(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 529 065**

(51) Int. CI.:

A61K 38/00 (2006.01)
A61K 38/16 (2006.01)
A61K 38/19 (2006.01)
C07K 7/00 (2006.01)
C07K 14/00 (2006.01)
C07K 1/00 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **13.12.2006 E 06848566 (3)**

(97) Fecha y número de publicación de la concesión europea: **17.09.2014 EP 1968635**

(54) Título: **Composiciones que contienen, procedimientos que implican y usos de aminoácidos no naturales y polipéptidos**

(30) Prioridad:

14.12.2005 US 743041 P
14.12.2005 US 743040 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
16.02.2015

(73) Titular/es:

AMBRX, INC. (100.0%)
10975 NORTH TORREY PINES ROAD, SUITE 100
LA JOLLA CA 92037, US

(72) Inventor/es:

MIAO, ZHENWEI;
LIU, JUNJIE y
NORMAN, THEA

(74) Agente/Representante:

IZQUIERDO BLANCO, María Alicia

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

Composiciones que contienen, procedimientos que implican y usos de aminoácidos no naturales y polipéptidos

SOLICITUDES RELACIONADAS

CAMPO DE LA INVENCIÓN

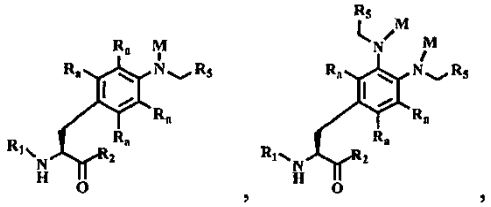
En el presente documento se describen procedimientos y composiciones de preparación y uso de agentes de aminoácidos no naturales.

ANTECEDENTES DE LA INVENCIÓN

La capacidad para incorporar aminoácidos no genéticamente codificados (es decir, "aminoácidos no naturales") en proteínas permite la introducción de grupos funcionales químicos que podrían proporcionar valiosas alternativas a los grupos funcionales que se producen naturalmente, tales como el épsilon -NH_2 de lisina, el sulfhidrilo -SH de cisteína, el grupo imino de histidina, etc. Ciertos grupos funcionales químicos son conocidos por ser inertes a los grupos funcionales encontrados en los 20 aminoácidos comunes genéticamente codificados, pero reaccionan limpia y eficazmente para formar enlaces estables con grupos funcionales que pueden incorporarse sobre aminoácidos no naturales.

Ahora están disponibles procedimientos para introducir selectivamente grupos funcionales químicos que no se encuentran en proteínas, que son químicamente inertes a todos de los grupos funcionales encontrados en los 20 aminoácidos comunes genéticamente codificados y que pueden usarse para reaccionar eficazmente y selectivamente con reactivos que comprenden ciertos grupos funcionales para formar enlaces covalentes estables.

En un aspecto, la presente invención proporciona un compuesto o sal del mismo seleccionado del grupo que consiste en:



en el que cada R_a está seleccionado independientemente del grupo que consiste en H, halógeno, alquilo, -NO_2, -CN, alquilo sustituido, -$N(R')_2$, -$C(O)_kR'$, -$C(O)N(R')_2$, -OR' y -$S(O)_kR'$ en la que k es 1, 2 ó 3;

M es H o -CH_2R_5;

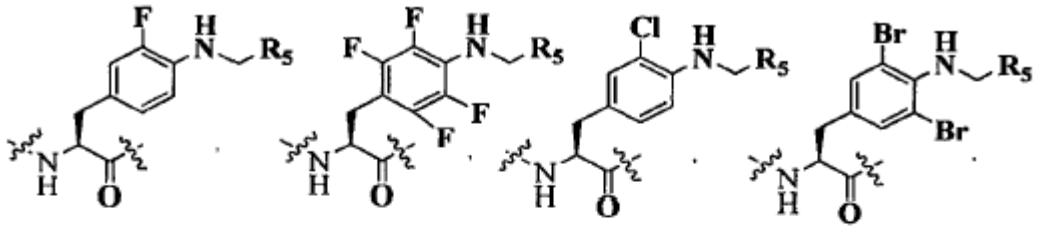
R_1 es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido;

