-, -C(S)(CR¹¹R¹²)_k-, -C(O)NR¹¹-, -NR¹¹C(O)-, -NR¹¹(CR¹¹R¹²)_k-, -CONR¹¹(CR¹¹R¹²)_k-, -N(R¹¹)CO(CR¹¹R¹²)_k-, -C(O)NR¹¹(CR¹¹R¹²)_k-, NR¹¹C(O)(CR¹¹R¹²)_k-, en donde cada R¹¹ y R¹² son independientemente H, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, o alquilo C₁₋₈ hidroxi-sustituido, y k es un número entero de 1 a 12, y
- 15 X¹ se selecciona de una marca, un colorante, un polímero, un polímero soluble en agua, un polialquilenglicol, un poli(etilenglicol), un derivado de poli(etilenglicol), un azúcar, un lípido, un fotorreticulador, un compuesto citotóxico, un fármaco, una marca de afinidad, una marca de fotoafinidad, un compuesto reactivo; una resina, un péptido, una segunda proteína o un polipéptido o un análogo de polipéptido, un anticuerpo o un fragmento de anticuerpo, un quelante metálico, un cofactor, un ácido graso, un hidrato de carbono, un polinucleótido, un ADN, un ARN, una sonda de la PCR, un polinucleótido antisentido, un ribooligonucleótido, un desoxirribonucleótido, ADN modificado con fosforotioato, ADN y ARN modificados, ácido nucleico peptídico, un sacárido, un disacárido, un oligosacárido, un polisacárido, un dendrímero soluble en agua, una ciclodextrina, una biomaterial, una nanopartícula, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un grupo funcional novedoso, un grupo que interactúa de forma covalente o no covalente con otras moléculas, un resto fotoestimulado, un resto excitable mediante radiación actínica, un ligando un resto fotoisomerizable, biotina, un análogo de biotina, un resto que incorpora un átomo pesado, un grupo escindible químicamente, un grupo fotoescindible, una cadena lateral alargada, un azúcar unido a carbono, un agente redox activo, un aminotioácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo cromóforo, un grupo quimioluminiscente, un resto fluorescente, un grupo denso en electrones, un grupo magnético, un grupo intercalante, un grupo quelante, un cromóforo, un agente de transferencia de energía, un agente biológicamente activo, una marca detectable, una molécula pequeña, un ácido ribonucleico inhibidor, un ARNip, un radionucleótido, un agente de captura de neutrones, un derivado de biotina, punto(s) cuántico(s), un nanotransmisor, un radiotransmisor, una abzima, una enzima, un activador complejo activado, un virus, una toxina, un adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, un epítipo de linfocito T, un fosfolípido, una molécula de tipo LPS, hemocianina de lapa californiana (KLH), un hapteno inmunógeno, un aglicano, un alérgeno, una angioestatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimotopo, un receptor, una micela inversa, detergentes, potenciadores de la respuesta inmune, colorantes fluorescentes, reactivos FRET, sondas de radioimagen, otra sonda de espectroscopía, profármacos, toxinas para inmunoterapia, un soporte sólido, -CH₂CH₂-(OCH₂CH₂O)_p-OX², -O-(CH₂CH₂O)_pCH₂CH₂-X², y cualquier combinación de los mismos, donde p es de 1 a 10.000 y X² es H, un alquilo C₁₋₈, un grupo protector o un grupo funcional terminal.

19. El método de la reivindicación 7-18, donde R_7 es LX¹.

20. El método de reivindicación 19, donde X es un azúcar, un polietilenglicol, un resto fluorescente, un ácido ribonucleico, un ácido desoxirribonucleico, una proteína, un fosfolípido, un agonista de TLR7, un hapteno inmunógeno

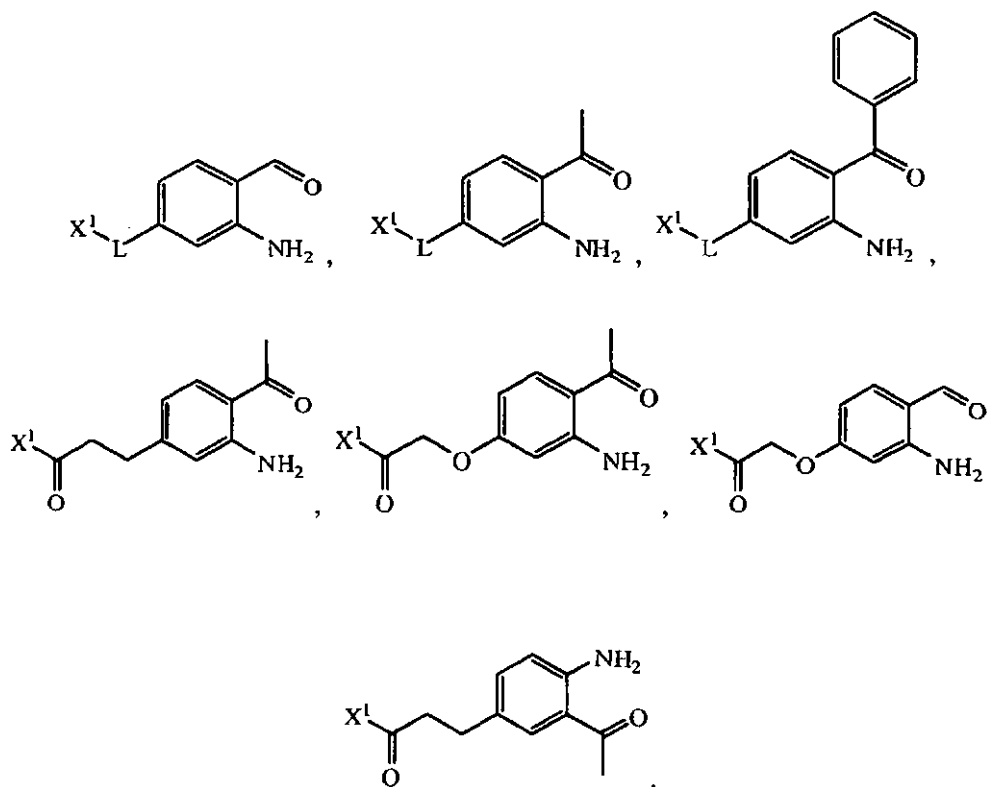
o un soporte sólido.

21. El método de la reivindicación 20, donde L es un poli(alquilenglicol), un poli(etilenglicol), alquileo C₁₋₈, un alquileo C₁₋₈ halo-sustituido o un alquileo C₁₋₈ hidroxi-sustituido.

5

22. El método de reivindicación 7, donde el reactivo de Fórmula (IV) se selecciona entre

i)



10

donde:

15

L se selecciona entre un enlace, alquileo C₁₋₈, alquileo C₁₋₈ halo-sustituido, alquileo C₁₋₈ hidroxi-sustituido, alquilenilo C₂₋₈, alquilenilo C₂₋₈ halo-sustituido, alquilenilo C₂₋₈ hidroxi-sustituido, poli(alquilenglicol), un poli(etilenglicol), -O(CR¹¹R¹²)_k-, -(CR¹¹R¹²)_kC(O)-, -S(CR¹¹R¹²)_k-, -S(O)_k(CR¹¹R¹²)_k-, -O(CR¹¹R¹²)_k-NR¹¹C(O)-, -O(CR¹¹R¹²)_kC(O)NR¹¹-, -C(O)-, -C(O)(CR¹¹R¹²)_k-, -C(S)-, -C(S)(CR¹¹R¹²)_k-, -C(O)NR¹¹-, -NR¹¹C(O)-, -NR¹¹(CR¹¹R¹²)_k-, -CONR¹¹(CR¹¹R¹²)_k-, -N(R¹¹)CO(CR¹¹R¹²)_k-, -C(O)NR¹¹(CR¹¹R¹²)_k-, -NR¹¹C(O)(CR¹¹R¹²)_k-, en donde cada R¹¹ y R¹² son independientemente H, alquilo C₁₋₈, alquilo C₁₋₈ halo-sustituido, o alquilo C₁₋₈ hidroxi-sustituido, y k es un número entero de 1 a 12, y

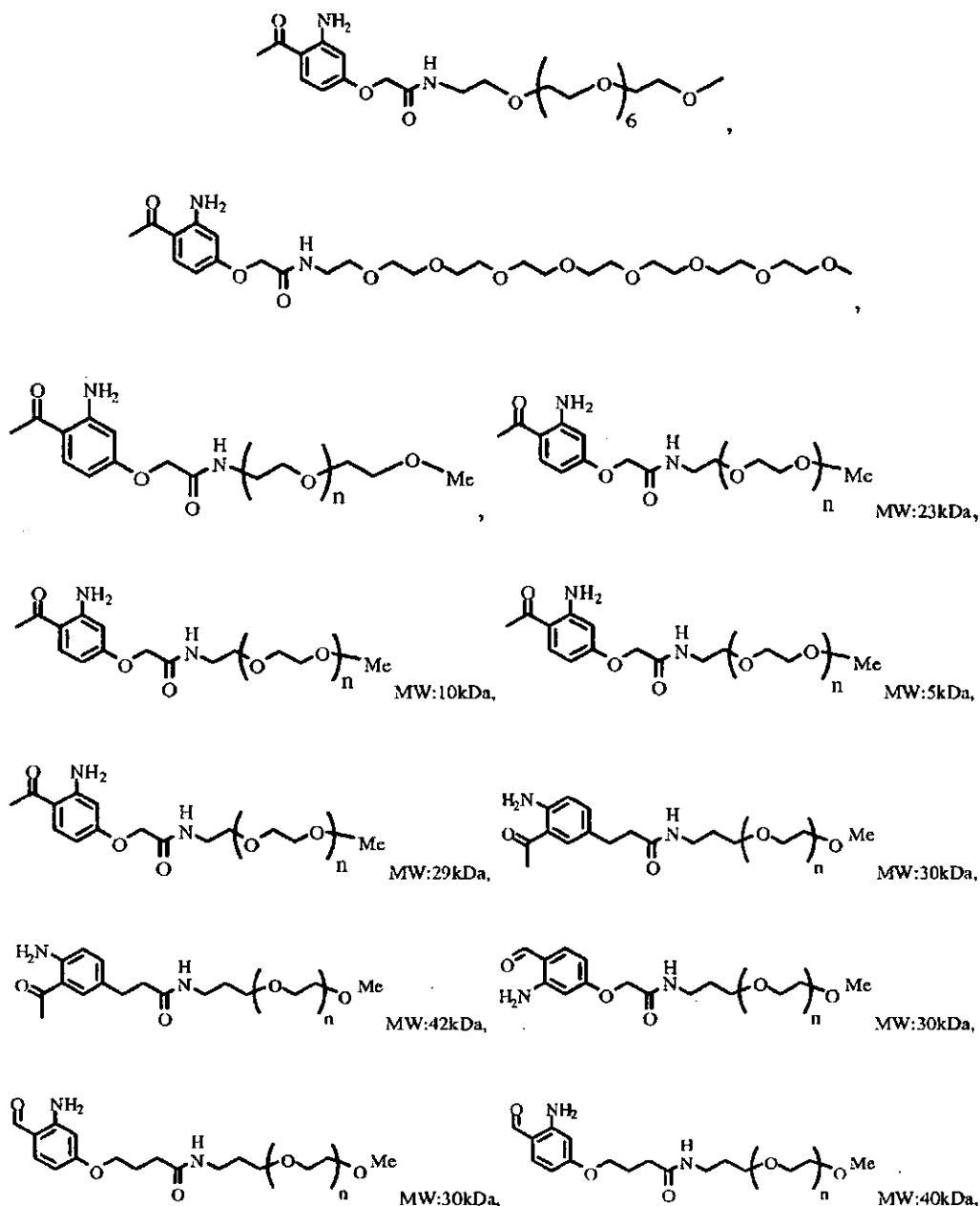
20

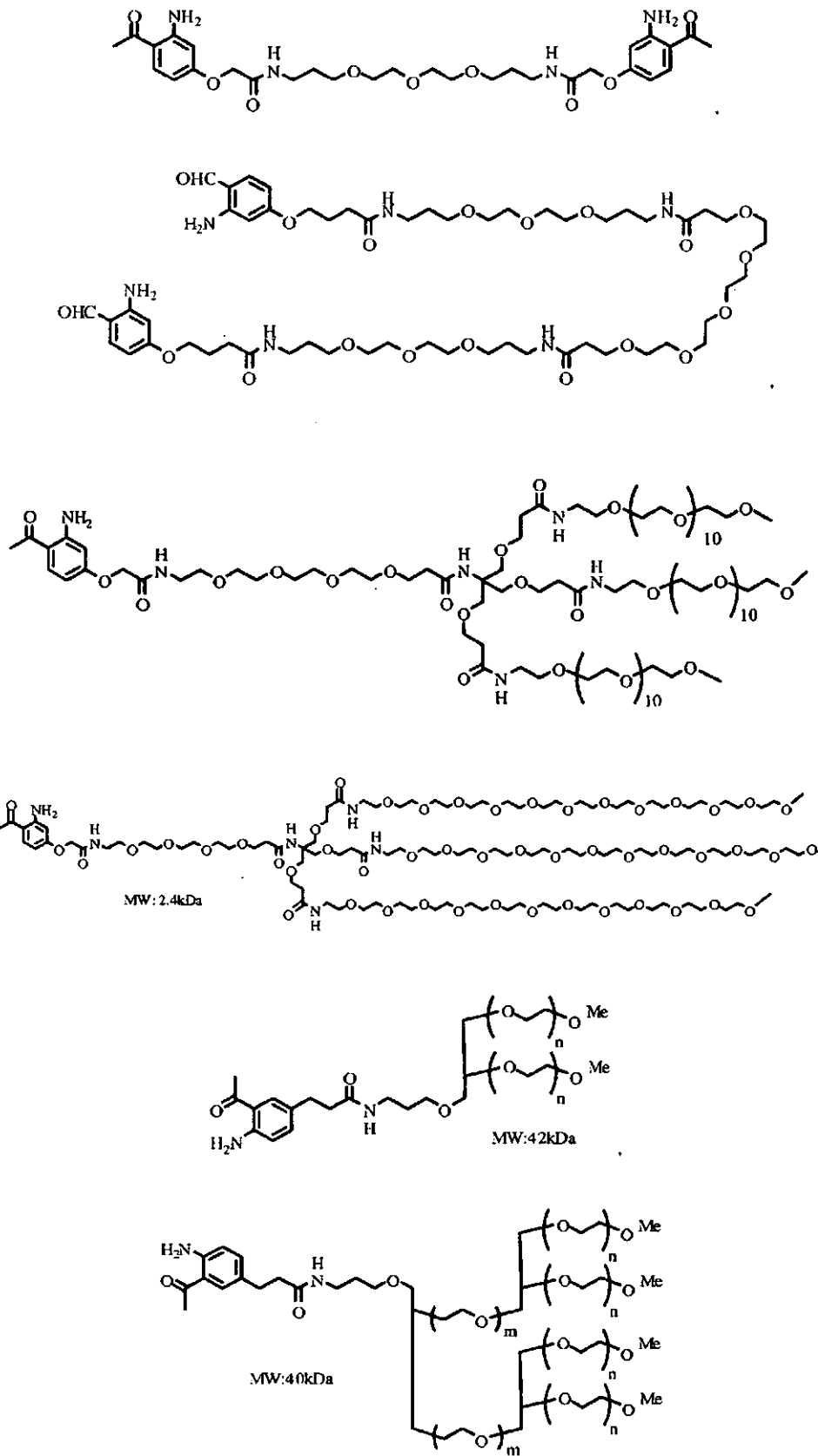
X¹ se selecciona de una marca, un colorante, un polímero, un polímero soluble en agua, un polialquilenglicol, un poli(etilenglicol), un derivado de poli(etilenglicol), un azúcar, un lípido, un fotorreticulador, un compuesto citotóxico, un fármaco, una marca de afinidad, una marca de fotoafinidad, un compuesto reactivo; una resina, un péptido, una segunda proteína o un polipéptido o un análogo de polipéptido, un anticuerpo o un fragmento de anticuerpo, un quelante metálico, un cofactor, un ácido graso, un hidrato de carbono, un polinucleótido, un ADN, un ARN, una sonda de la PCR, un polinucleótido antisentido, un ribooligonucleótido, un desoxirribooligonucleótido, ADN modificado con fosforotioato, ADN y ARN modificados, ácido nucleico peptídico, un sacárido, un disacárido, un oligosacárido, un polisacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial, una nanopartícula, una marca de espín, un fluoróforo, un resto que contiene metal, un resto radioactivo, un grupo funcional novedoso, un grupo que interactúa de forma covalente o no covalente con otras moléculas, un resto fotoestimulado, un resto excitable mediante radiación actínica, un ligando, un resto fotoisomerizable, biotina, un análogo de biotina, un resto que incorpora un átomo pesado, un grupo escindible químicamente, un grupo fotoescindible, una cadena lateral alargada, un azúcar unido a carbono, un agente redox activo, un aminotioácido, un resto tóxico, un resto marcado isotópicamente, una sonda biofísica, un grupo fosforescente, un grupo cromóforo, un grupo quimioluminiscente, un resto fluorescente, un grupo denso en electrones, un grupo magnético, un grupo intercalante, un grupo quelante, un cromóforo, un agente de

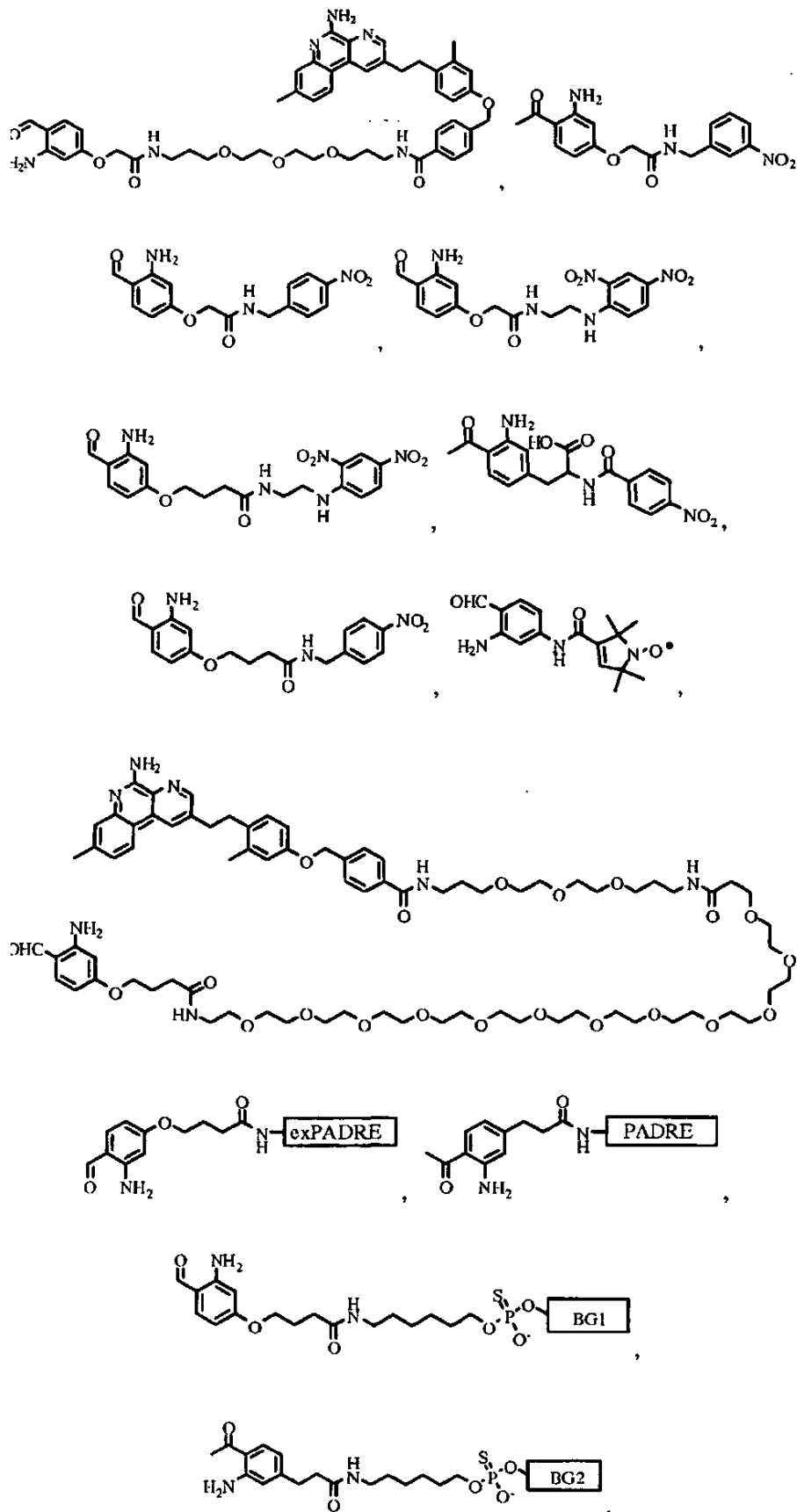
35

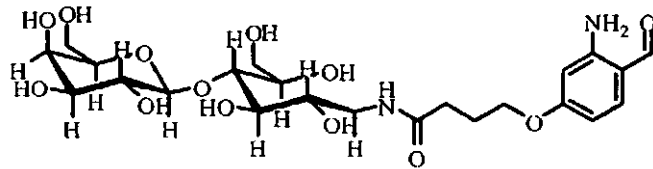
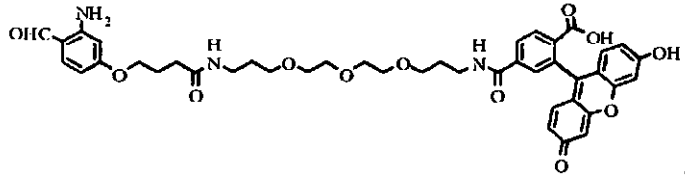
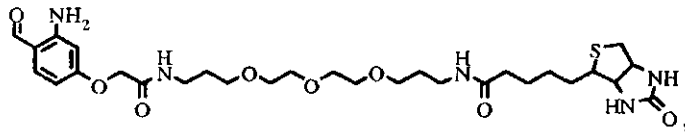
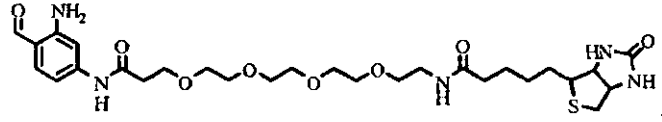
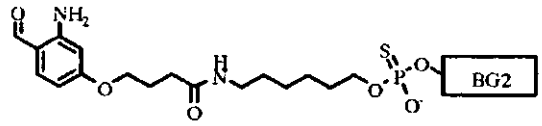
transferencia de energía, un agente biológicamente activo, una marca detectable, una molécula pequeña, un ácido ribonucleico inhibidor, un ARNip, un radionucleótido, un agente de captura de neutrones, un derivado de biotina, punto(s) cuántico(s), un nanotransmisor, un radiotransmisor, una abzima, una enzima, un activador complejo activado, un virus, una toxina, un adyuvante, un agonista de TLR2, un agonista de TLR4, un agonista de TLR7, un agonista de TLR9, un agonista de TLR8, un epítipo de linfocito T, un fosfolípido, una molécula de tipo LPS, hemocianina de lapa californiana (KLH), un hapteno inmunógeno, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimotopo, un receptor, una micela inversa, detergentes, potenciadores de la respuesta inmune, colorantes fluorescentes, reactivos FRET, sondas de radioimagen, otra sonda de espectroscopía, profármacos, toxinas para inmunoterapia, un soporte sólido, $-\text{CH}_2\text{CH}_2-(\text{OCH}_2\text{CH}_2\text{O})_p-\text{OX}^2$, $-\text{O}-(\text{CH}_2\text{CH}_2\text{O})_p\text{CH}_2\text{CH}_2-\text{X}^2$, y cualquier combinación de los mismos, donde p es de 1 a 10.000 y X^2 es H, un alquilo C_{1-8} , un grupo protector o un grupo funcional terminal o

ii) el reactivo de Fórmula (IV) se selecciona entre

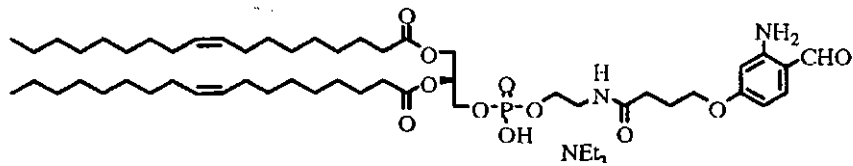








y



donde:

5

exPADRE es un péptido de acuerdo con la SEC ID N°28;
 PADRE es un péptido de acuerdo con la SEC ID N°29;
 BG1 es un polinucleótido de acuerdo con la SEC ID N°30; y
 BG2 es un polinucleótido de acuerdo con la SEC ID N°31.