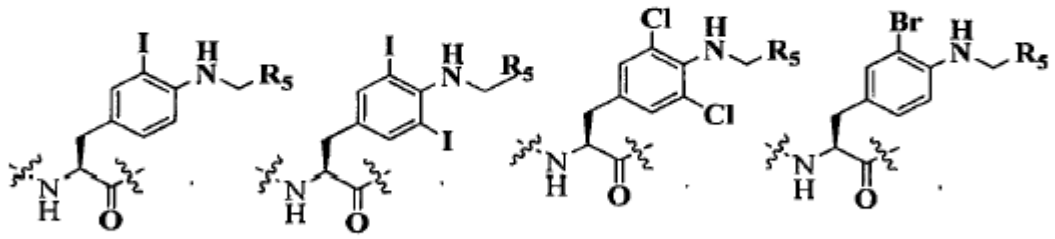
R_2 es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;

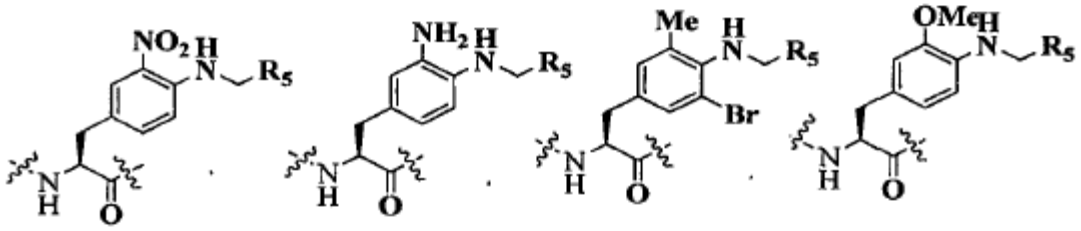
R_5 es L-X, en la que X es un polímero que comprende alquilo, alquenilo, alquinilo, alcoxi, alquilalcoxi, poli(óxido de alquileno), arilo, heteroarilo, alcarilo o aralquilo; y L es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno, alquenileno, -O-, -O-(alquilen)-, -(alquilen)-O-, -C(O)-, -C(O)-(alquilen)-, -(alquilen)-C(O)-, -C(O)N(R')-, -C(O)N(R')-(alquilen)-, -(alquilen)-C(O)N(R'), -OC(O)N(R')-, -OC(O)N(R')-(alquilen)-, -(alquilen)-OC(O)N(R')-, -N(R')C(O)-, -NR'C(O)-(alquilen)-, -(alquilen)-NR'C(O)-, -S-, -S-(alquilen)-, -$S(O)_k$- en la que k es 1, 2 ó 3, -$S(O)_k$(alquilen)-, -C(S)-, -C(S)-(alquilen)-, -CSN(R')-, -CSN(R')-(alquilen)-, -N(R')CO-(alquilen)-, -N(R')C(O)O-, -(alquilen)-O-N=CR'-, -(alquilen)-C(O)NR'-(alquilen)-, -(alquilen)-$S(O)_k$-(alquilen o alquilen sustituido)-S-, -(alquilen)-S-S-, -$S(O)_kN(R')$-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')$S(O)_kN(R')$-, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, -$C(R')_2$-N=N- y -$C(R')_2$-N(R')-N(R')-; y

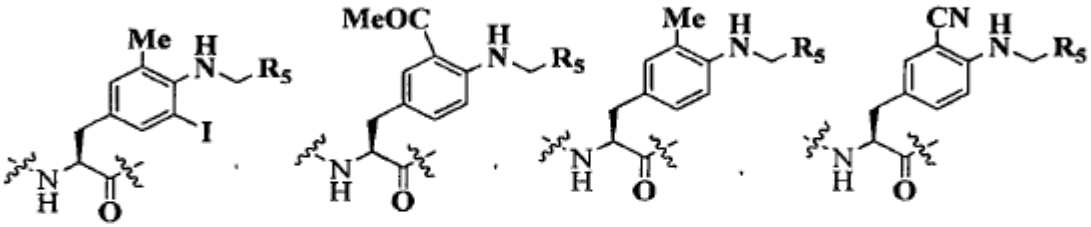
cada R' es independientemente H, alquilo o alquilo sustituido.

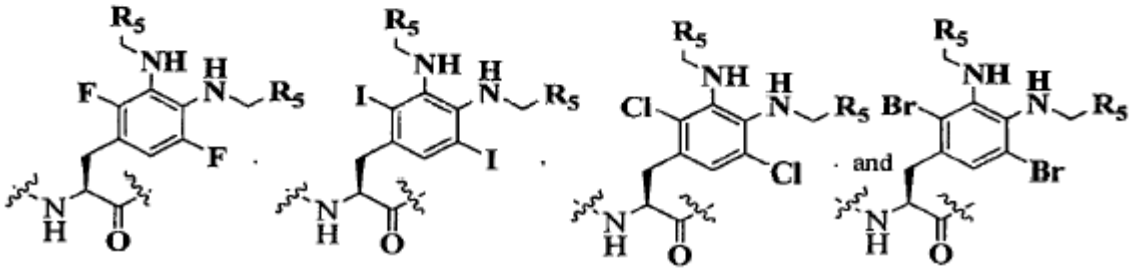
en el que el compuesto está seleccionado del grupo:



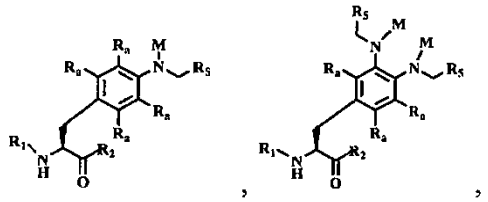




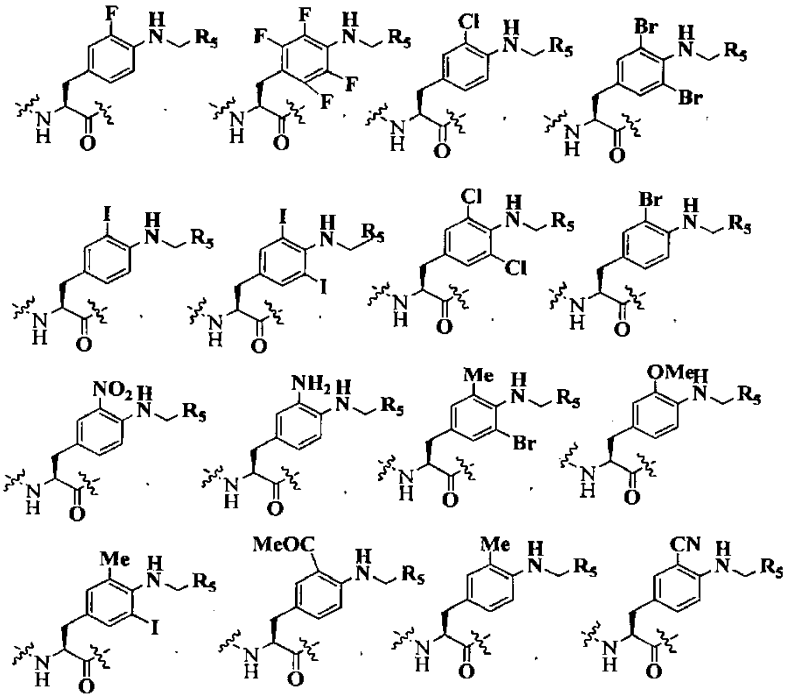




En otro aspecto, la presente invención proporciona un procedimiento de preparación de un compuesto que contiene al menos un aminoácido no natural seleccionado del grupo que consiste en:



en el que el compuesto se forma por una alquilación reductora de un resto amino aromático sobre al menos un aminoácido no natural con al menos un reactante que comprende al menos un resto de aldehído, en el que R_1, R_2, R_5, R_a y M son como antes,
en el que el al menos un aminoácido no natural está seleccionado de un grupo:



RESUMEN DE LA INVENCIÓN

En el presente documento se describen procedimientos, composiciones, técnicas y estrategias de preparación, purificación, caracterización y uso de aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados. Un aspecto de la invención es un procedimiento de derivatización de un aminoácido no natural y/o un polipéptido de aminoácidos no naturales. El procedimiento, implica derivatización química. Tales derivatizaciones pueden ser regioselectivas. Tales derivatizaciones pueden ser regioespecíficas. Tales derivaciones pueden ser estequiométricas, casi estequiométricas o similares a estequiométricas en tanto el reactivo que contiene el aminoácido no natural como el reactivo derivatizante. El procedimiento permite la incorporación estequiométrica, casi estequiométrica o similar a estequiométrica de un grupo deseado sobre un polipéptido de aminoácidos no naturales. También se desvelan estrategias, mezclas de reacción, condiciones sintéticas que permiten la incorporación estequiométrica, casi estequiométrica o similar a estequiométrica de un grupo deseado sobre el polipéptido de aminoácidos no naturales. Tales derivatizaciones pueden ser rápidas a temperatura ambiente. Tales derivatizaciones pueden producirse en disoluciones acuosas. En otras realizaciones o realizaciones adicionales, tales derivatizaciones se producen a un pH entre aproximadamente 4 y aproximadamente 10. En otras realizaciones o realizaciones adicionales, tales derivatizaciones se producen a un pH entre aproximadamente 4 y aproximadamente 7. En otras realizaciones o realizaciones adicionales, tales derivatizaciones se producen a un pH entre aproximadamente 4 y aproximadamente 5. En otras realizaciones o realizaciones adicionales, tales derivatizaciones se producen a un pH de aproximadamente 5. En otras realizaciones o realizaciones adicionales, tales derivatizaciones se producen a un pH de aproximadamente 4.

En un aspecto de la invención son aminoácidos no naturales para la derivatización química de péptidos y proteínas basada en la reactividad de un grupo de amina aromática. En otras realizaciones o realizaciones adicionales, al menos uno de los aminoácidos no naturales anteriormente mencionados se incorpora en un polipéptido, es decir, tales realizaciones son polipéptidos de aminoácidos no naturales. Los aminoácidos no naturales pueden funcionalizarse en sus cadenas laterales de forma que su reacción con una molécula derivatizante genere un enlace

de amina. Los aminoácidos no naturales están seleccionados de aminoácidos que tienen cadenas laterales de amina aromática. Los aminoácidos no naturales pueden comprender una cadena lateral enmascarada, que incluye un grupo de amina aromática enmascarada.

Los aminoácidos no naturales comprenden cadenas laterales de amina aromática en las que la amina aromática es una arilamina o realización adicional, los aminoácidos no naturales se parecen a un aminoácido natural en la estructura, pero contienen grupos amina aromática. Los aminoácidos no naturales se parecen a fenilalanina o tirosina (aminoácidos aromáticos). Los aminoácidos no naturales pueden tener propiedades que son distintas de aquellas de los aminoácidos naturales. Tales propiedades distintas son la reactividad química de la cadena lateral; esta reactividad química distinta permite que la cadena lateral del aminoácido no natural se someta a una reacción mientras que es una unidad de un polipéptido aún cuando las cadenas laterales de las unidades de aminoácido que se producen naturalmente en el mismo polipéptido no se someten a la reacción anteriormente mencionada. En otra realización, la cadena lateral del aminoácido no natural tiene una química ortogonal a las de los aminoácidos que se producen naturalmente. En otra realización, la cadena lateral del aminoácido no natural comprende un resto que contiene nucleófilo; en otra realización, el resto que contiene nucleófilo sobre la cadena lateral del aminoácido no natural pueden someterse a una reacción para generar una proteína derivatizada de amina. También se desvelan compuestos en los que la cadena lateral del aminoácido no natural comprende un resto que contiene electrófilo; el resto que contiene electrófilo sobre la cadena lateral del aminoácido no natural puede someterse a ataque nucleófilo para generar una proteína derivatizada de amina. En cualquiera de las realizaciones anteriormente mencionadas en este párrafo, el aminoácido no natural puede existir como una molécula separada o puede incorporarse en un polipéptido de cualquier longitud; si se produce lo último, entonces el polipéptido puede incorporar adicionalmente aminoácidos que se producen naturalmente o no naturales.