10

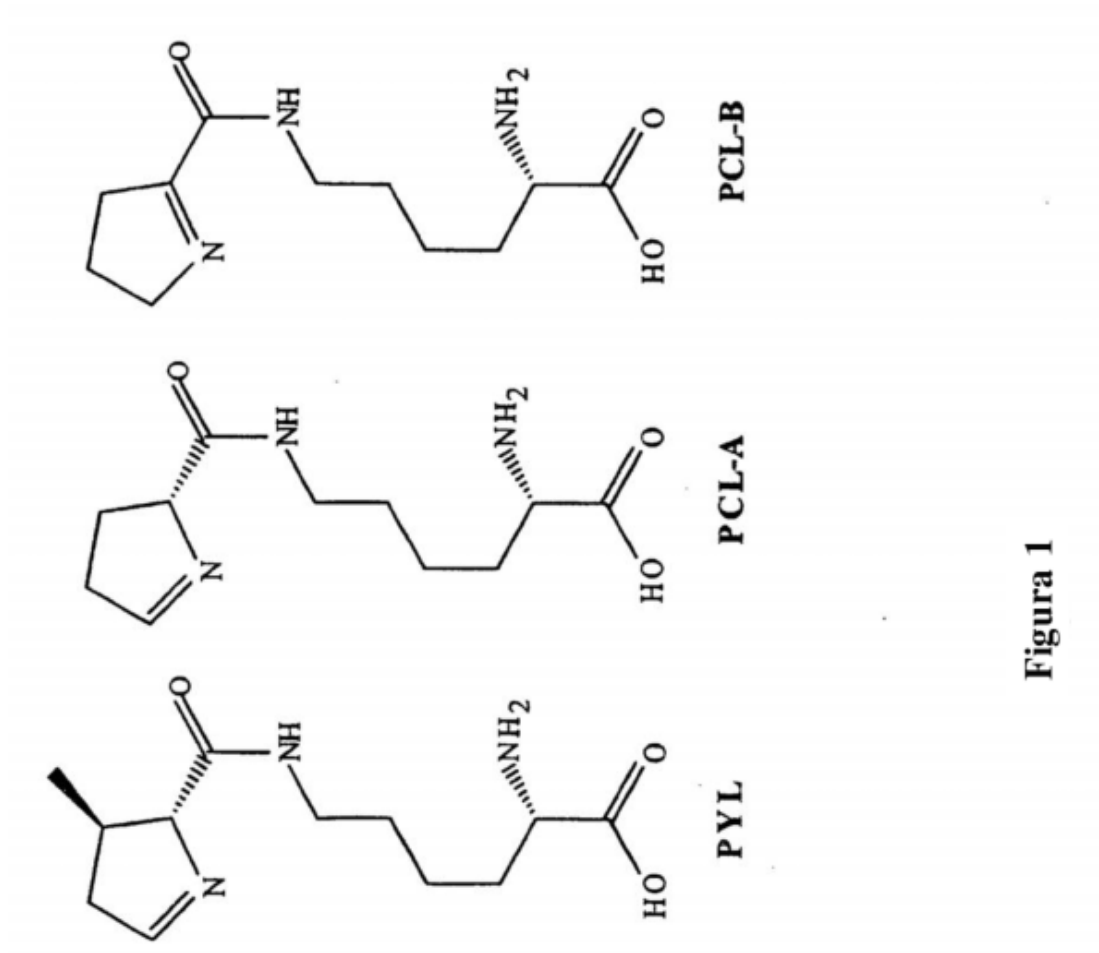


Figure 1

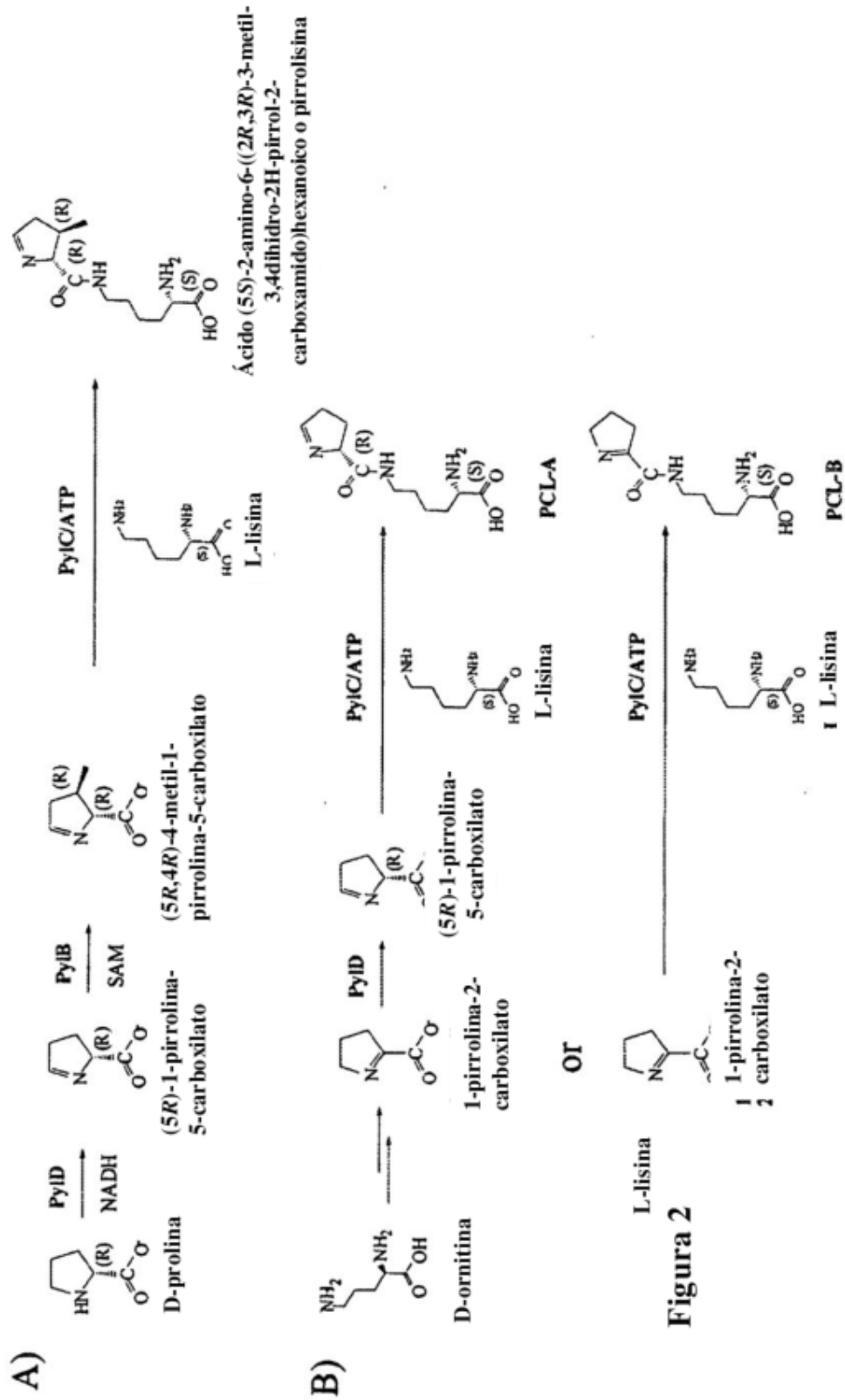
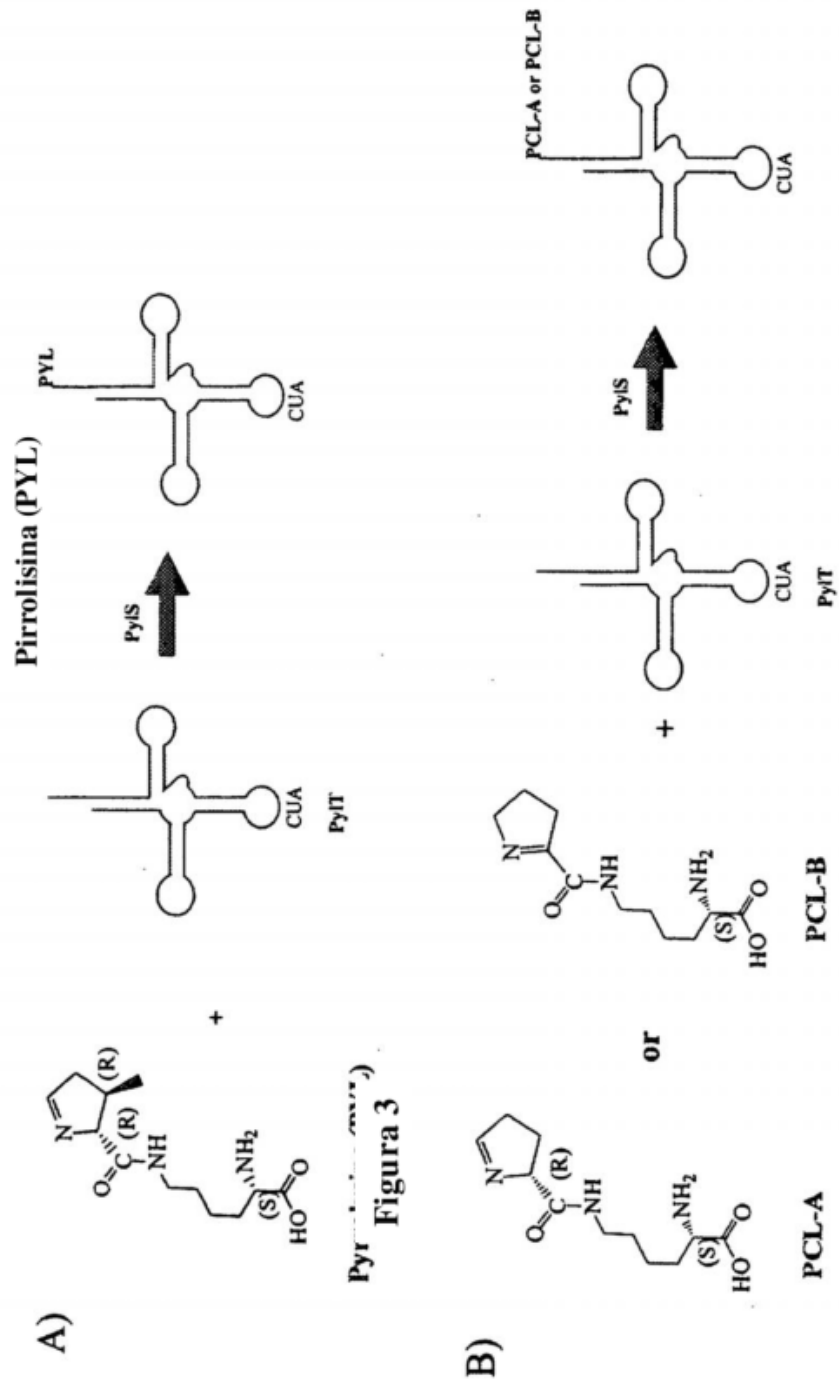