En un aspecto del procedimiento de la invención se usan moléculas sustituidas con carbonilo, tales como aldehídos, para la producción de polipéptidos de aminoácidos no naturales derivatizados basados en un enlace de amina. Se usan moléculas sustituidas con aldehído para derivatizar polipéptidos de aminoácidos no naturales que contienen aminas aromáticas mediante la formación de un enlace de amina entre la molécula derivatizante y el polipéptido de aminoácidos no naturales que contienen aminas aromáticas. Las moléculas sustituidas con aldehído del procedimiento de la invención también desveladas como sustituyentes de aldehído comprenden un grupo R_5 como se define en la reivindicación adjunta: una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor, una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos. En otras realizaciones o realizaciones adicionales, las moléculas sustituidas con aldehído son moléculas de polietilenglicol sustituidas con aldehído (PEG). La cadena lateral del aminoácido no natural usada en el procedimiento de la invención puede tener una química ortogonal a aquellos de los aminoácidos que se producen naturalmente que permite al aminoácido no natural reaccionar selectivamente con las moléculas sustituidas con aldehído. La cadena lateral del aminoácido no natural comprende un resto que contiene amina aromática que reacciona selectivamente con la molécula que contiene aldehído; el resto que contiene amina aromática sobre la cadena lateral del aminoácido no natural pueden someterse a reacción para generar una proteína derivatizada de amina alquilada. En otro aspecto relacionado con las realizaciones descritas en este párrafo son los polipéptidos de aminoácidos no naturales modificados que resultan de la reacción de la molécula derivatizante con los polipéptidos de aminoácidos no naturales. Otras realizaciones incluyen cualquier modificación adicional de los polipéptidos de aminoácidos no naturales ya modificados.

Las moléculas sustituidas con amina aromática son arilaminas que pueden usarse para la producción de polipéptidos de aminoácidos no naturales derivatizados basados en un enlace de amina. También se desvelaron moléculas sustituidas con amina aromática usadas para derivatizar polipéptidos de aminoácidos no naturales que contienen aldehído mediante la formación de un enlace de amina entre la molécula derivatizante y el polipéptido de aminoácidos no naturales que contienen aldehído. En otras realizaciones o realizaciones adicionales, las moléculas sustituidas con amina aromática comprenden un grupo seleccionado de: una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco;

una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos. En otras realizaciones o realizaciones adicionales, las moléculas sustituidas con amina aromática son moléculas de polietilenglicol (PEG) sustituidas con amina aromática. En otra realización, la cadena lateral del aminoácido no natural tiene una química ortogonal a aquellos de los aminoácidos que se producen naturalmente que permite al aminoácido no natural reaccionar selectivamente con las moléculas sustituidas con amina aromática. En otra realización, la cadena lateral del aminoácido no natural comprende un resto que contiene aldehído que reacciona selectivamente con la molécula que contiene amina aromática; en otra realización, el resto que contiene aldehído sobre la cadena lateral del aminoácido no natural puede someterse a reacción para generar una proteína derivatizada de amina alquilada. En otro aspecto relacionado con las realizaciones descritas en este párrafo son los polipéptidos de aminoácidos no naturales modificados que resultan de la reacción de la molécula derivatizante con los polipéptidos de aminoácidos no naturales. Otras realizaciones incluyen cualquier modificación adicional de los polipéptidos de aminoácidos no naturales ya modificados.

Los compuestos de la invención pueden comprender conectores mono-, bi- y multi-funcionales para la generación de polipéptidos de aminoácidos no naturales derivatizados basados en un enlace de amina. En una realización son conectores moleculares (bi- y multi-funcionales) que pueden usarse para conectar los polipéptidos de aminoácidos no naturales que contienen aminas aromáticas a otras moléculas. Los polipéptidos de aminoácidos no naturales que contienen aminas aromáticas comprenden una cadena lateral de arilamina en una realización que utiliza un polipéptido de aminoácidos no naturales que contienen aminas aromáticas. El conector molecular puede contener un grupo aldehído en uno de sus extremos. En otras realizaciones o realizaciones adicionales, las moléculas de conector sustituidas con aldehído son moléculas de conector de polietilenglicol (PEG) sustituidas con aldehído. En otras realizaciones, el término “otras moléculas” incluye, a modo de ejemplo solo, proteínas, otros polímeros y moléculas pequeñas. En otras realizaciones o realizaciones adicionales, los conectores moleculares que contienen aldehído comprenden los mismos grupos o equivalentes en todos los extremos de manera que tras la reacción con un polipéptido de aminoácidos no naturales que contienen aminas aromáticas, el producto resultante sea la homo-multimerización del polipéptido de aminoácidos no naturales que contienen aminas aromáticas. En otras realizaciones, la homo-multimerización es una homo-dimerización. En otras realizaciones o realizaciones adicionales, los conectores moleculares comprenden al menos un grupo aldehído y un grupo diferente en todos los extremos de manera que tras la reacción con un polipéptido de aminoácidos no naturales que contienen aminas aromáticas, el producto resultante sea la hetero-multimerización del polipéptido de aminoácidos no naturales que contienen aminas aromáticas. En otras realizaciones, la hetero-multimerización es una hetero-dimerización. En otra realización, la cadena lateral del aminoácido no natural tiene una química ortogonal a aquellos de los aminoácidos que se producen naturalmente que permite al aminoácido no natural reaccionar selectivamente con las moléculas de conector sustituidas con aldehído. En otro aspecto relacionado con las realizaciones descritas en este párrafo son los polipéptidos de aminoácidos no naturales modificados o no modificados ligados que resultan de la reacción de la molécula de conector con los polipéptidos de aminoácidos no naturales. Otras realizaciones incluyen cualquier modificación adicional de los polipéptidos de aminoácidos no naturales modificados o no modificados ya ligados.

El procedimiento de la invención derivatiza proteínas mediante la reacción de aminas aromáticas y reactantes de aldehído con proteína derivatizada de amina alquilada. Incluidos dentro de este aspecto están procedimientos para la derivatización de proteínas basada en la alquilación reductora de reactantes que contiene amina aromática y aldehído para generar aducto de proteína derivatizada de amina alquilada. En realizaciones adicionales u otras realizaciones son procedimientos de derivatización de proteínas que contienen aminas aromáticas con moléculas de polietilenglicol (PEG) funcionalizadas con aldehído. En todavía aspectos adicionales u otros aspectos, la molécula sustituida con aldehído puede incluir proteínas, otros polímeros (no ramificados y ramificados).

En el presente documento se describen procedimientos para la síntesis química de moléculas sustituidas con aldehído para la derivatización de proteínas sustituidas con amina aromática. La molécula sustituida con aldehído puede comprender péptidos, otros polímeros (no ramificados y ramificados) y moléculas pequeñas. También se describen procedimientos para la preparación de moléculas sustituidas con aldehído adecuadas para la derivatización de polipéptidos de aminoácidos no naturales que contienen aminas aromáticas. Los aminoácidos no

naturales, que incluyen, pero no se limitan a, aminoácidos no naturales que contienen aminas aromáticas, pueden incorporarse específicamente en el sitio durante la traducción *in vivo* de proteínas. Los aminoácidos no naturales, que incluyen aminoácidos no naturales que contienen aminas aromáticas, pueden incorporarse específicamente en el sitio por traducción ribosómica. Los aminoácidos no naturales, que incluyen, pero no se limitan a, aminoácidos no naturales que contienen aminas aromáticas, pueden incorporarse específicamente en el sitio durante la traducción *in vitro*. Las moléculas sustituidas con aldehído permiten la derivatización específica del sitio de aminoácidos no naturales que contienen aminas aromáticas mediante alquilación reductora del resto de amina aromática para formar un polipéptido derivatizado de amina alquilada en un modo específico de sitio. El procedimiento para la preparación de moléculas sustituidas con aldehído proporciona acceso a una amplia variedad de polipéptidos derivatizados específicamente en el sitio. También se describen procedimientos de síntesis de moléculas de polietilenglicol (PEG) funcionalizadas con aldehído.

Otro aspecto del procedimiento de la invención es la derivatización química de polipéptidos de aminoácidos no naturales sustituidos con aminas aromáticas usando un conector bifuncional que contiene aldehído. En una realización son procedimientos de unión de un conector sustituido con aldehído a una proteína sustituida con amina aromática mediante una reacción de alquilación reductora para generar un enlace de amina. El aminoácido no natural sustituido con amina aromática es un aminoácido no natural sustituido con arilamina. En otras realizaciones o realizaciones adicionales, los polipéptidos de aminoácidos no naturales se derivatizan específicamente en el sitio y/o con control preciso de la estructura tridimensional, usando un conector bifuncional que contiene aldehído. En una realización, tales procedimientos se usan para unir conectores moleculares (mono- bi- y multi-funcionales) a polipéptidos de aminoácidos no naturales que contienen aminas aromáticas, en los que al menos uno de los extremos del conector contiene un grupo aldehído que pueden ligarse con los polipéptidos de aminoácidos no naturales que contienen aminas aromáticas mediante un enlace de amina. Estos conectores se usan para conectar los polipéptidos de aminoácidos no naturales que contienen aminas aromáticas a otras moléculas, que incluyen, a modo de ejemplo, proteínas, otros polímeros (ramificados y no ramificados).