Figura 2



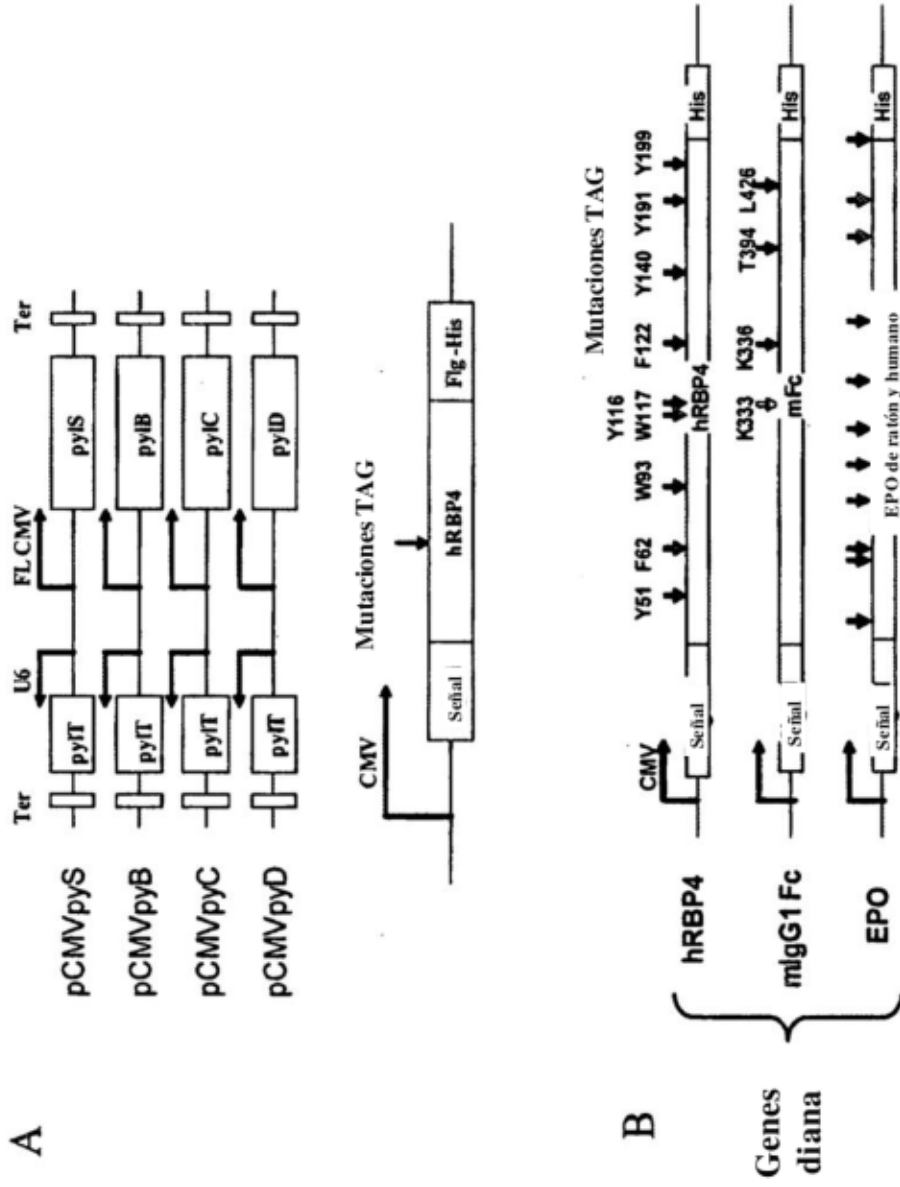


Figura 4

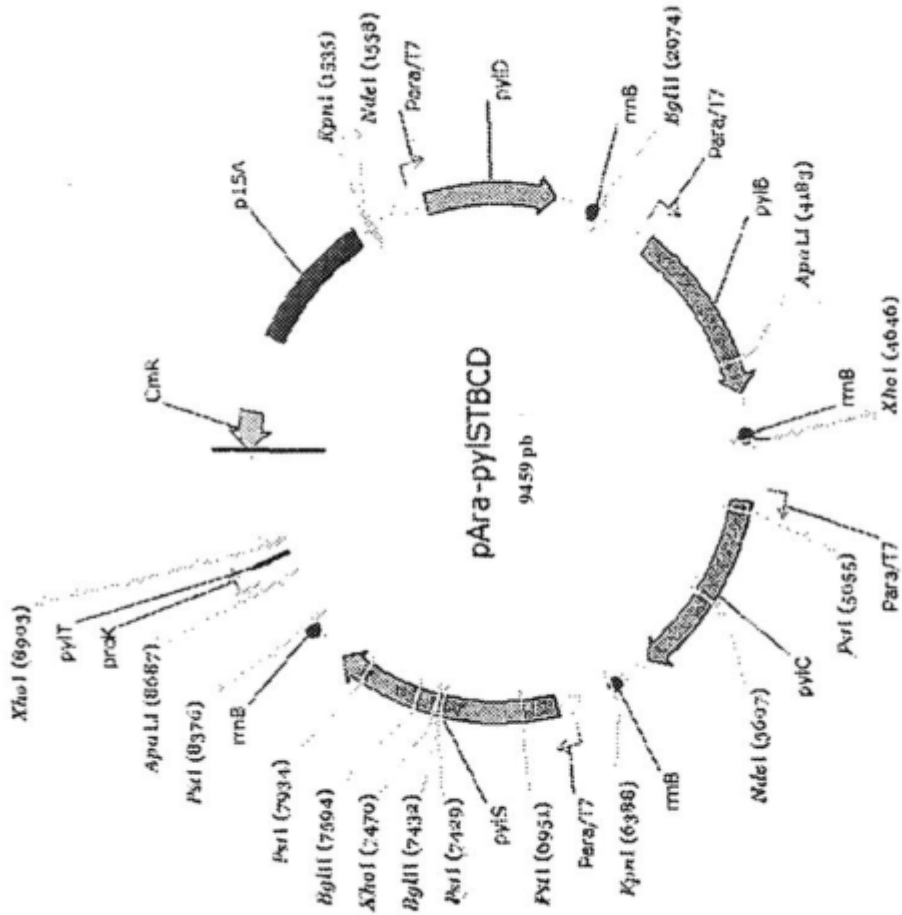


Figure 5

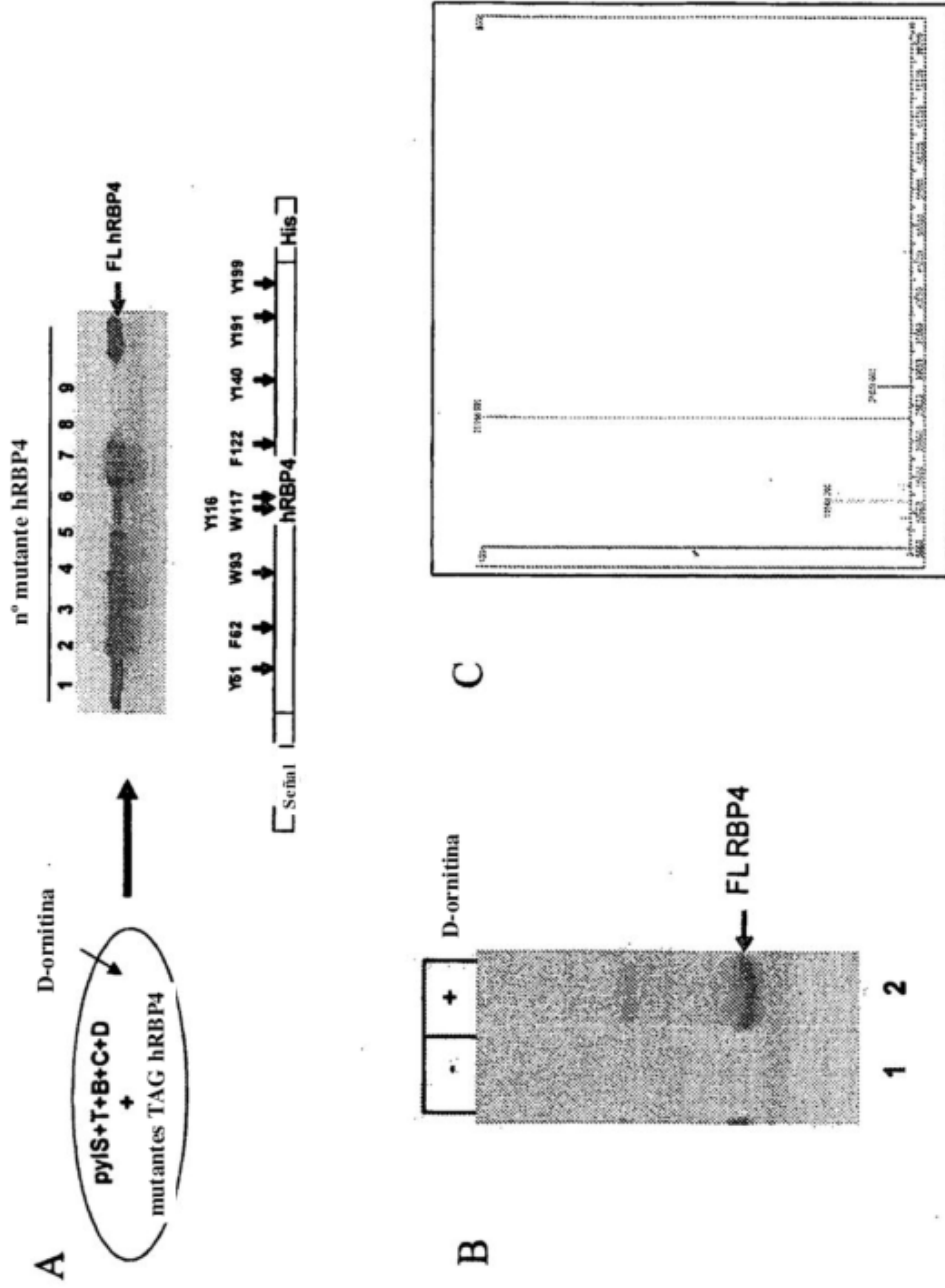


Figura 6

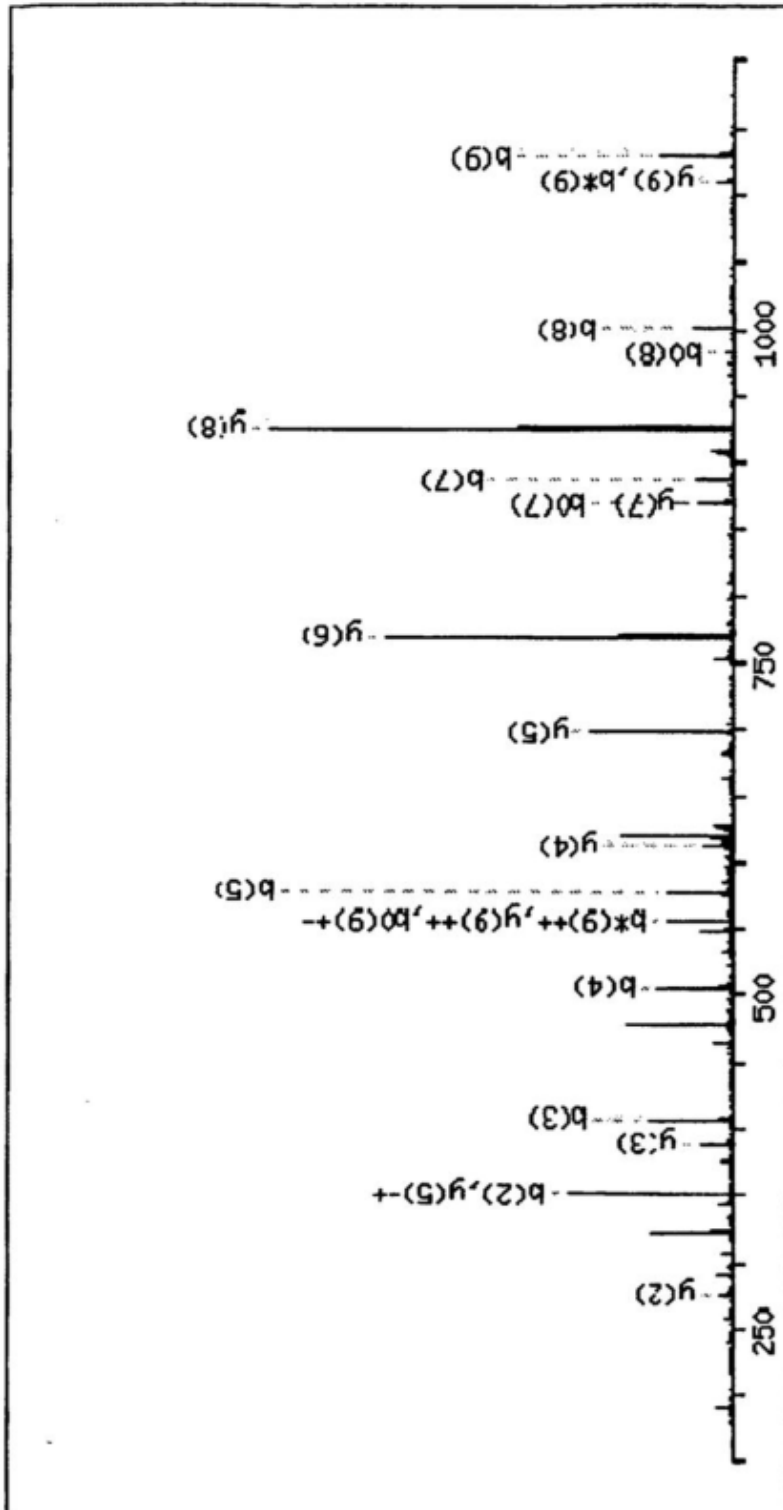


Figura 7

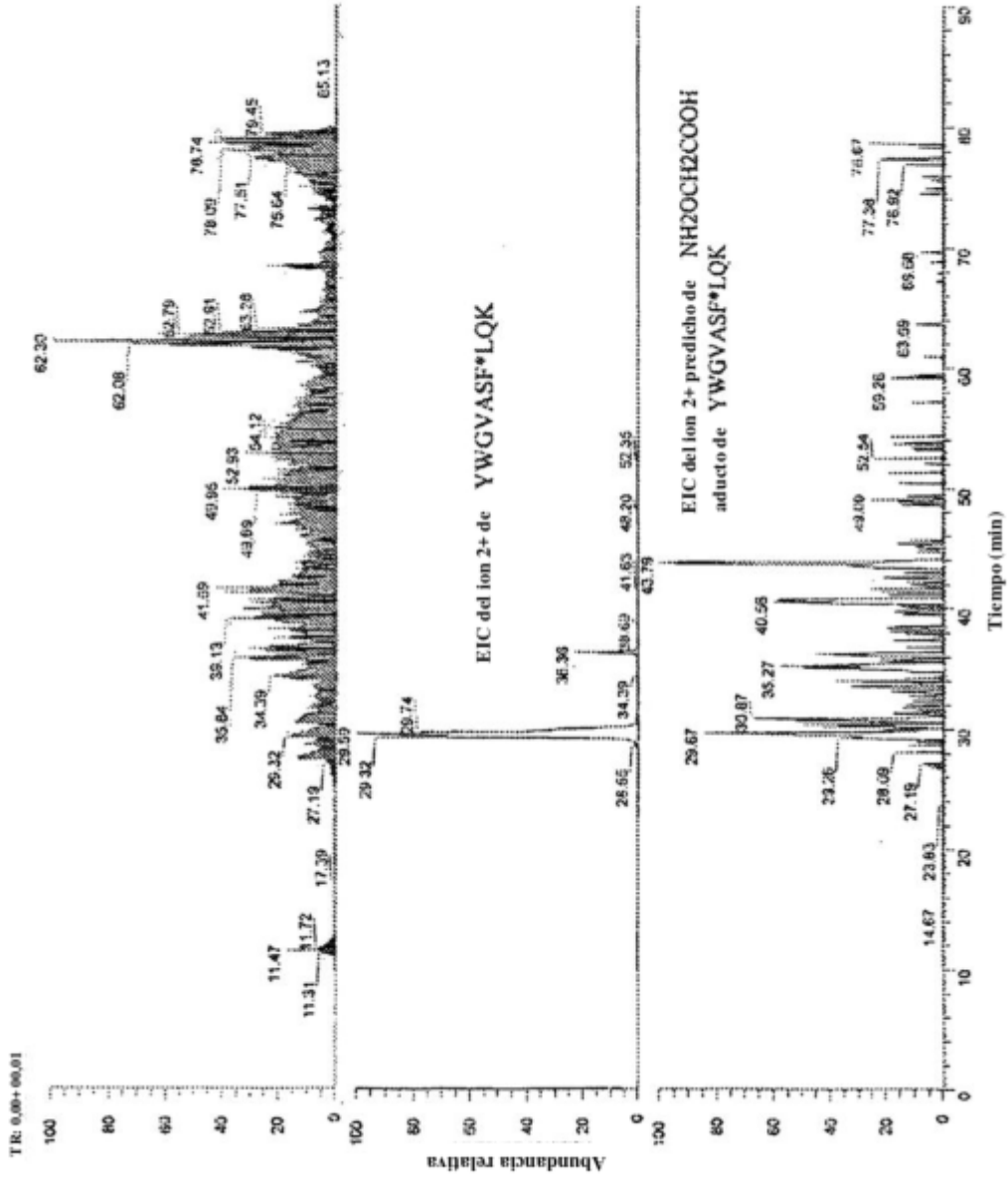


Figura 8

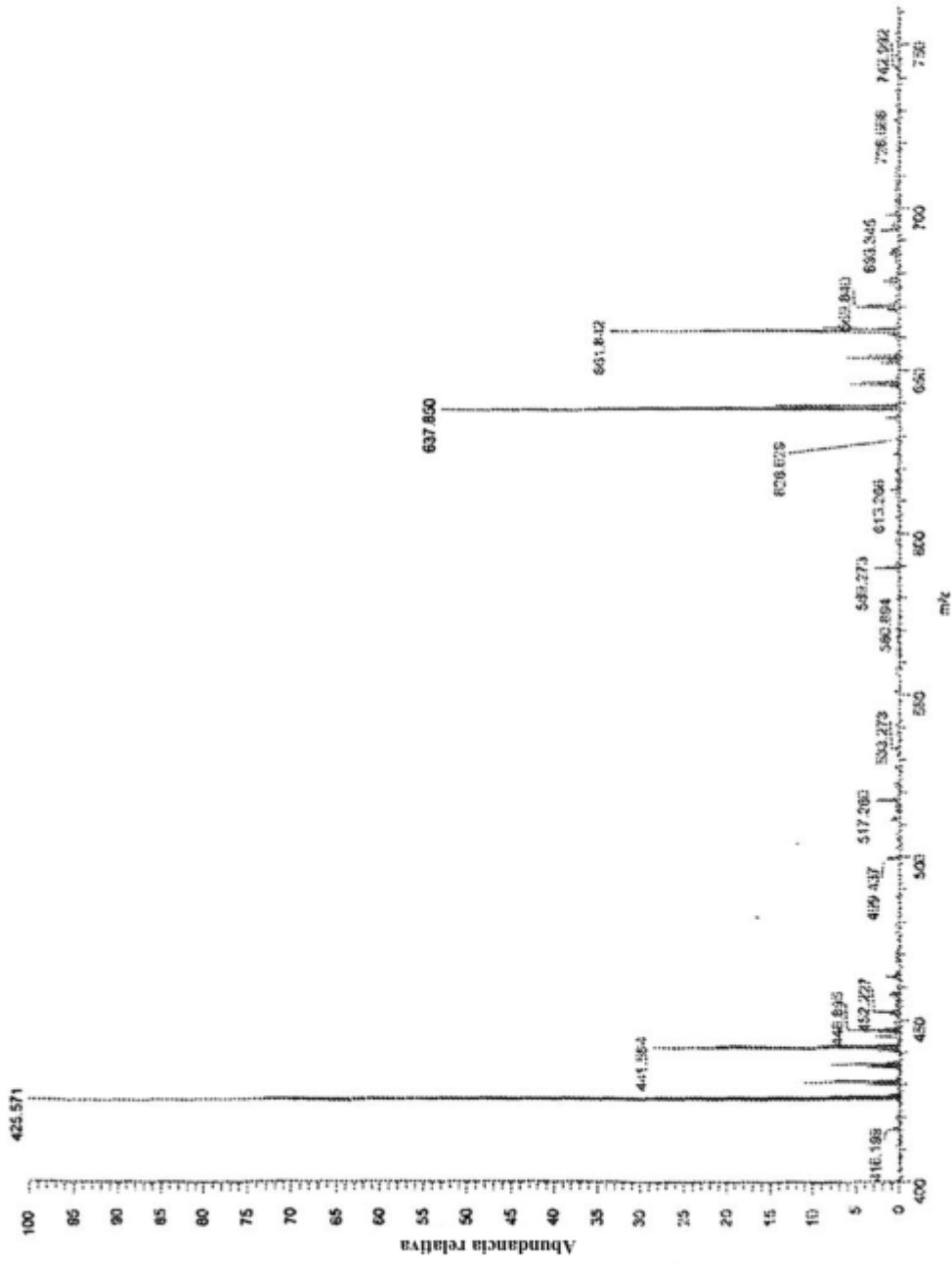


Figura 9

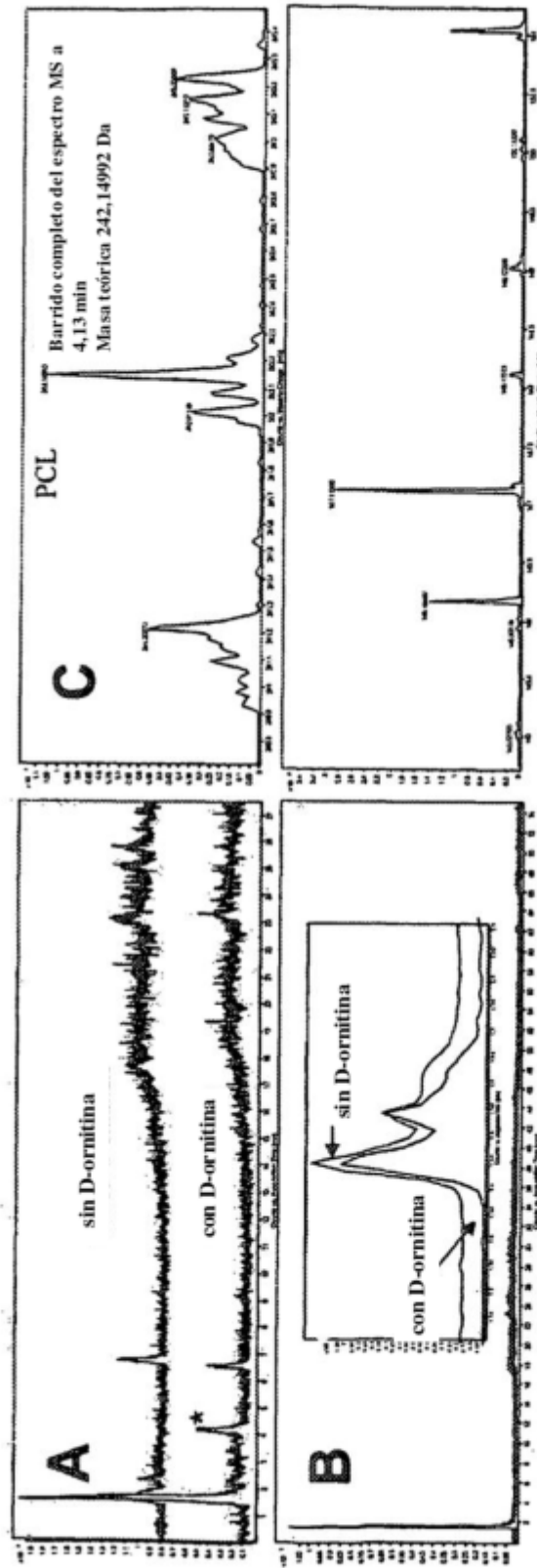


Figura 10

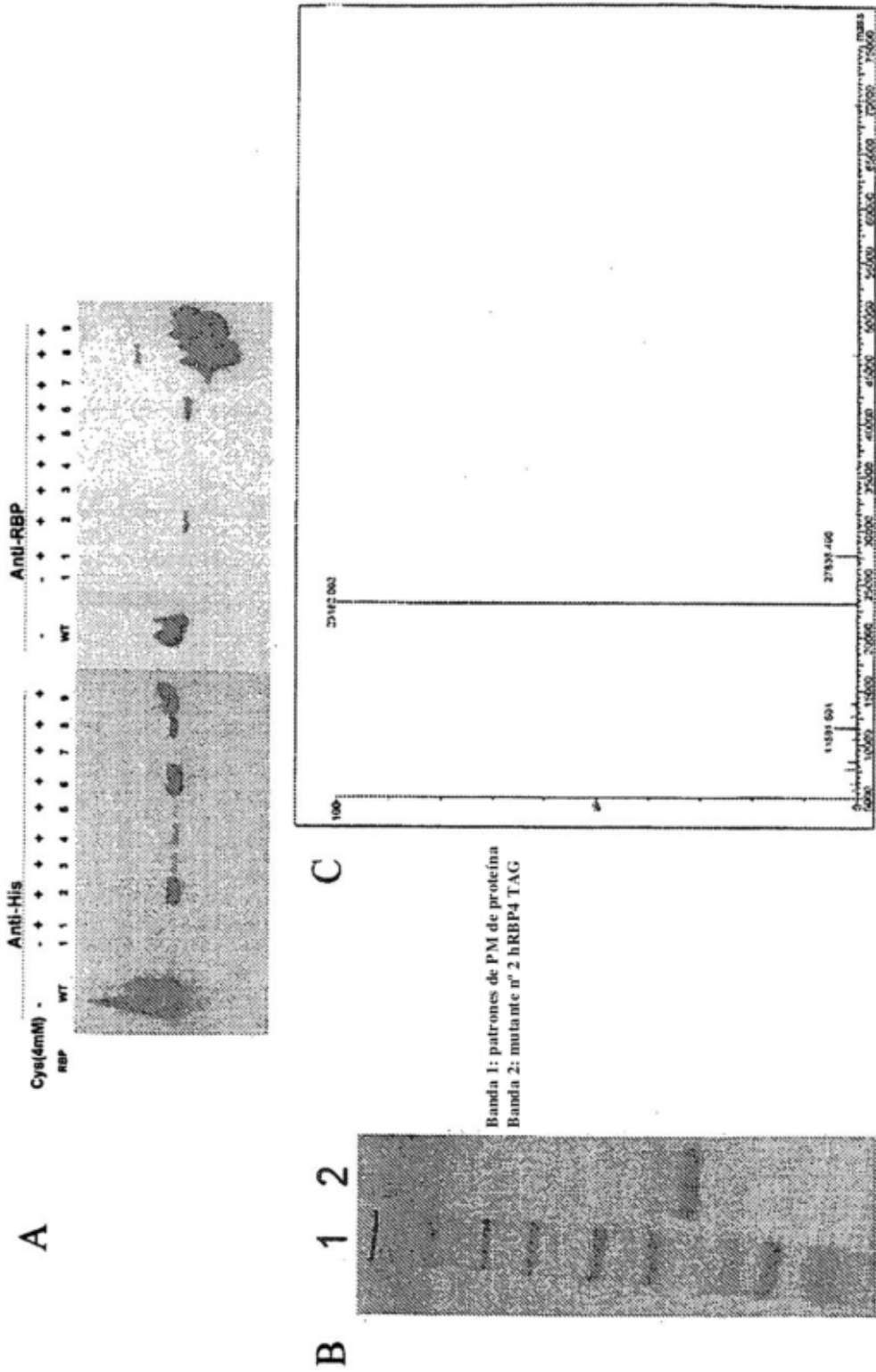


Figura 11

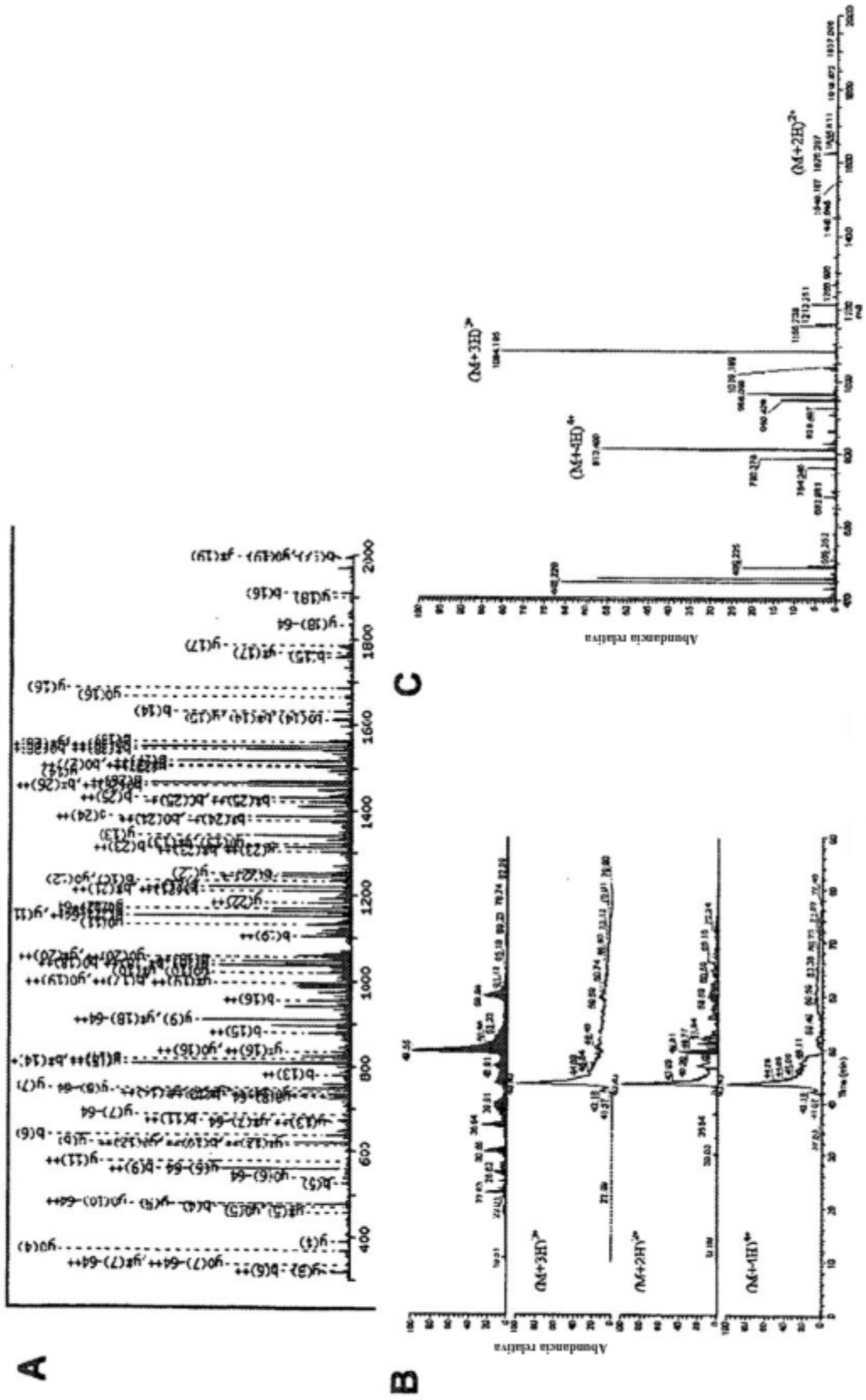


Figura 12

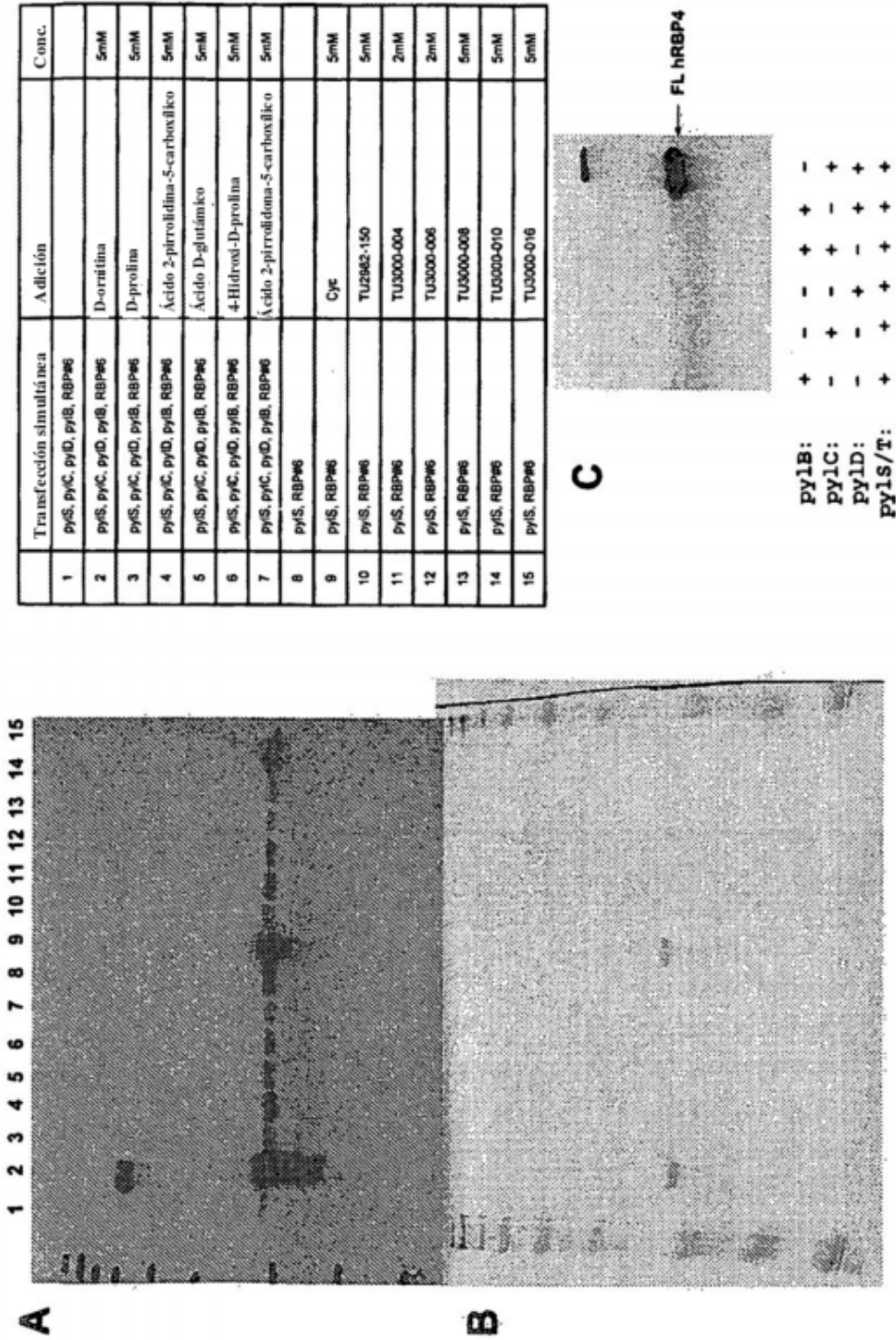


Figura 13

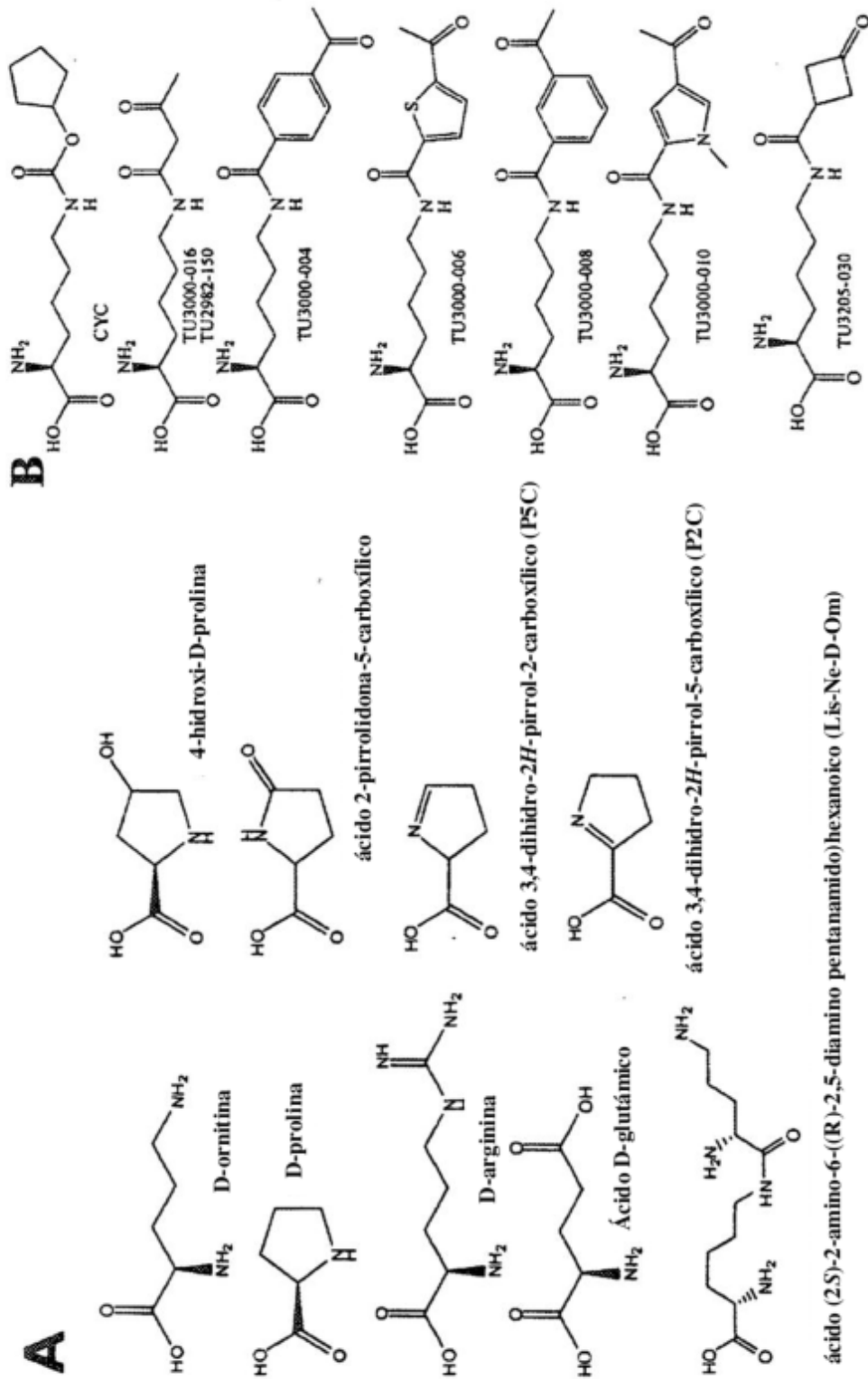


Figura 14

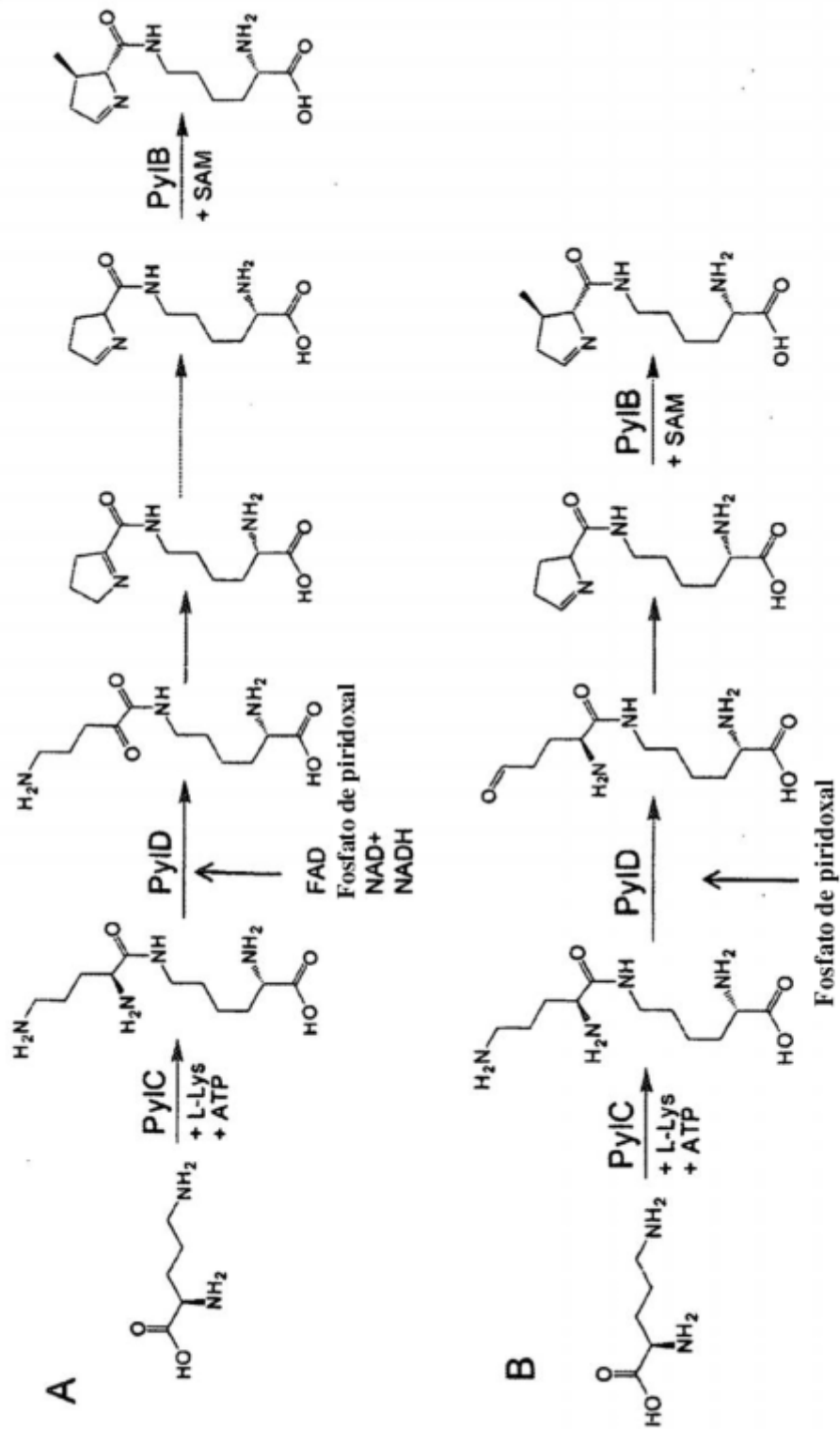


Figura 16

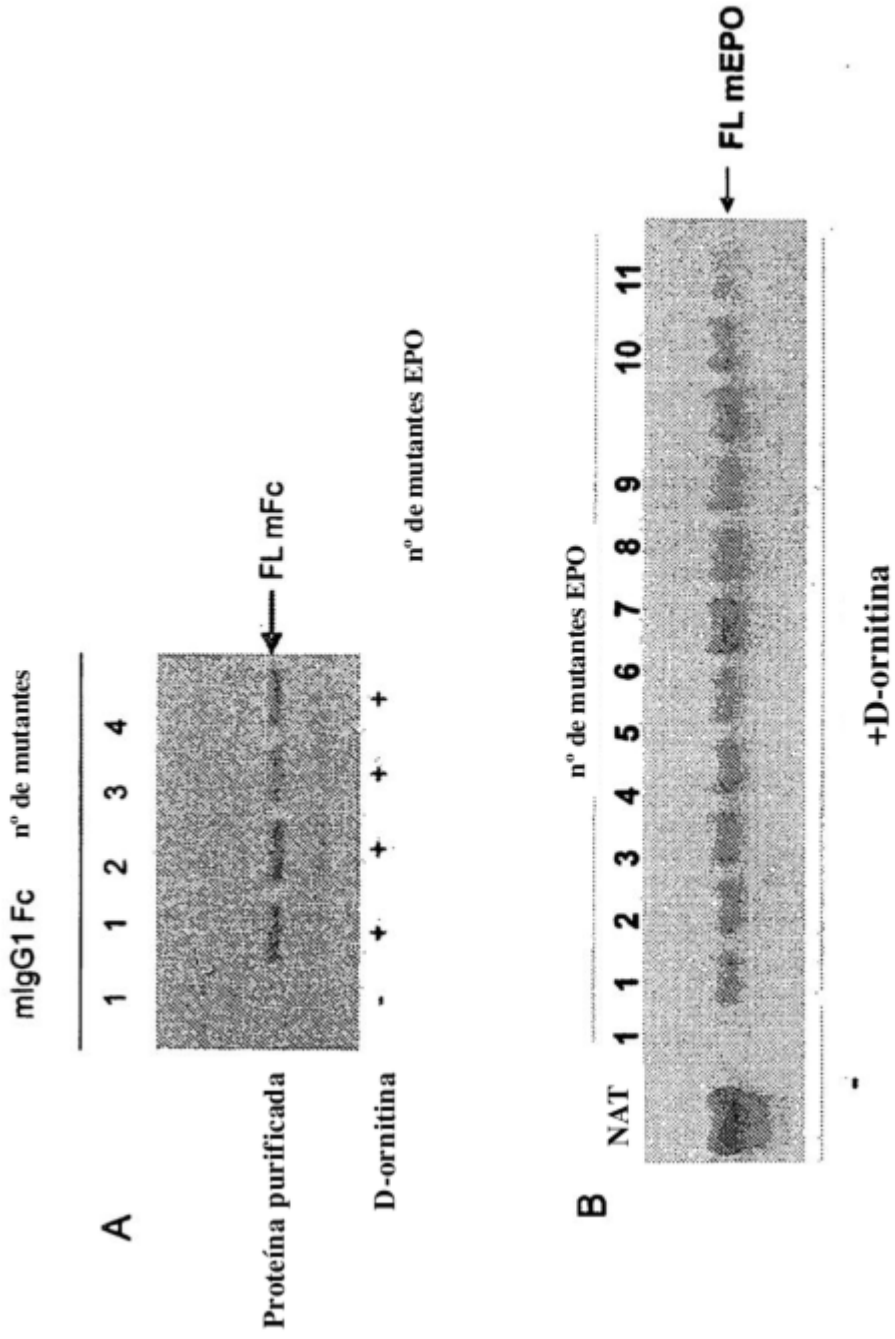
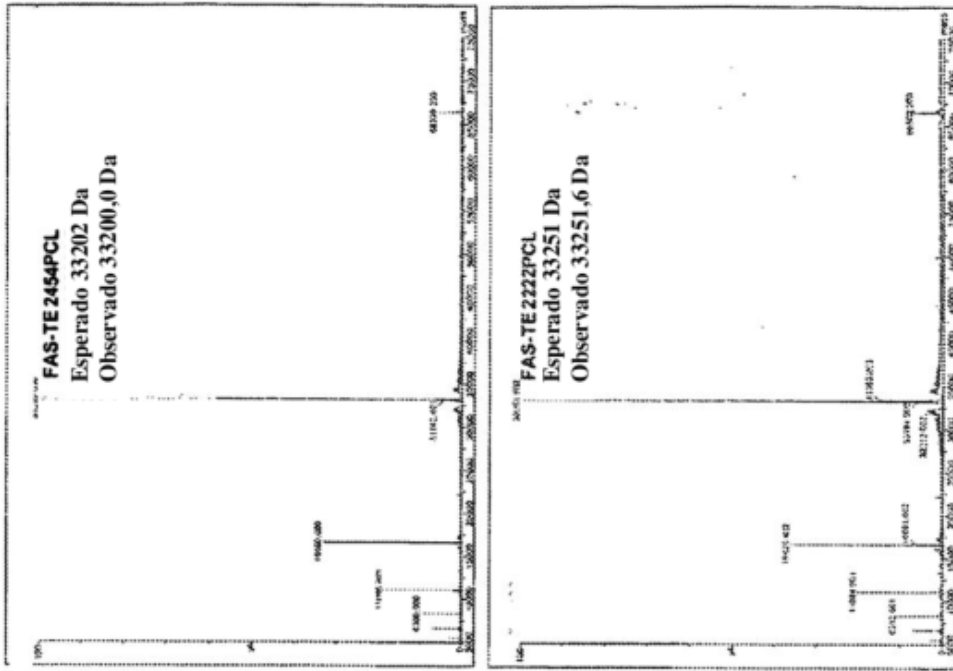
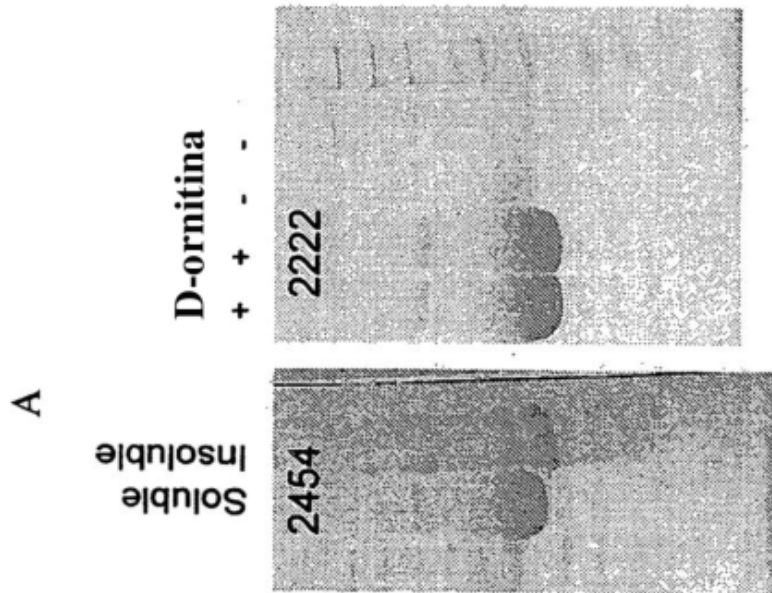


Figura 17



B

Figura 18



A

Soluble

2454

D-ornitina

+ + - -

2222

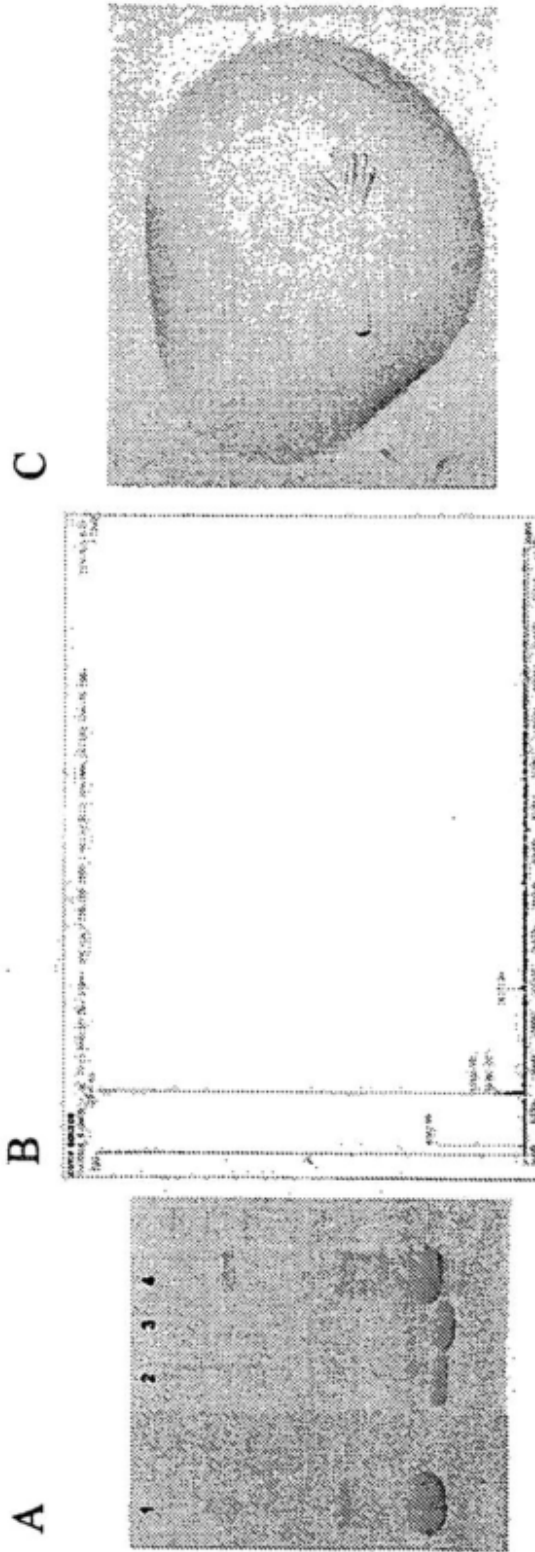


Figura 19

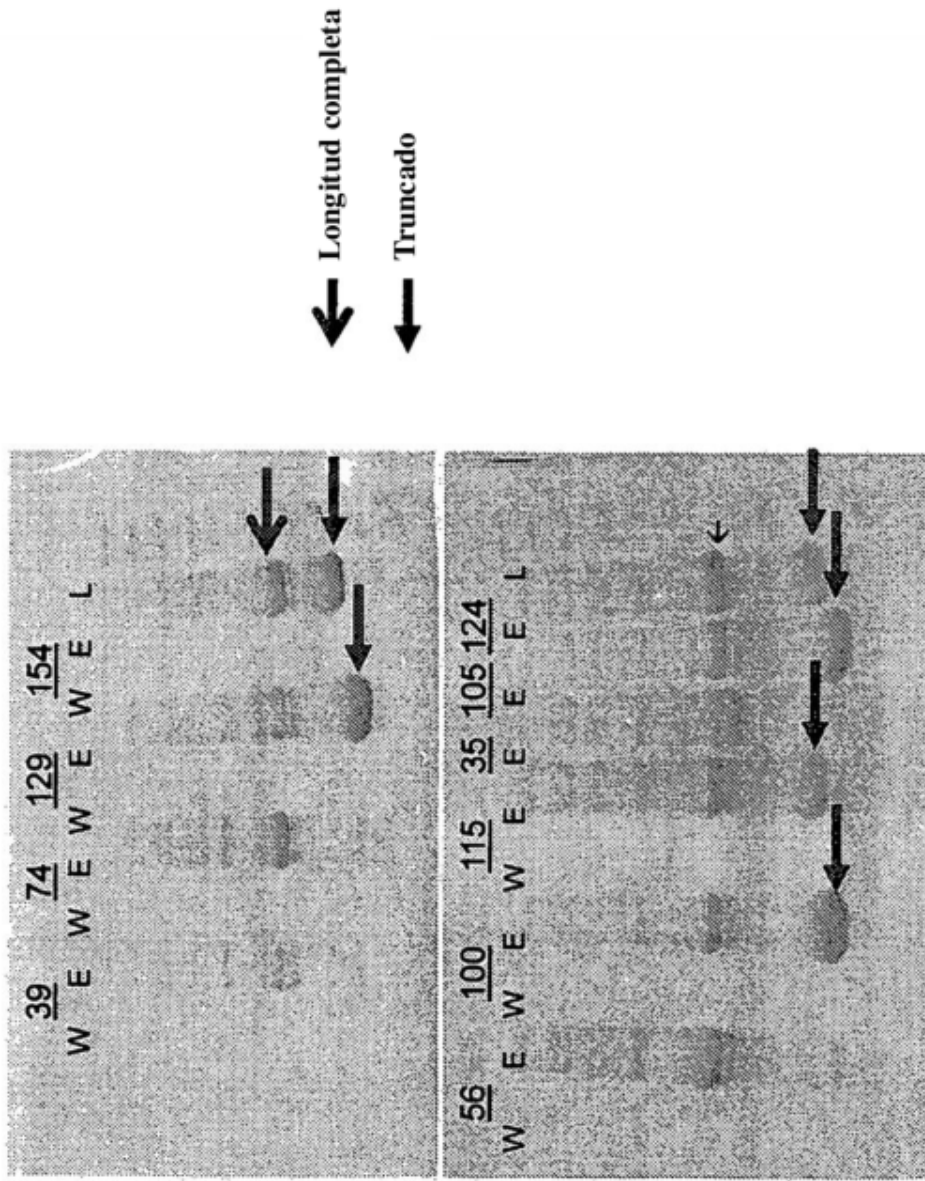


Figura 20

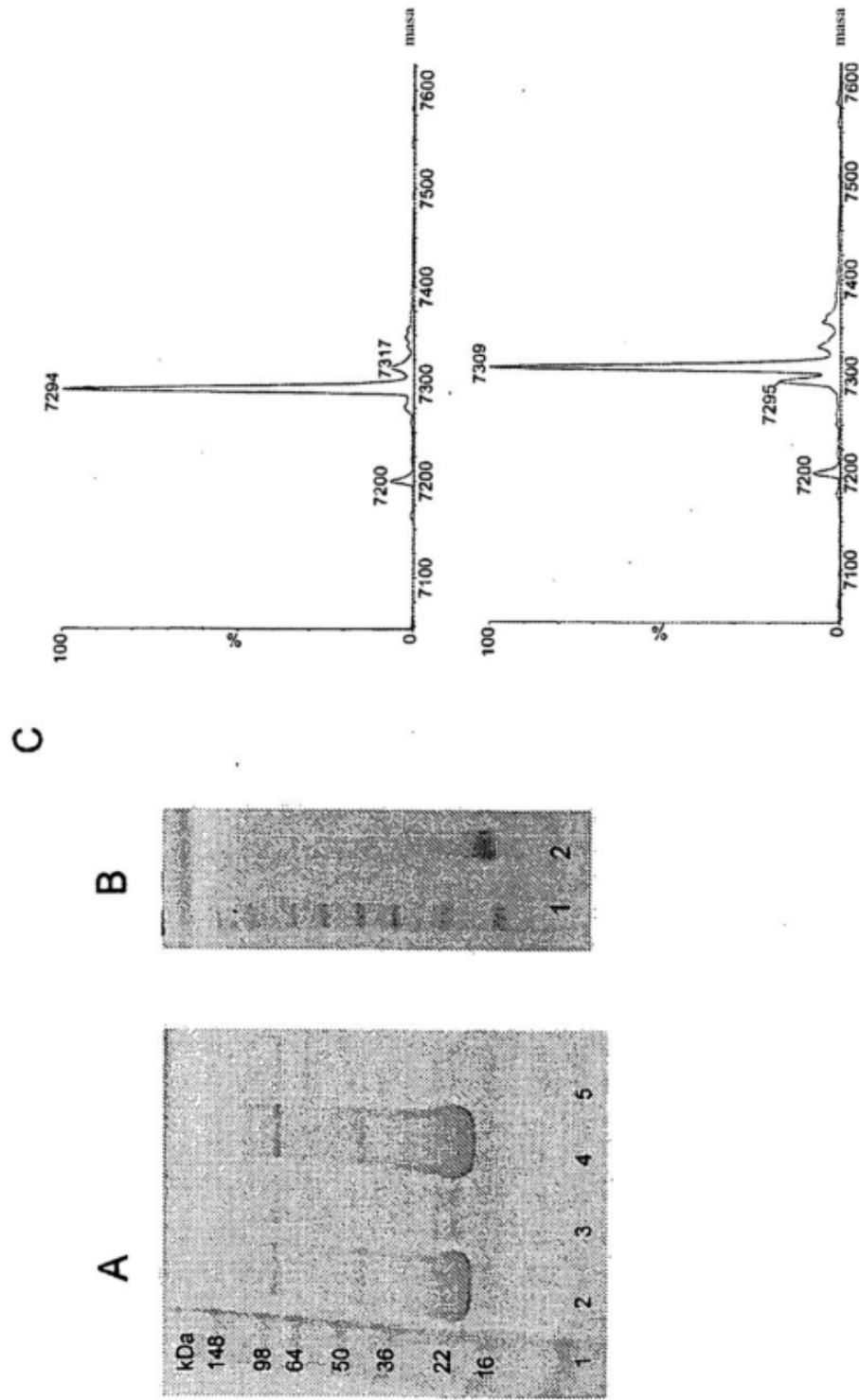


Figure 21

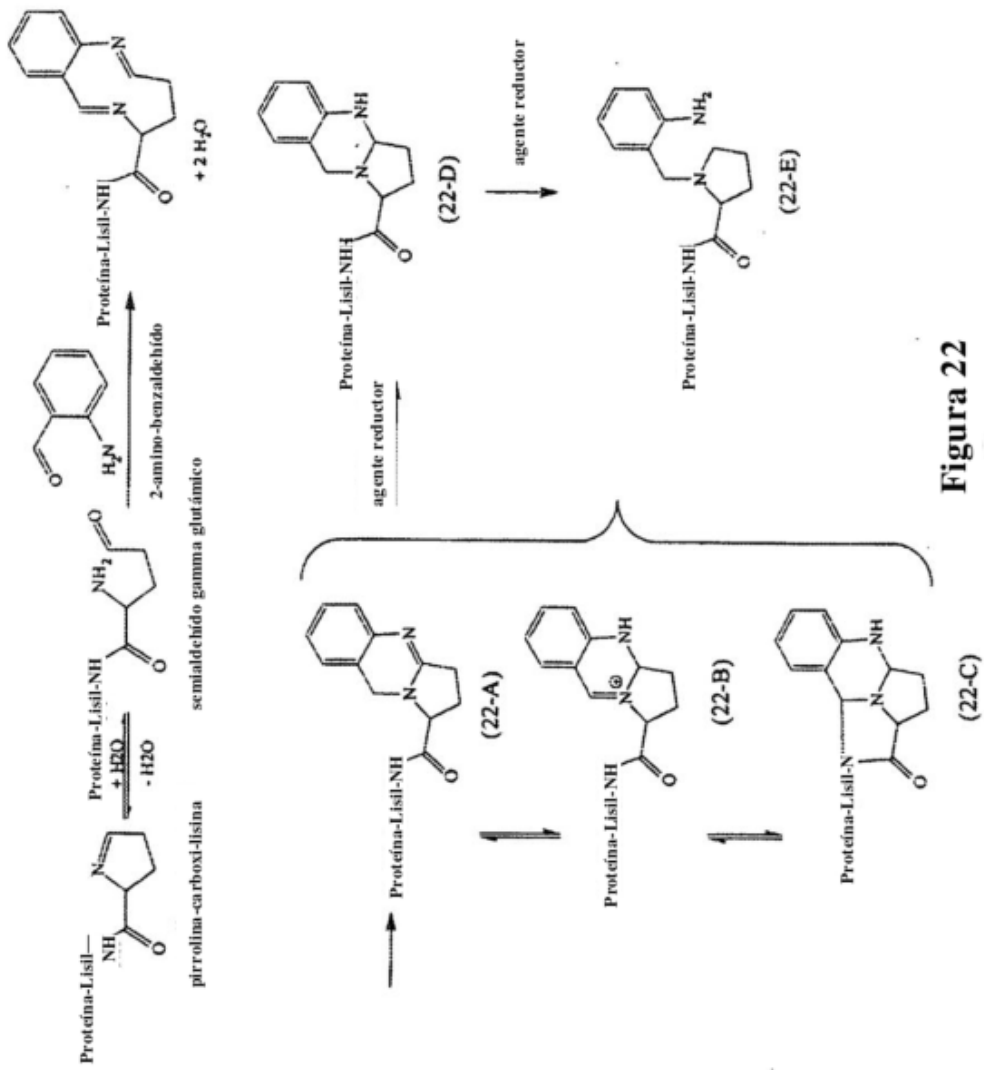


Figura 22

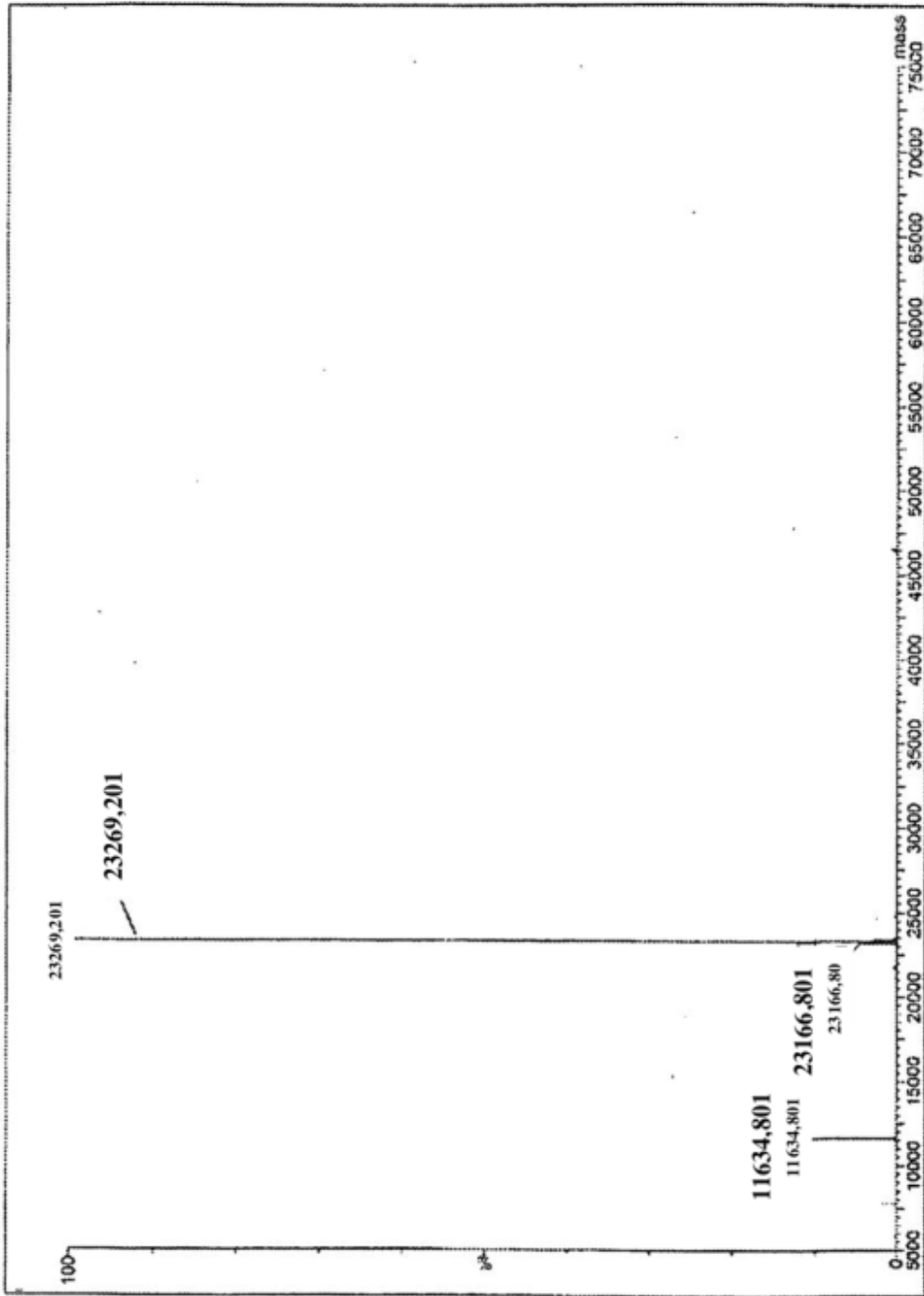
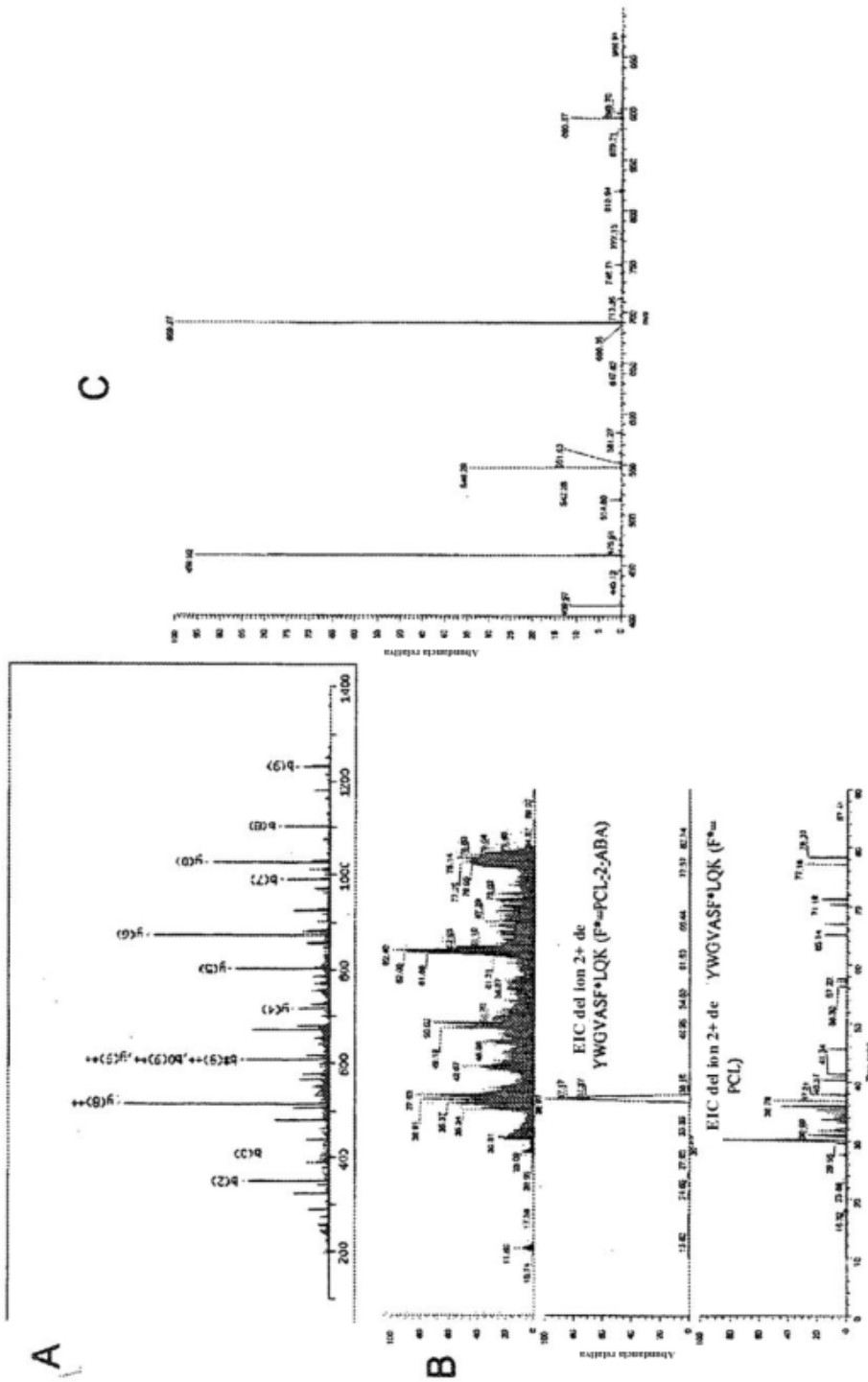


Figura 24



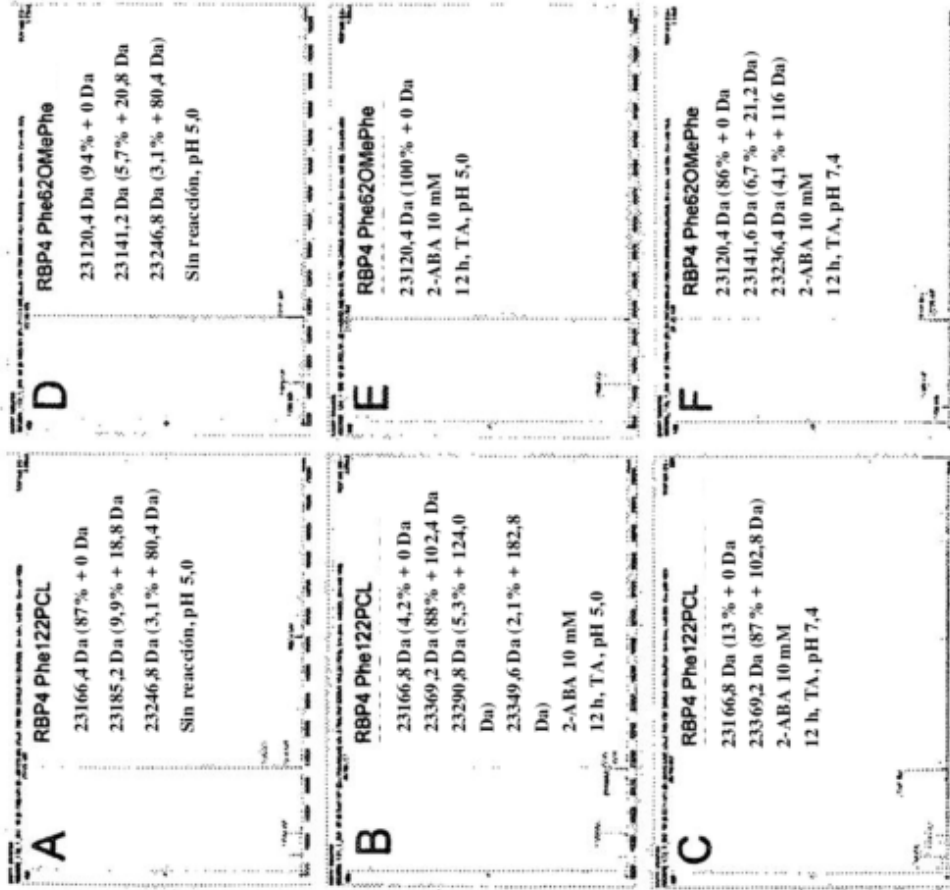


Figura 26

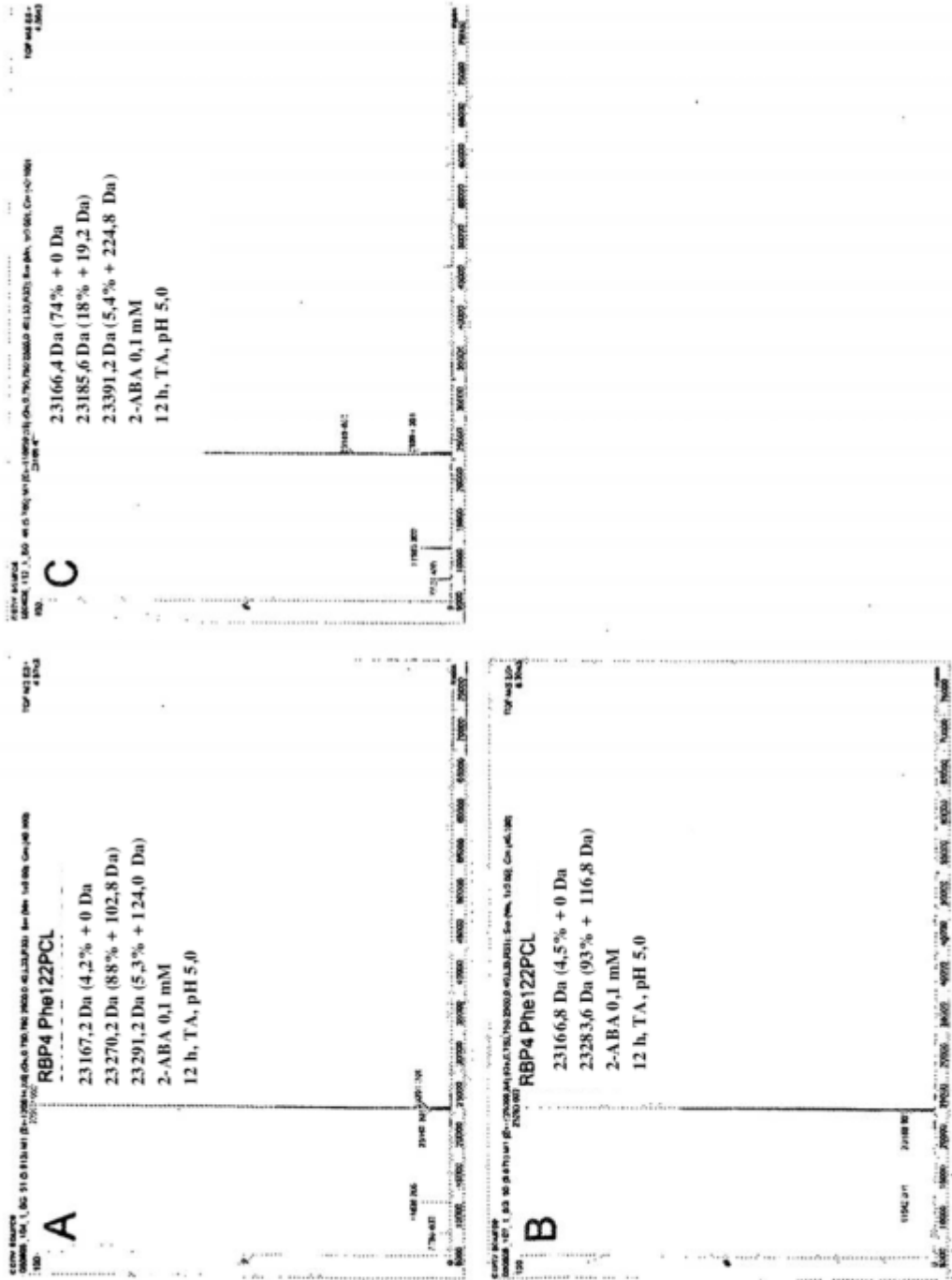


Figura 27

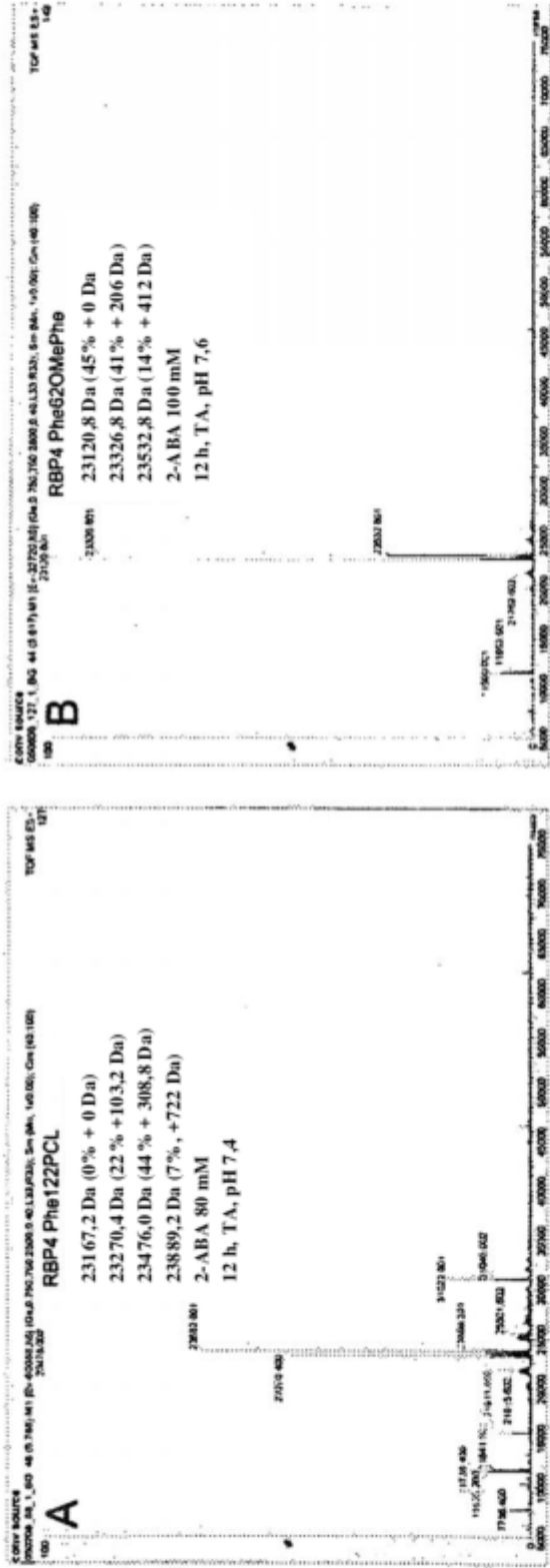


Figura 28

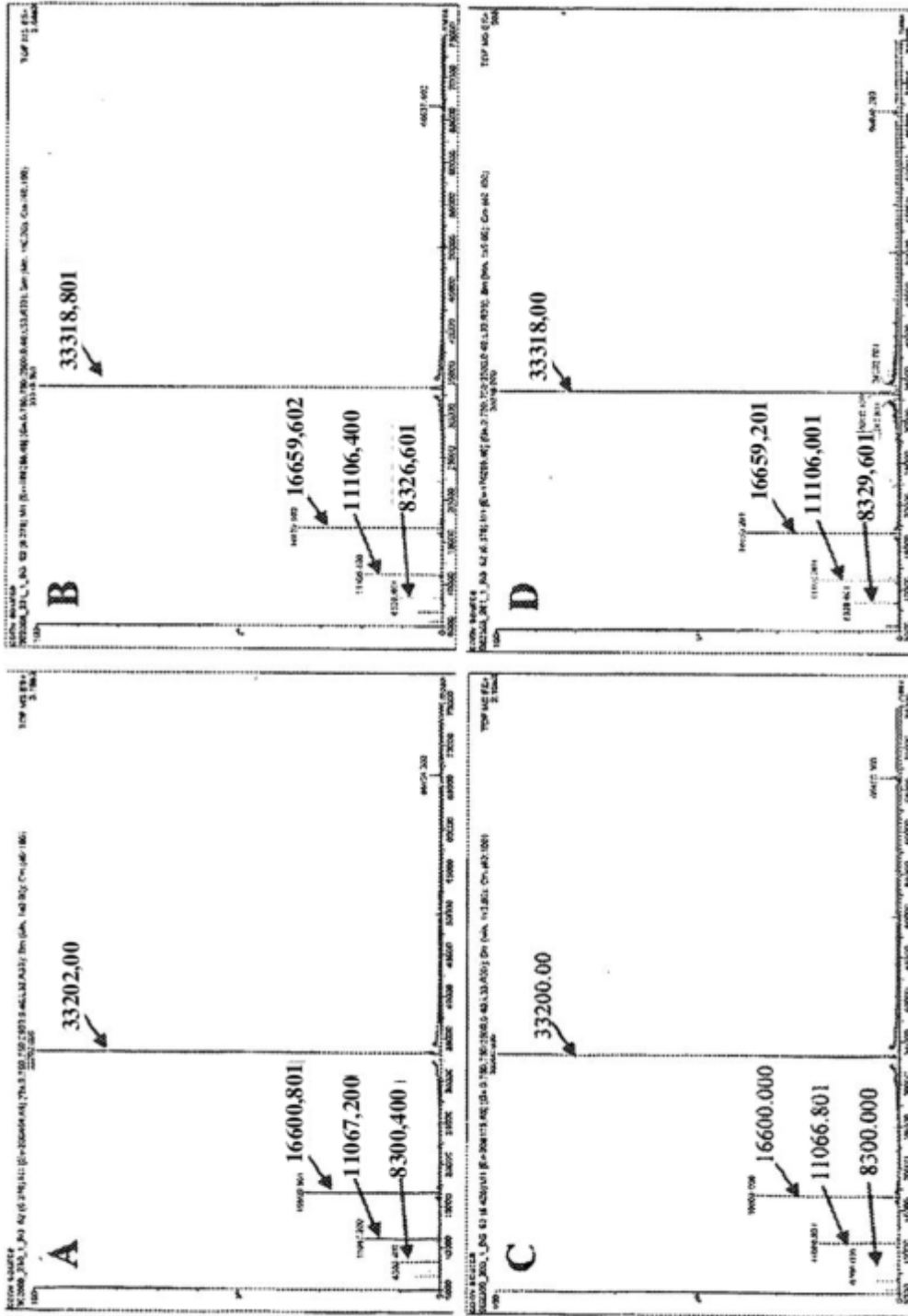


Figura 29

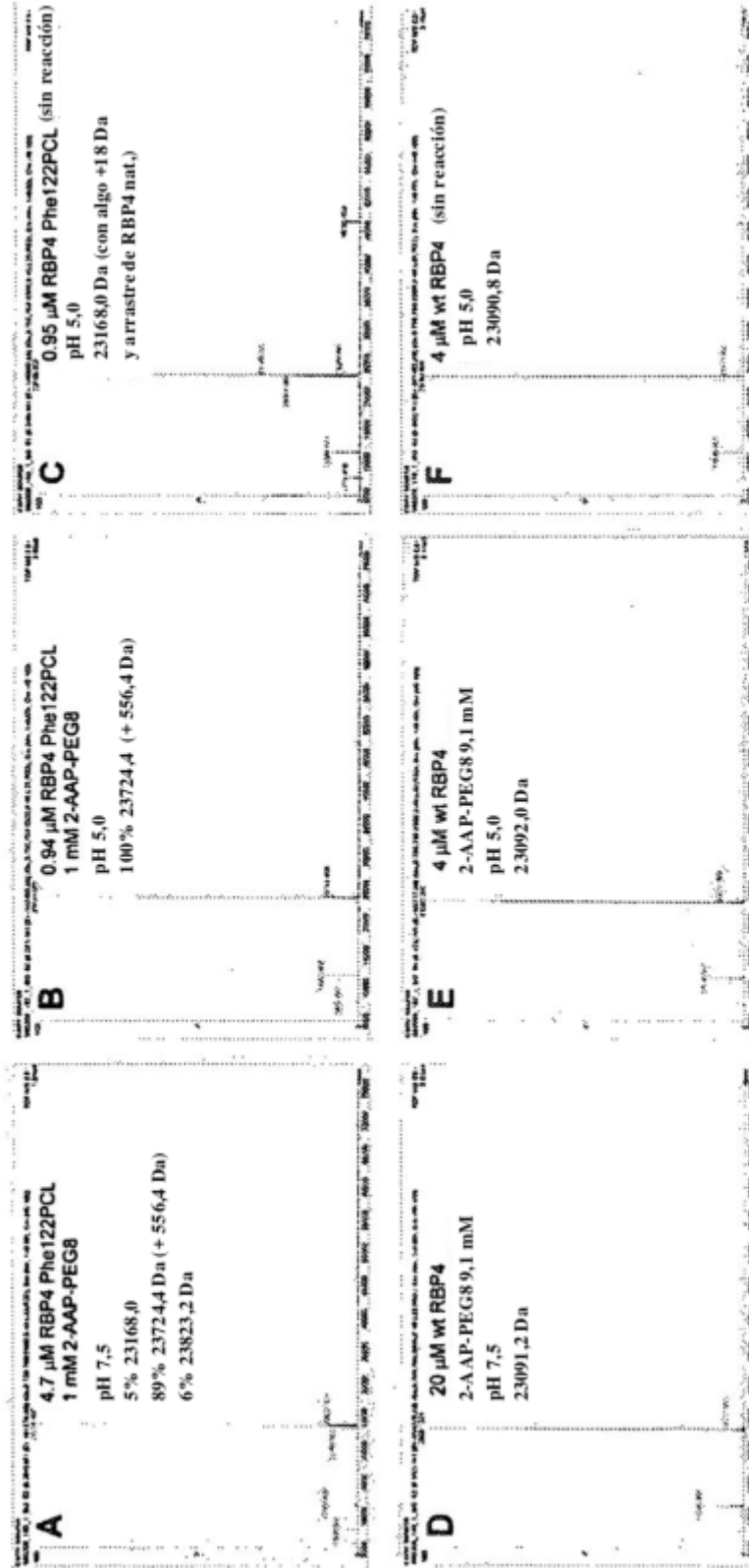


Figura 32

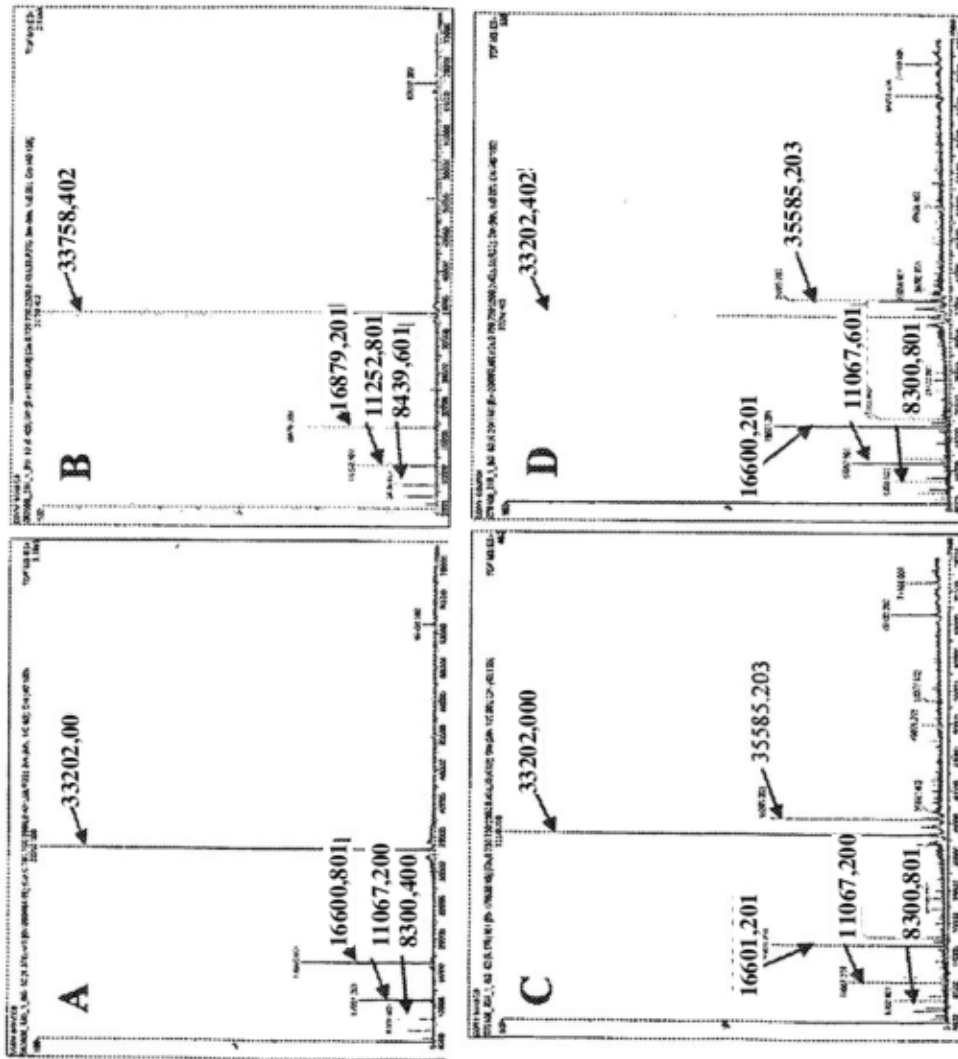


Figura 33

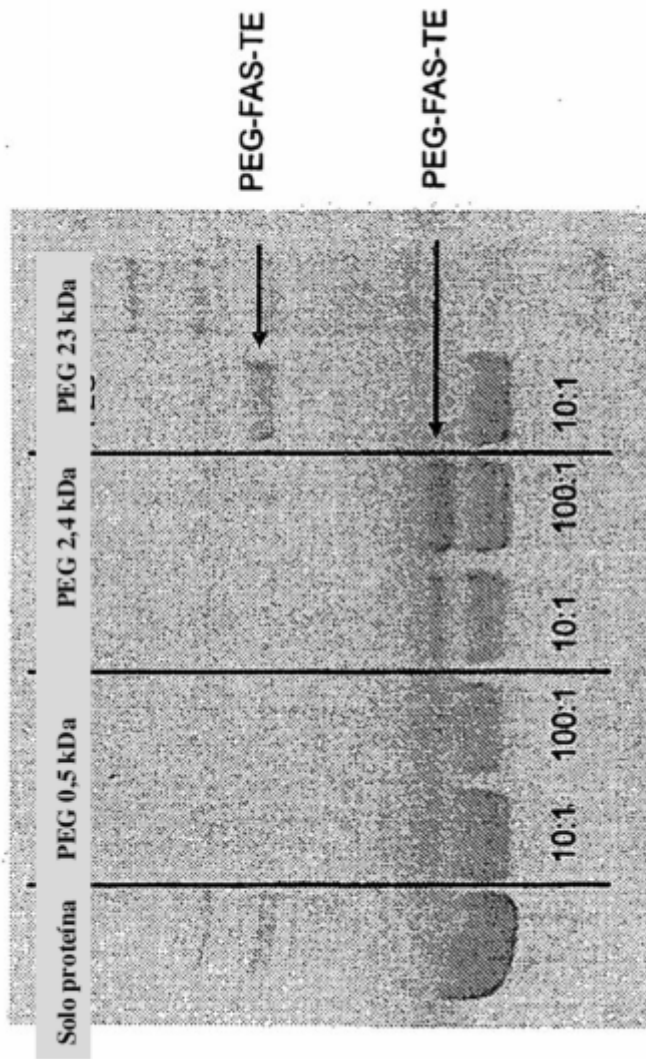


Figura 34

Gel SDS de las reacciones de PEGilación: (después de 60 h a 4° C)

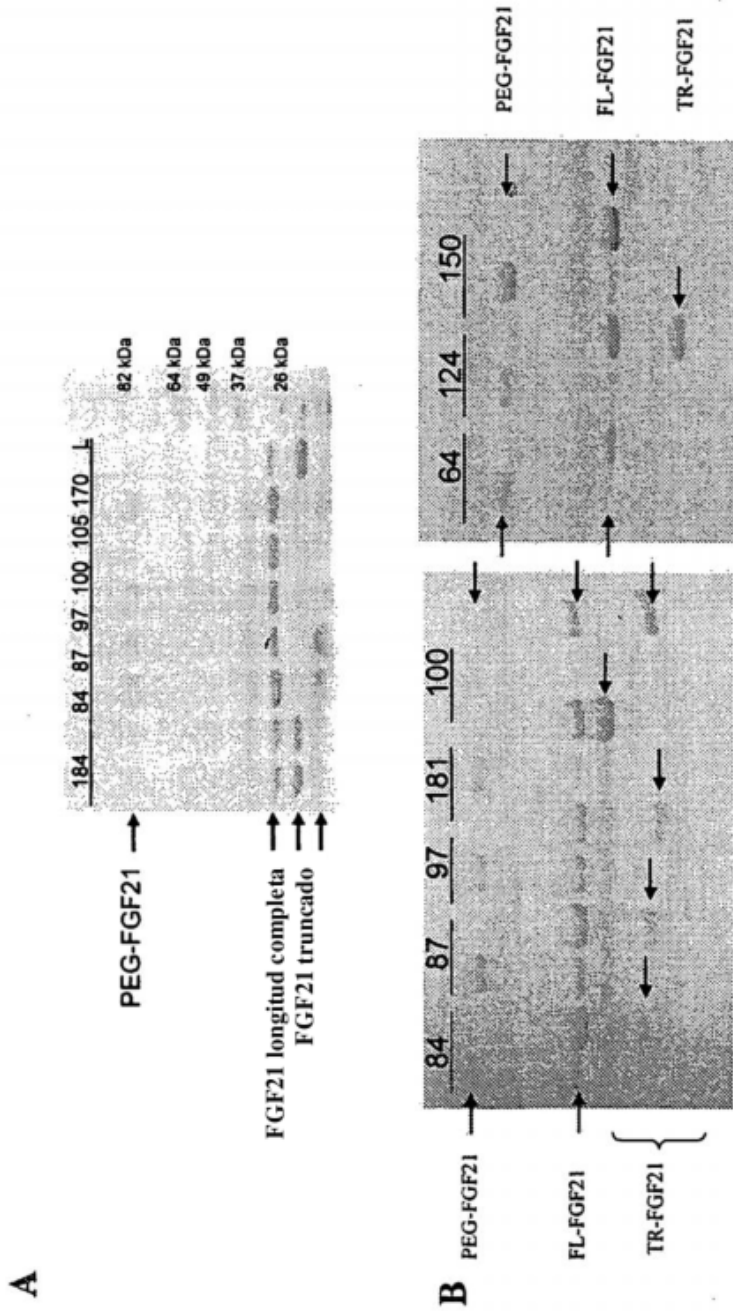


Figura 36

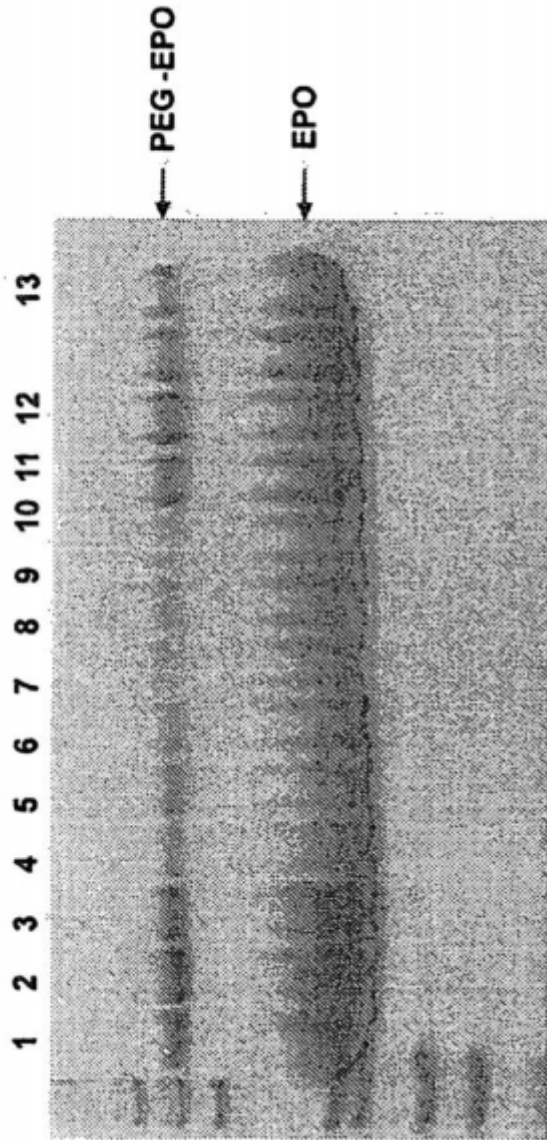


Figura 37

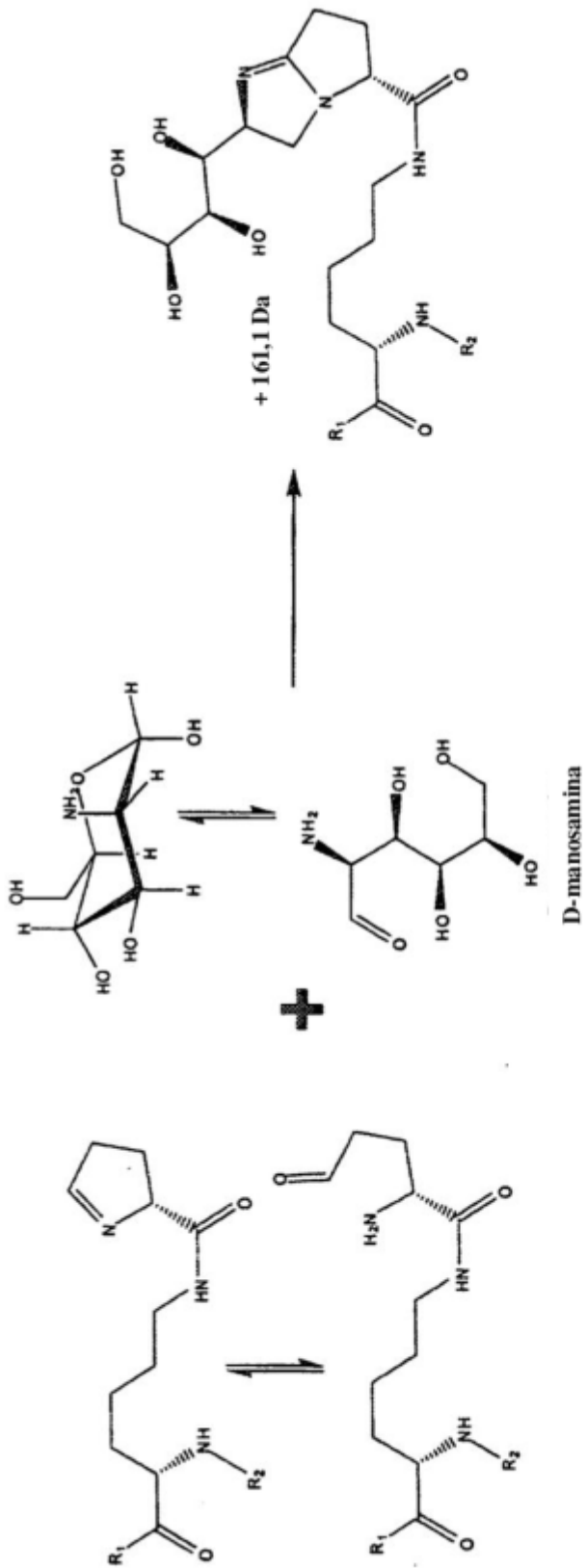


Figura 38

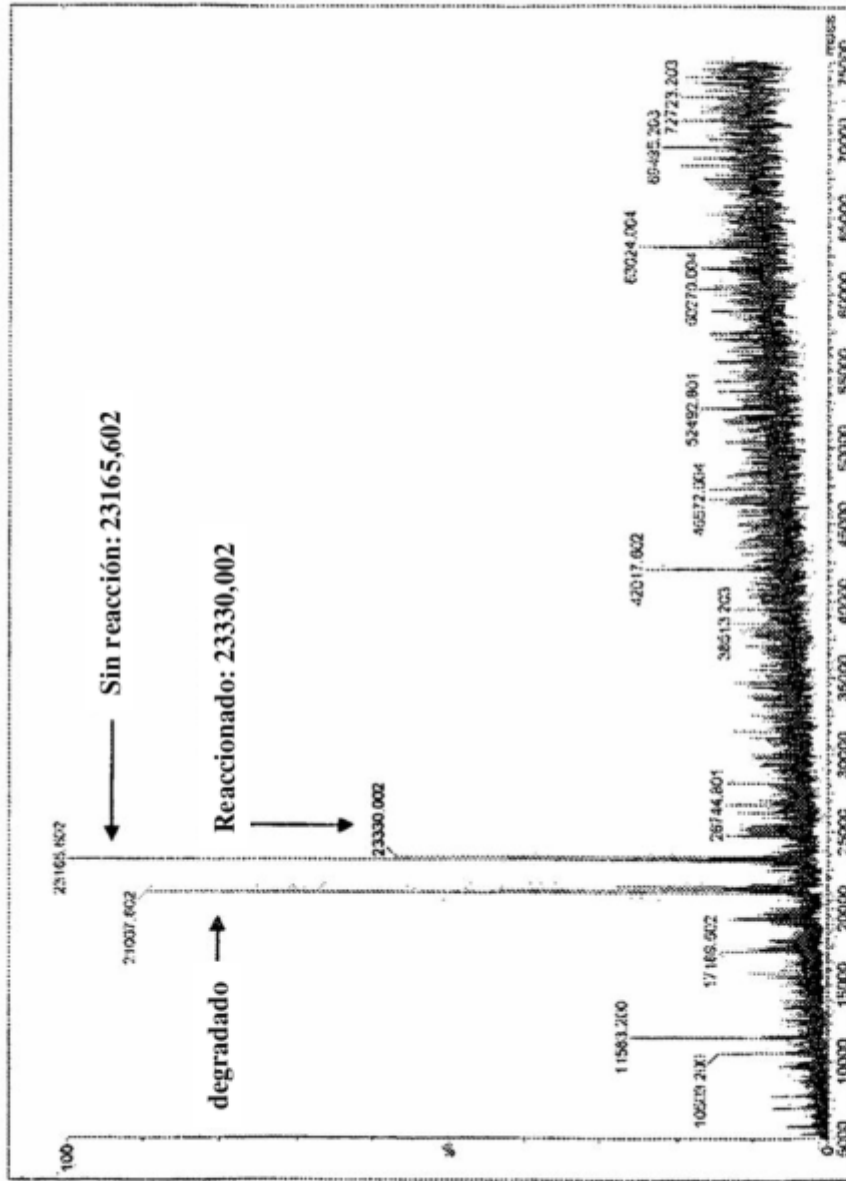


Figura 39

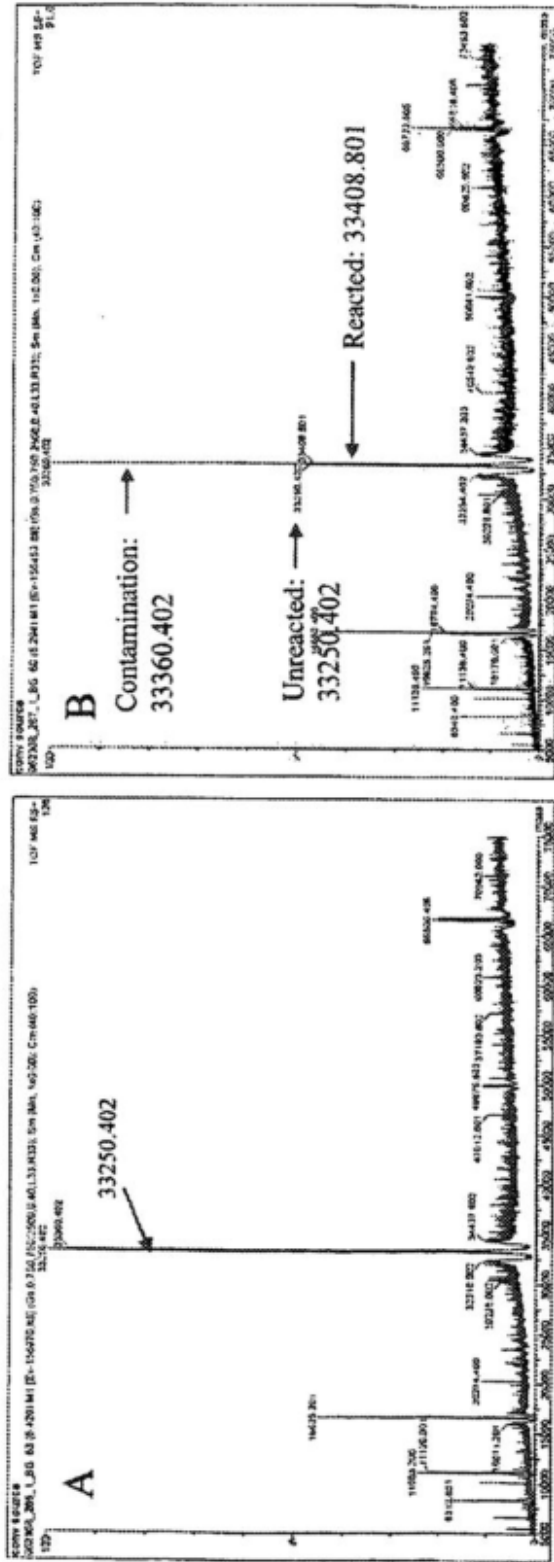


Figure 40

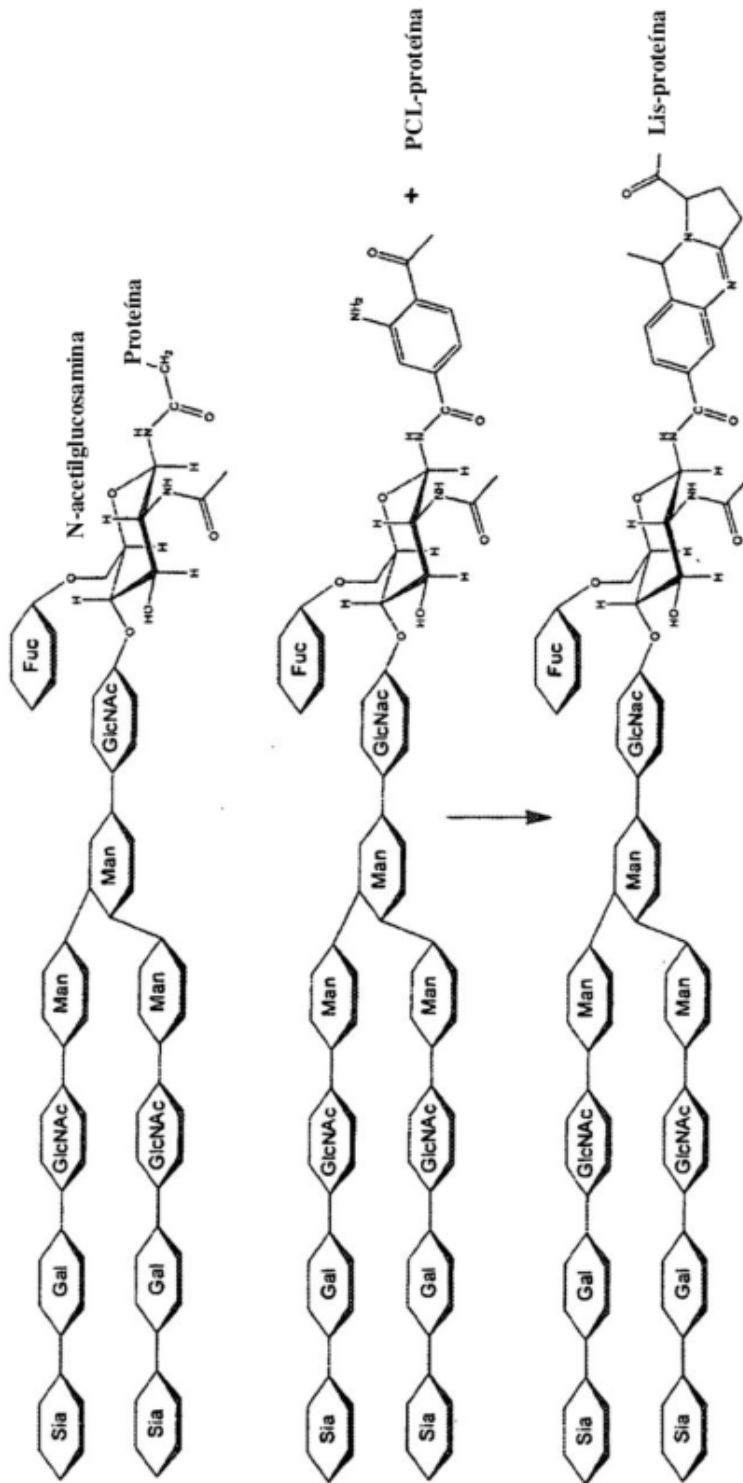


Figura 41

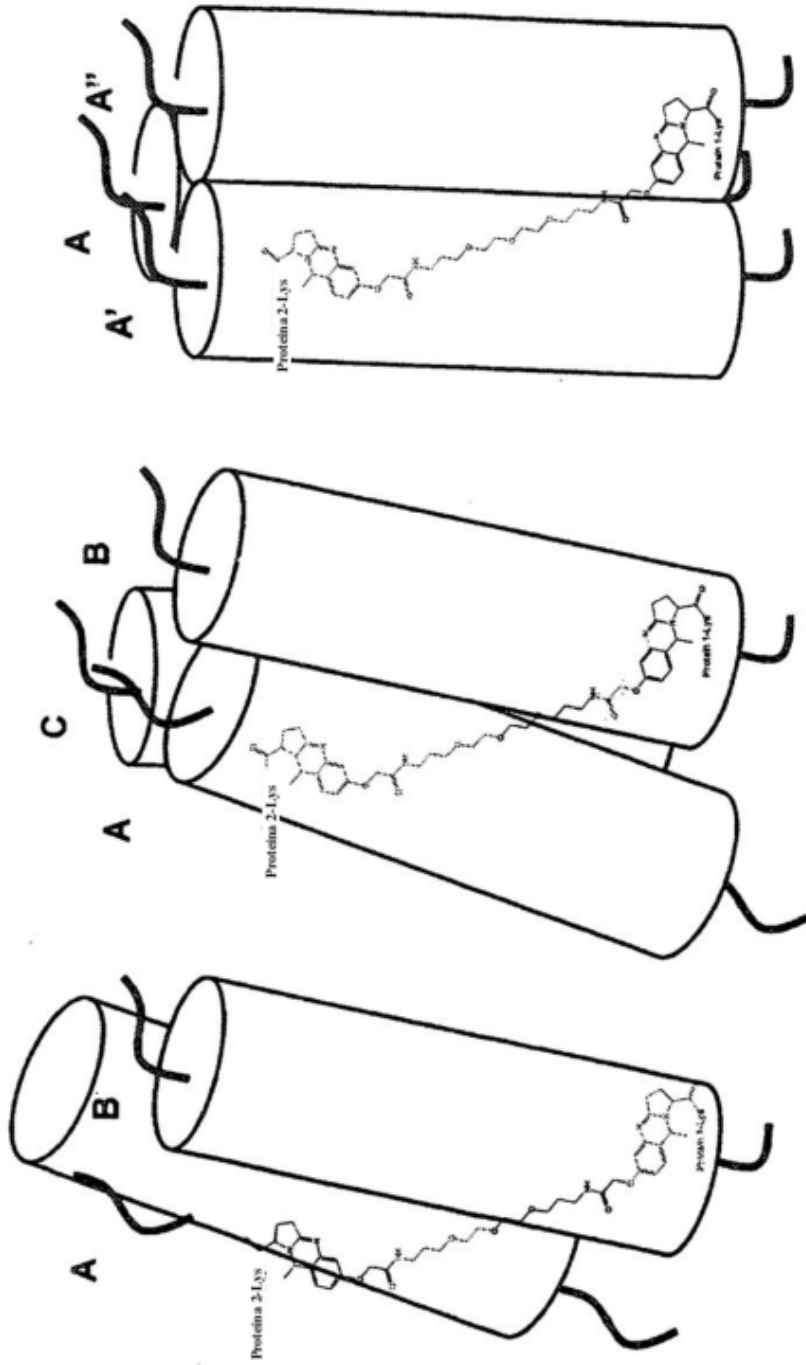


Figura 42

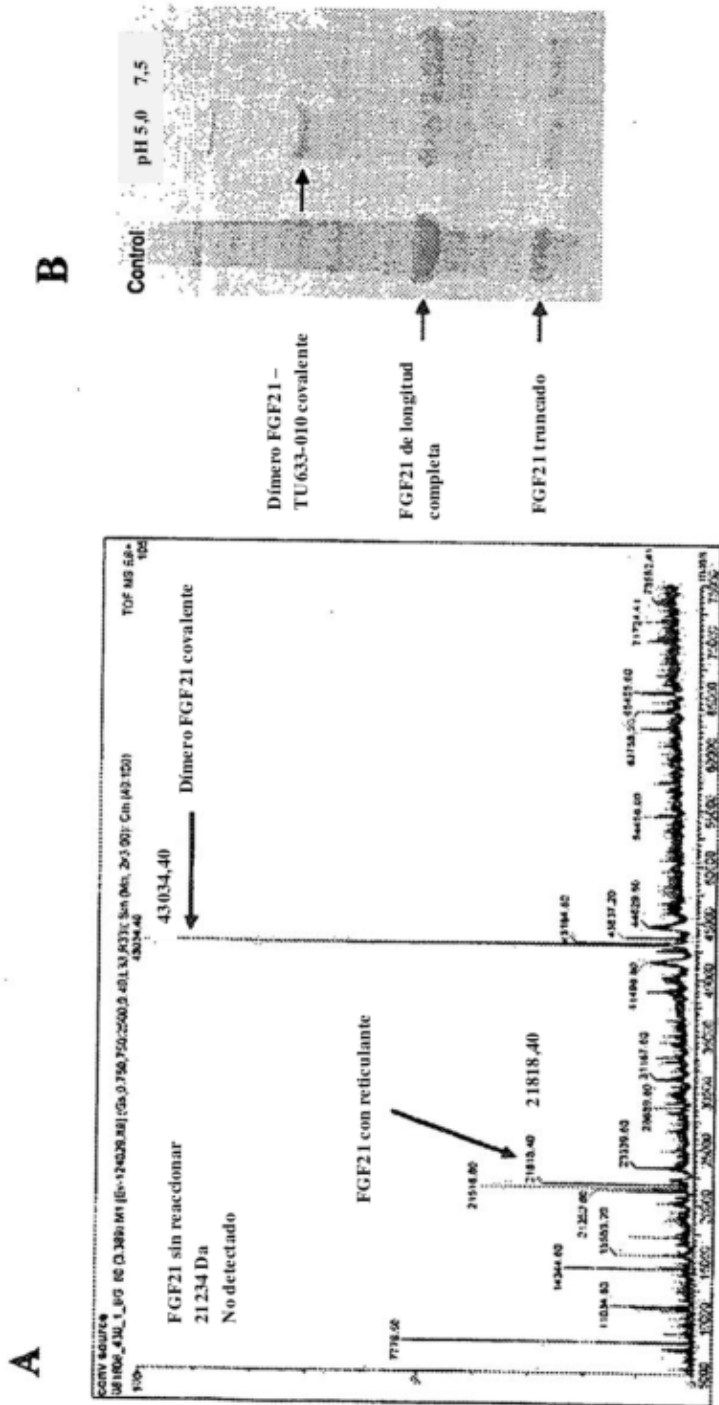


Figura 44

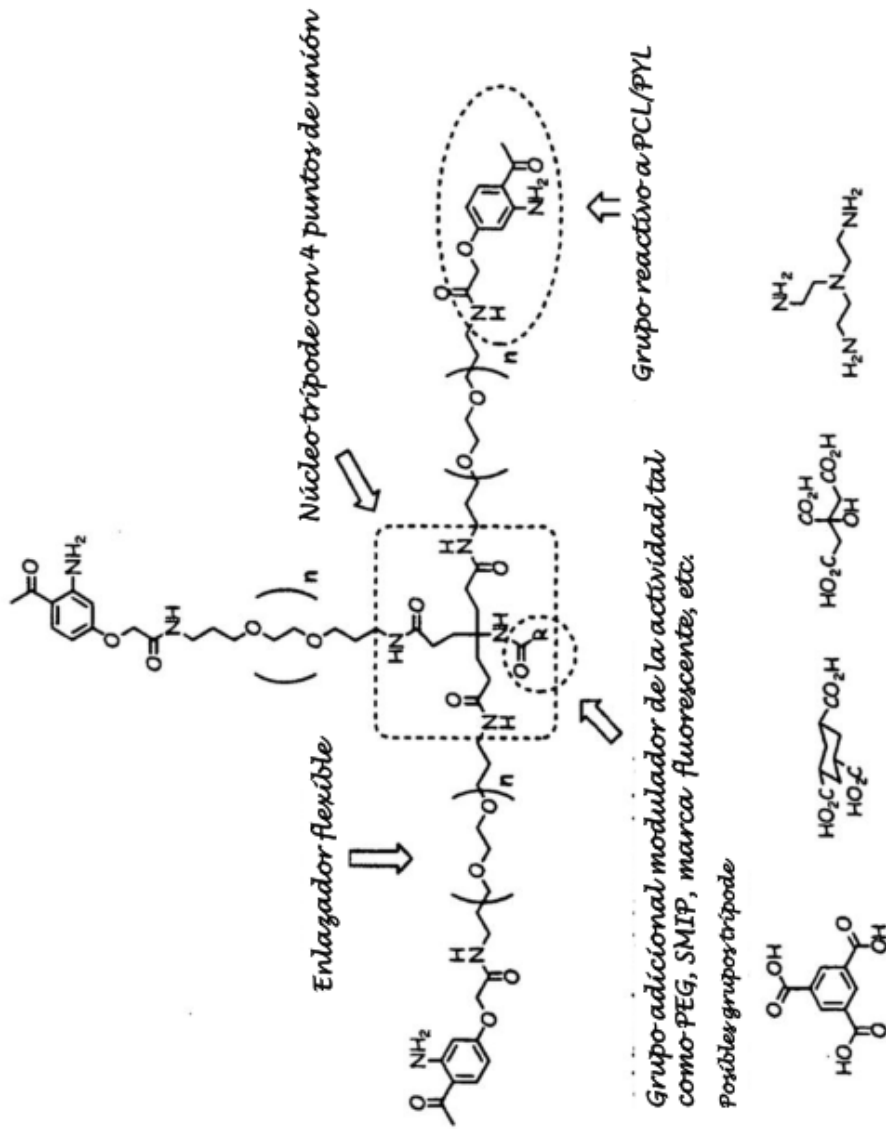


Figura 45

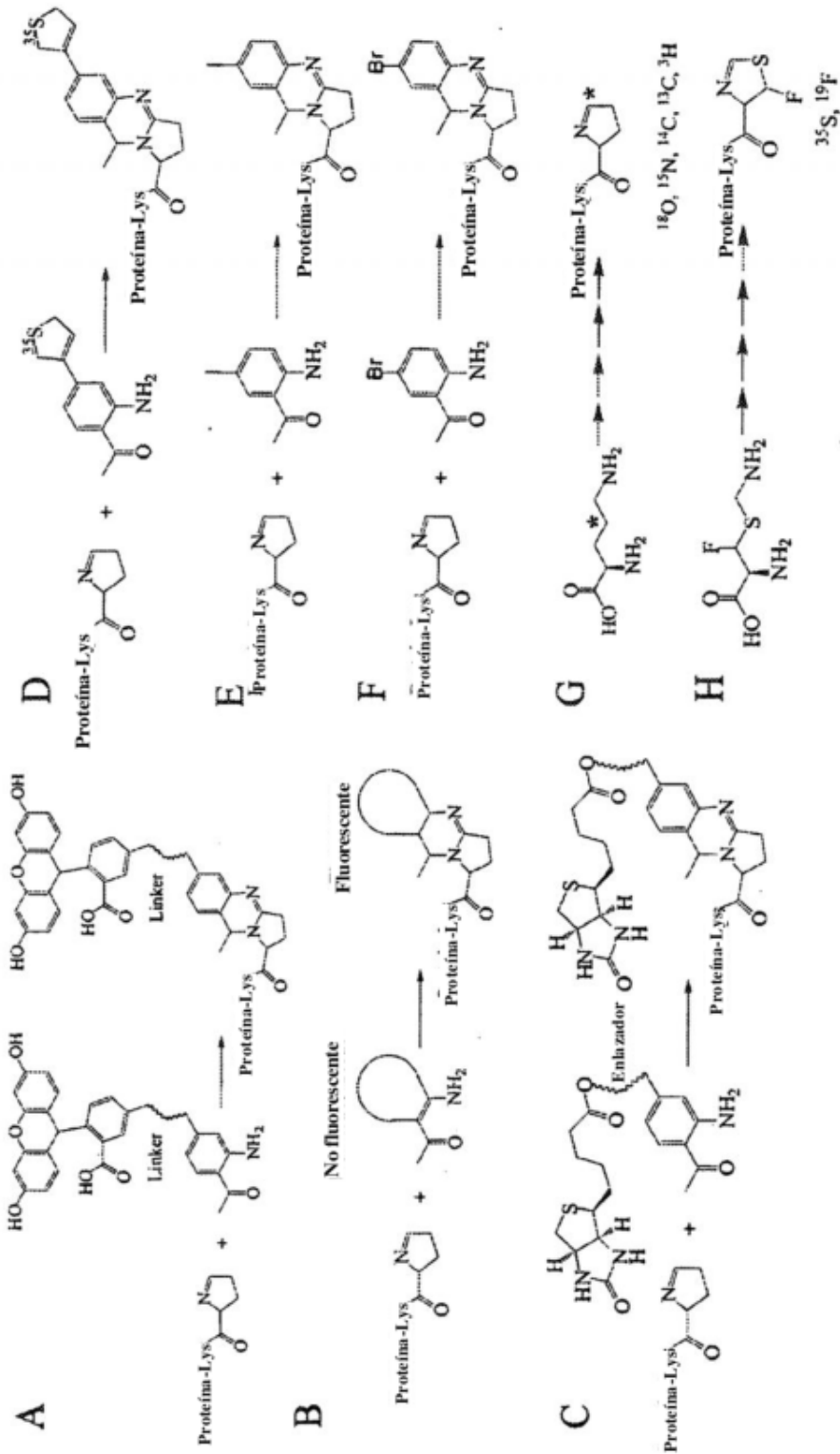


Figura 46

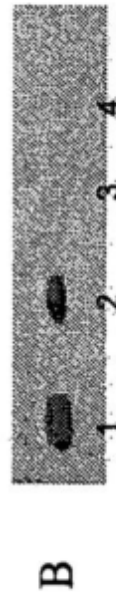
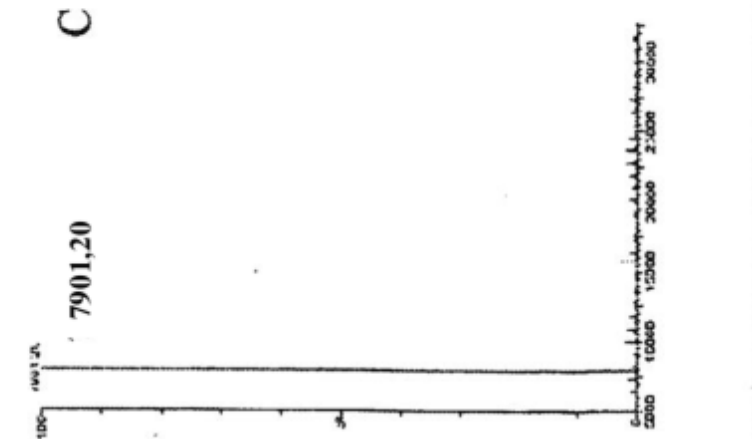
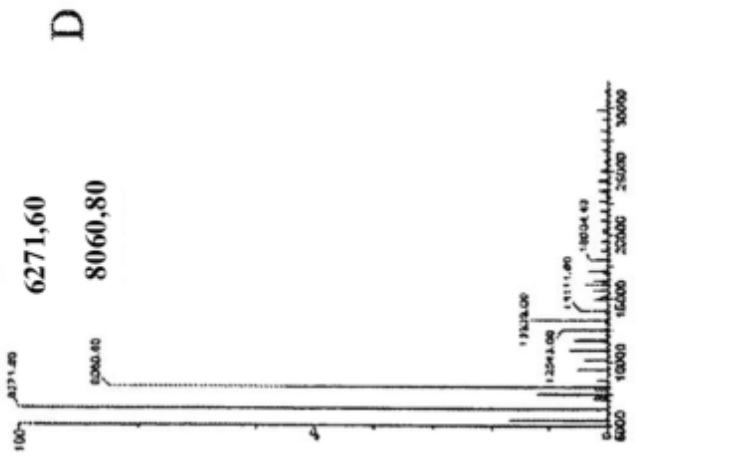
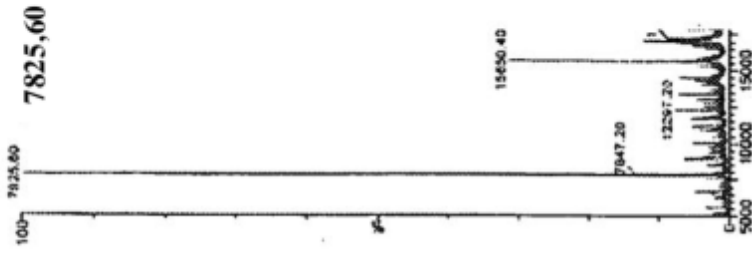


Figure 47

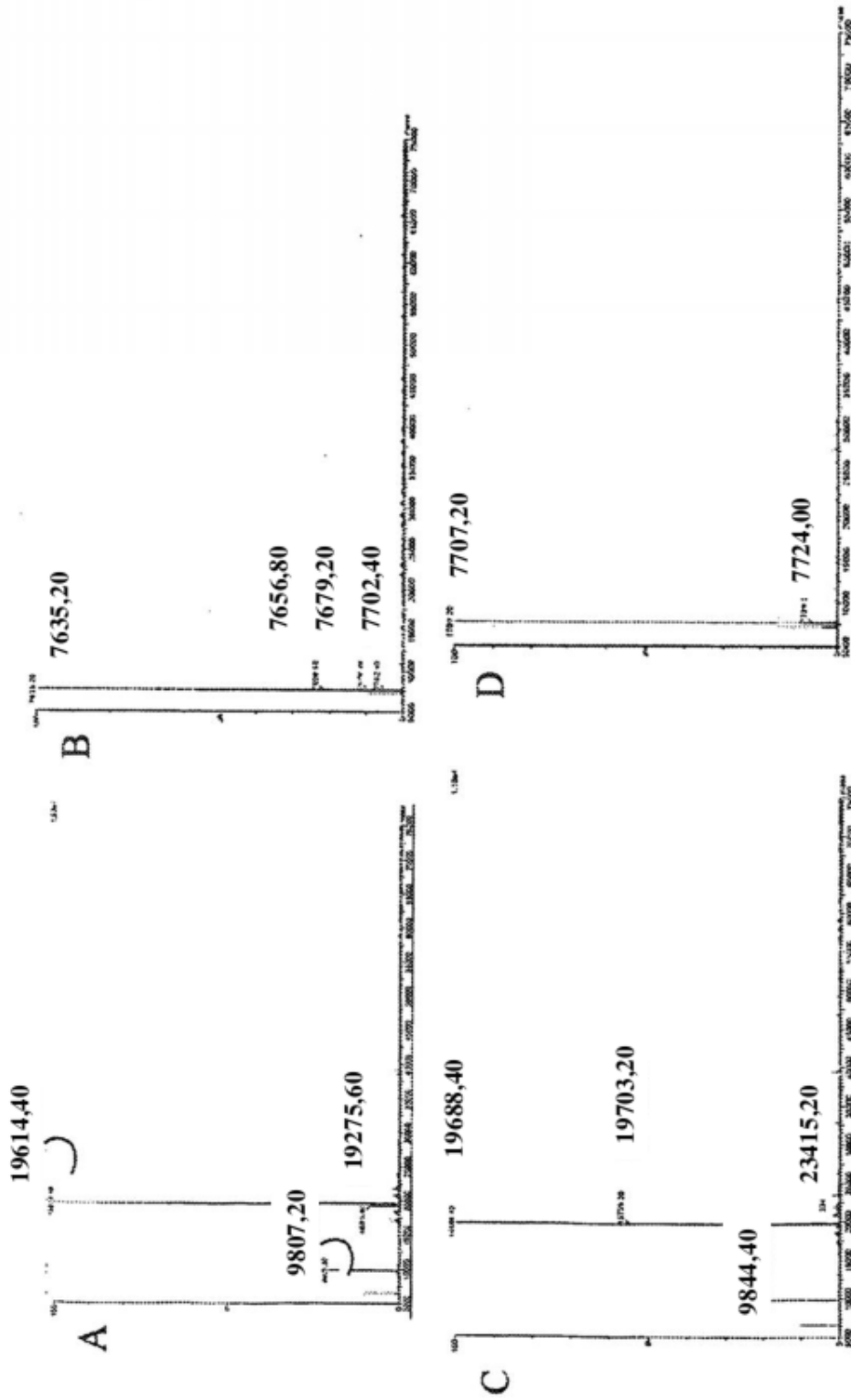


Figura 48

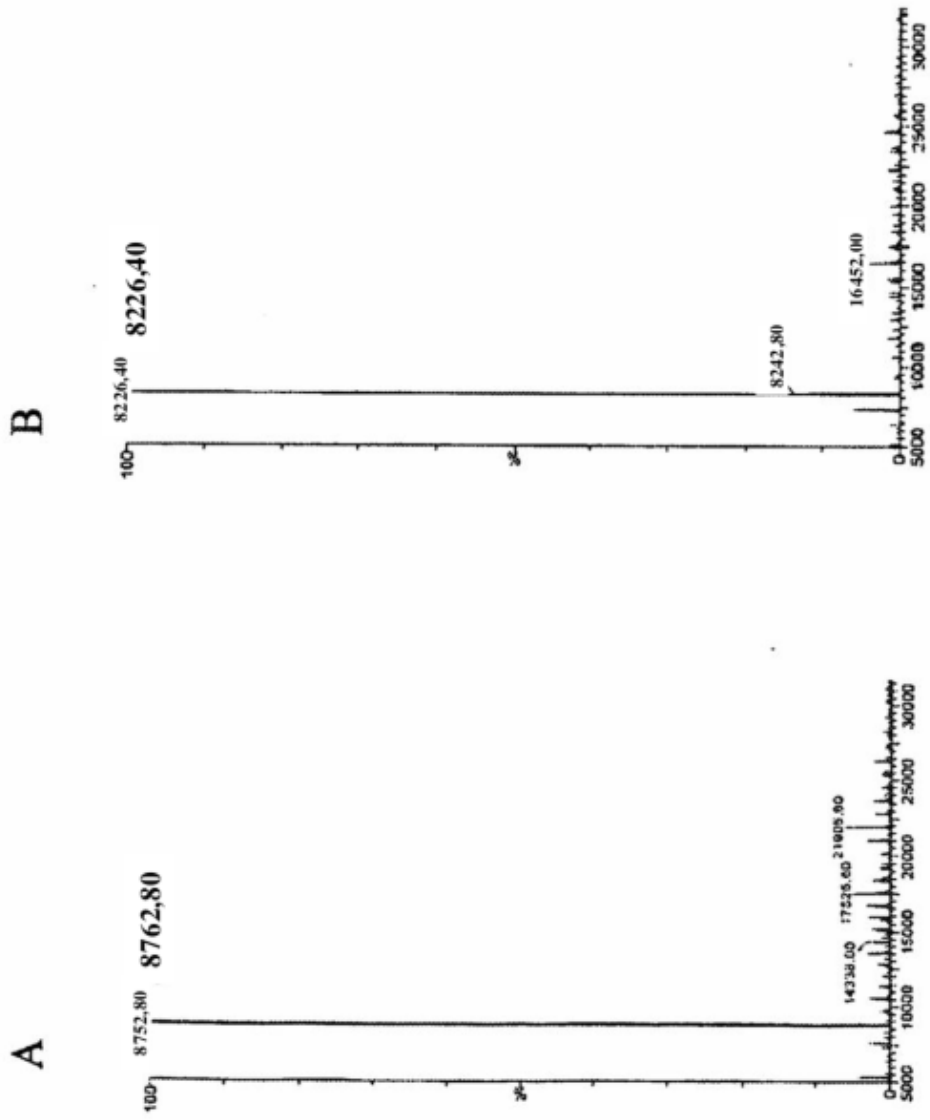


Figura 49

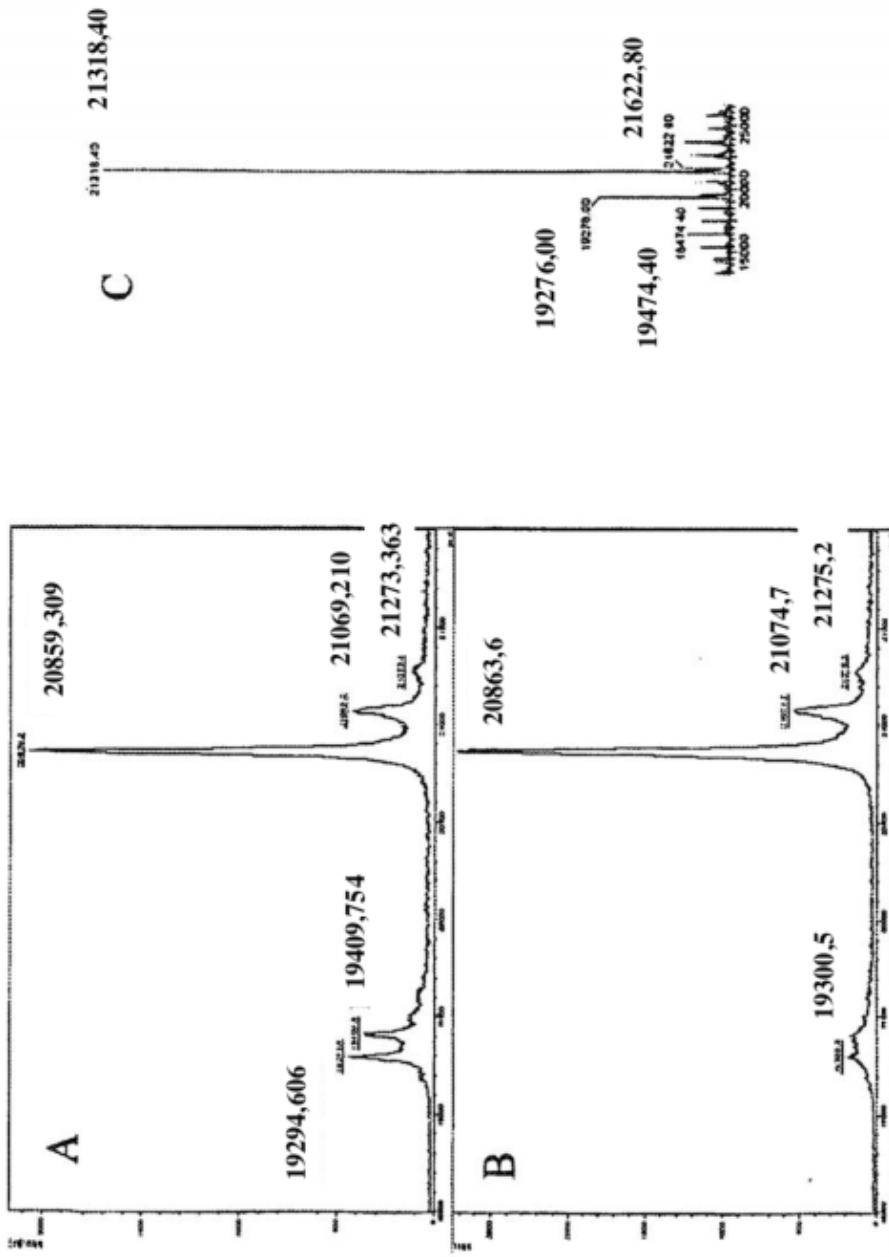


Figura 50

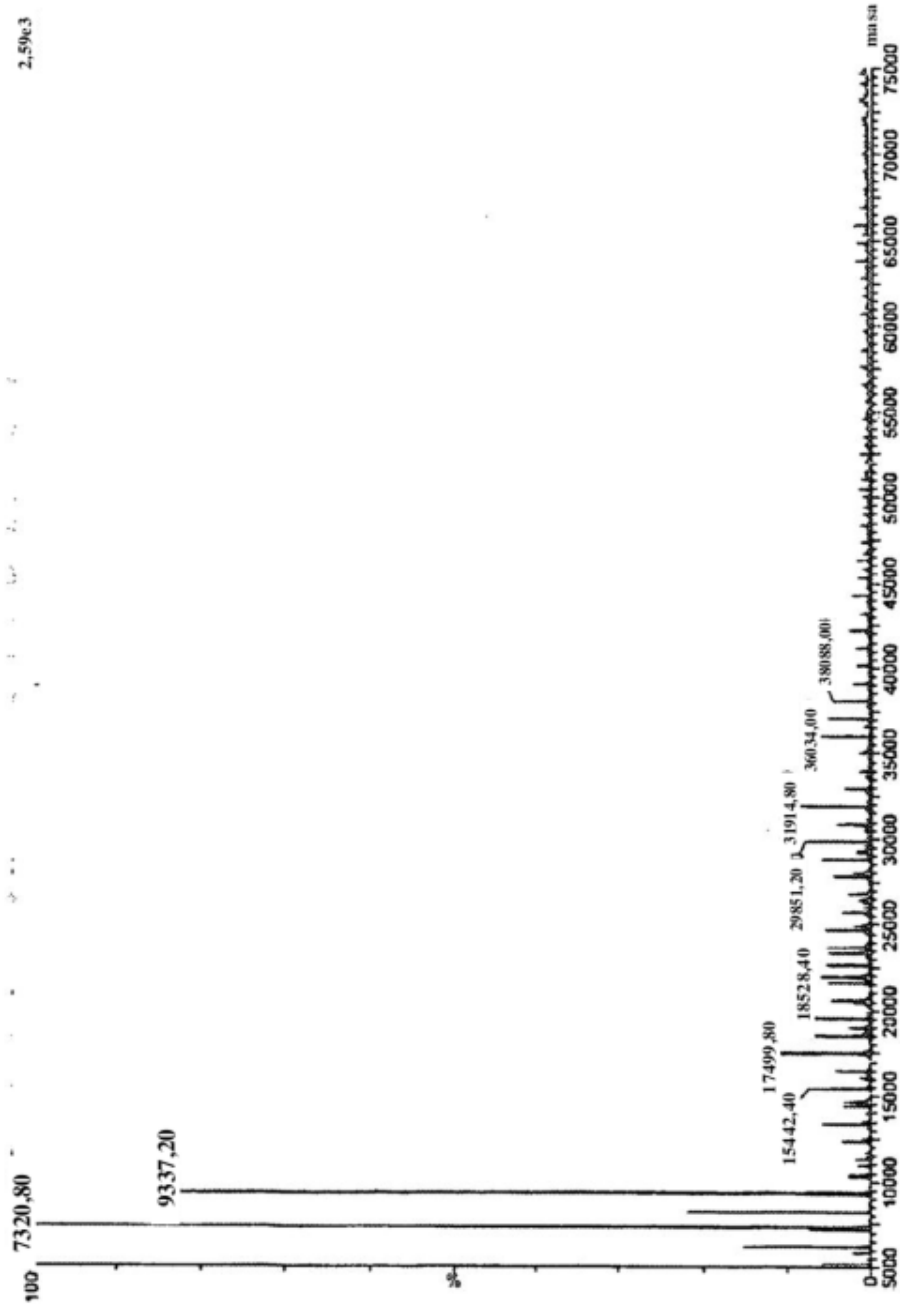


Figura 51

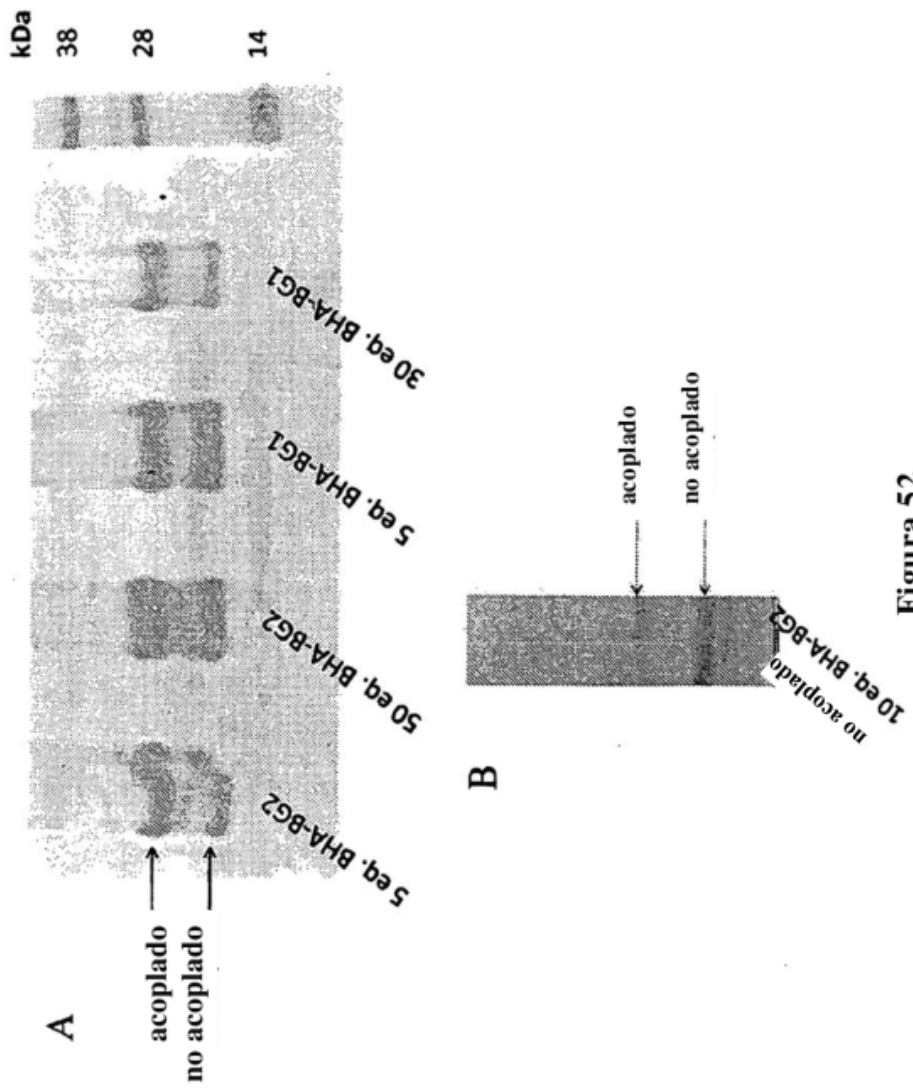


Figura 52

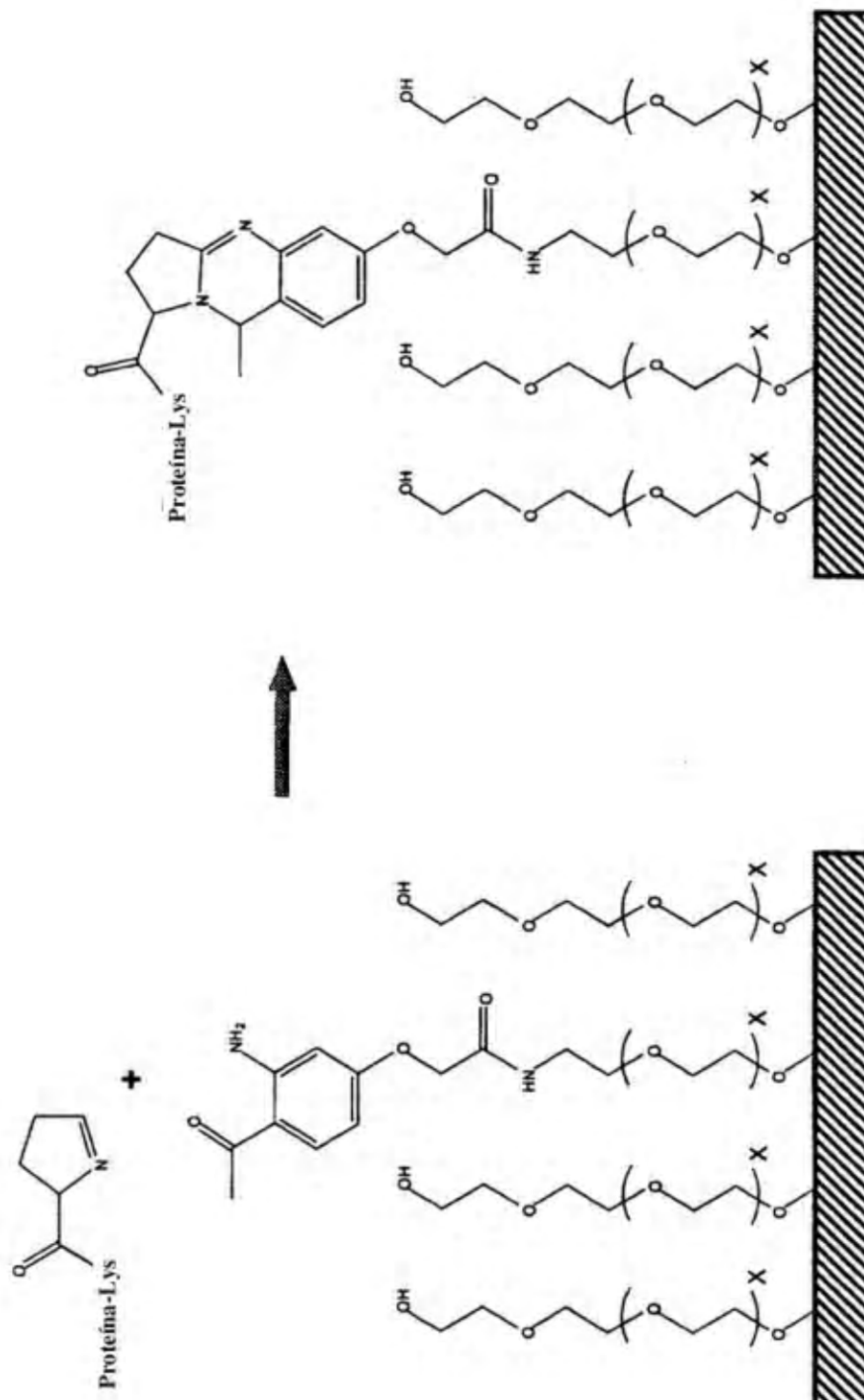


Figura 53

La reducción de NaCNBH3 evita la disociación de los conjugados de proteína de tipo PCL

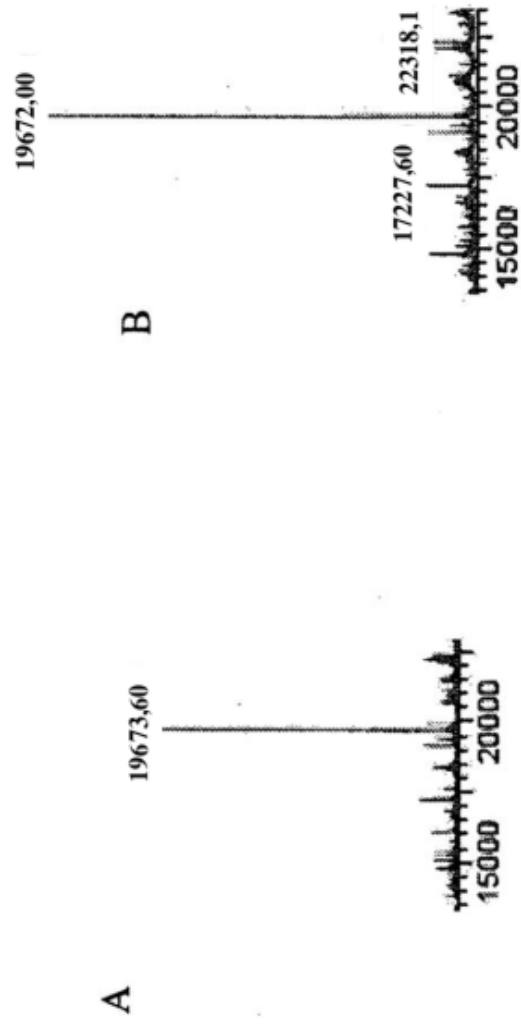


Figura 54

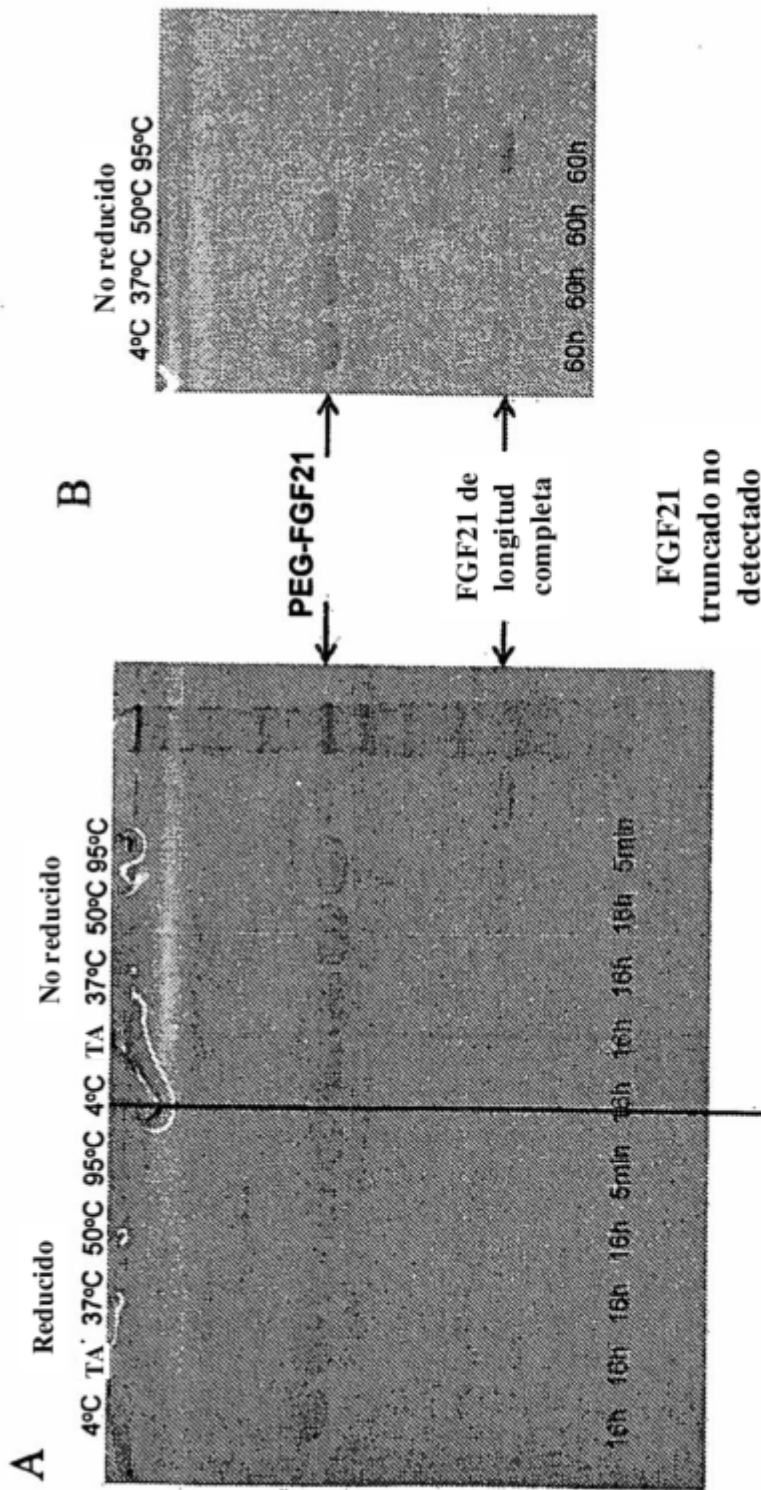


Figura 55

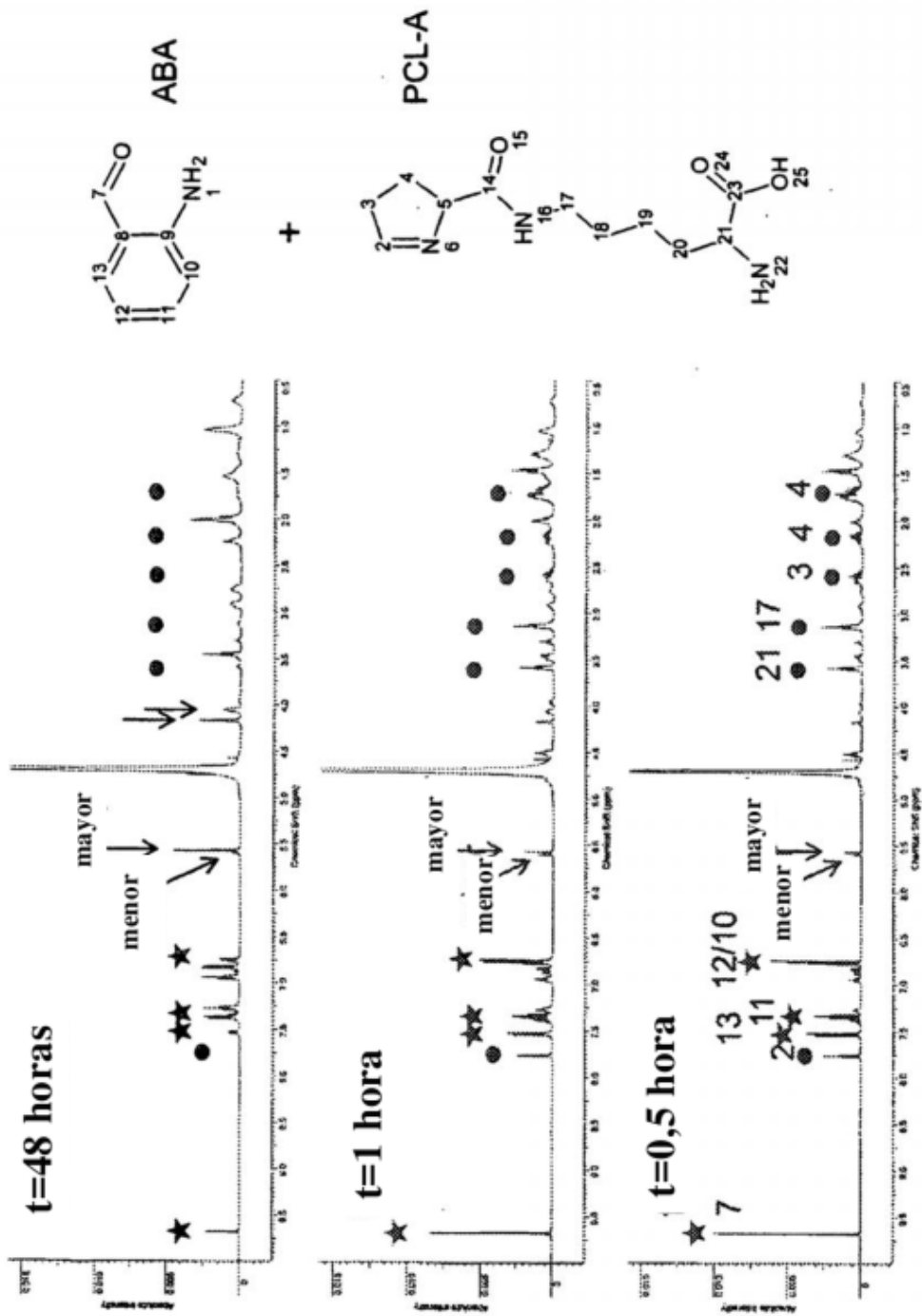


Figura 56

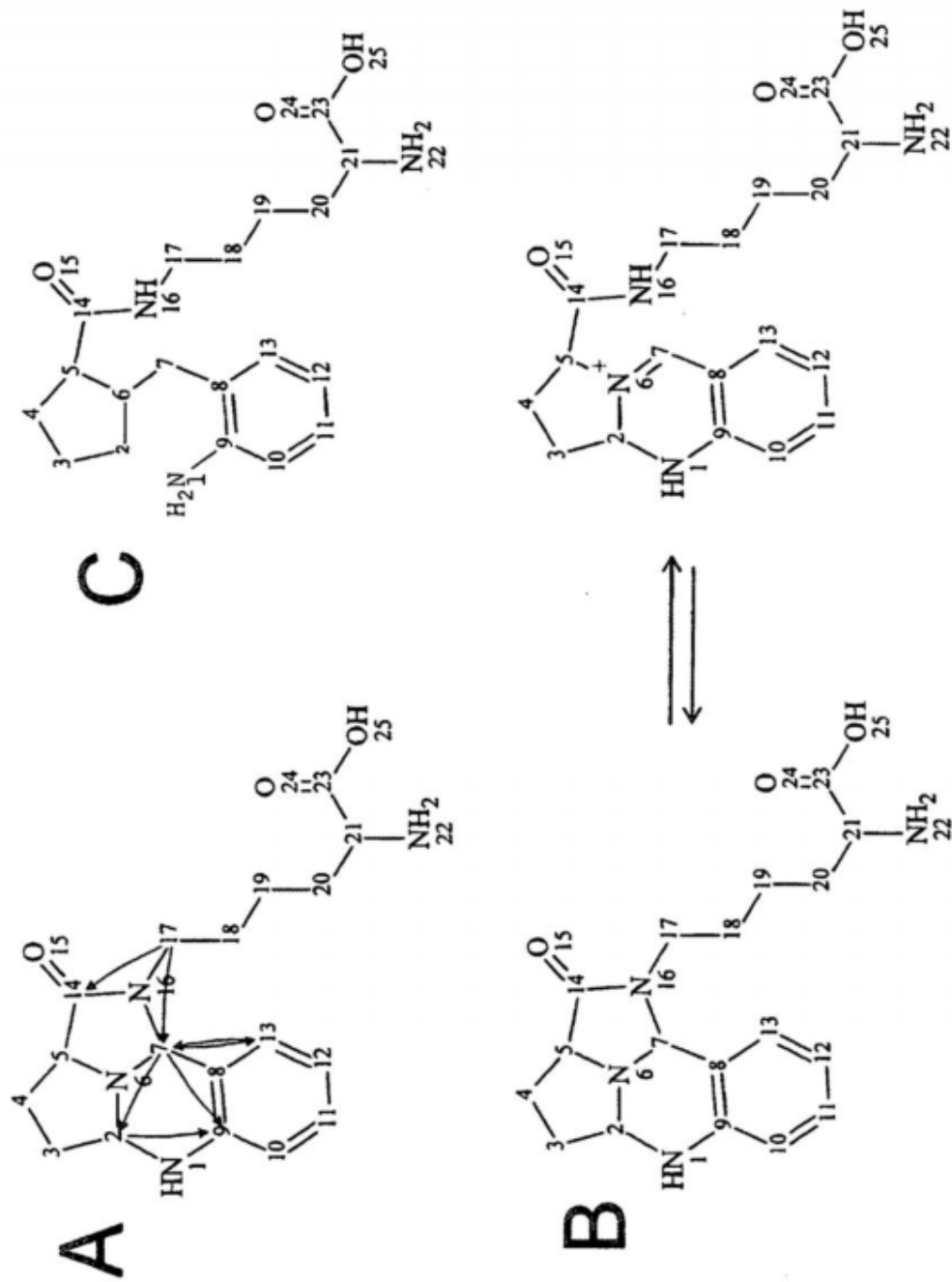


Figura 57

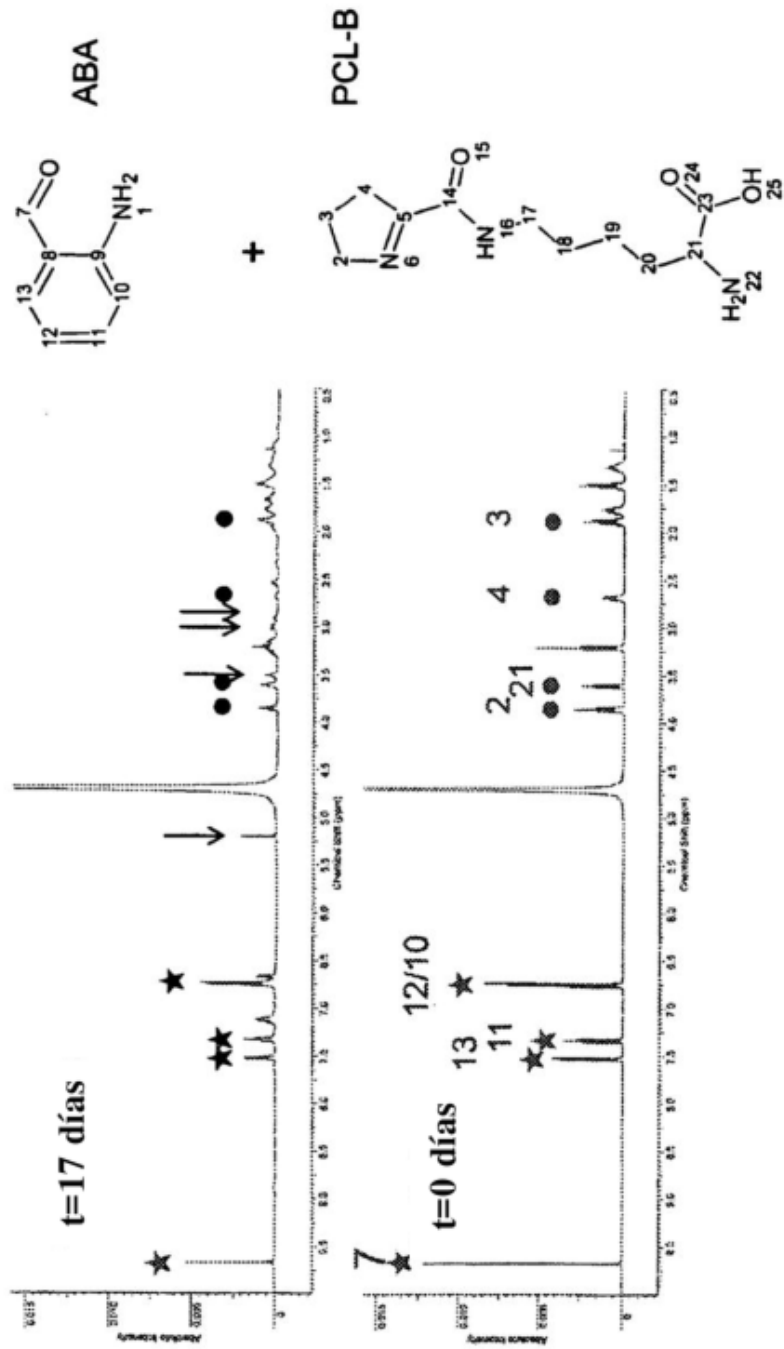


Figura 58