En algunas realizaciones, el polipéptido de aminoácidos no naturales de la invención está ligado a un polímero soluble en agua. En algunas realizaciones, el polímero soluble en agua comprende un resto de polietilenglicol. En algunas realizaciones, la molécula de polietilenglicol es un polímero bifuncional. En algunas realizaciones, el polímero bifuncional está ligado a un segundo polipéptido. En algunas realizaciones, el segundo polipéptido es idéntico al primer polipéptido; en otras realizaciones, el segundo polipéptido es un polipéptido diferente. En algunas realizaciones, el polipéptido de aminoácidos no naturales comprende al menos dos aminoácidos ligados a un polímero soluble en agua que comprende un resto de polietilenglicol.

El polipéptido de aminoácidos no naturales de la invención puede comprender una sustitución, adición o deleción que aumenta la afinidad del polipéptido de aminoácidos no naturales por un receptor. En algunas realizaciones, el polipéptido de aminoácidos no naturales comprende una sustitución, adición o deleción que aumenta la estabilidad del polipéptido de aminoácidos no naturales. En algunas realizaciones, el polipéptido de aminoácidos no naturales comprende una sustitución, adición o deleción que aumenta la solubilidad acuosa del polipéptido de aminoácidos no naturales. En algunas realizaciones, el polipéptido de aminoácidos no naturales comprende una sustitución, adición o deleción que aumenta la solubilidad del polipéptido de aminoácidos no naturales producida en una célula huésped. En algunas realizaciones, el polipéptido de aminoácidos no naturales comprende una sustitución, adición o deleción que modula la resistencia a proteasas, semivida en suero, inmunogenicidad y/o expresión con respecto al polipéptido de aminoácido sin la sustitución, adición o deleción.

Los polipéptidos de aminoácidos no naturales de la invención pueden ser un agonista, agonista parcial, antagonista, antagonista parcial o agonista inverso. En algunas realizaciones, el agonista, agonista parcial, antagonista, antagonista parcial o agonista inverso comprende un aminoácido no natural ligado a un polímero soluble en agua. En algunas realizaciones, el polímero acuoso comprende un resto de polietilenglicol. En algunas realizaciones, el polipéptido que comprende un aminoácido no natural ligado a un polímero soluble en agua, por ejemplo, puede prevenir la dimerización del receptor correspondiente. En algunas realizaciones, el polipéptido que comprende un aminoácido no natural ligado a un polímero soluble en agua modula la unión del polipéptido a un componente de unión, ligando o receptor. En algunas realizaciones, el polipéptido que comprende un aminoácido no natural ligado a un polímero soluble en agua modula una o más propiedades o actividades del polipéptido.

En algunas realizaciones, el codón selector está seleccionado del grupo que consiste en un codón ámbar, codón ocre, codón ópalo, un codón único, un codón raro, un codón no natural, un codón de cinco bases y un codón de cuatro bases.

En otras realizaciones o realizaciones adicionales, los polipéptidos de aminoácidos de aminas aromáticas no naturales y/o polipéptidos de aminoácidos de aminas aromáticas no naturales modificados descritos en el presente documento tienen al menos una de las siguientes características: (1) el resto de amina de la amina aromática es una amina primaria; (2) el resto de amina de la amina aromática es una amina secundaria; (3) el resto aromático de la amina aromática es un resto heteroaromático; (4) el resto aromático de la amina aromática es un resto arilo; (5) la amina aromática reacciona con grupos funcionalizados con aldehído; (6) se acopla a un polímero soluble en agua; (7) está PEGilada; (8) tiene elevada semivida terapéutica con respecto al polipéptido correspondiente sin el

aminoácido de amina aromática no natural; (9) tiene elevada semivida en suero con respecto al polipéptido correspondiente sin el aminoácido de amina aromática no natural; (10) tiene elevado tiempo en circulación con respecto al polipéptido correspondiente sin el aminoácido de amina aromática no natural; (11) tiene elevada solubilidad en agua con respecto al polipéptido correspondiente sin el aminoácido de amina aromática no natural; (12) tiene biodisponibilidad potenciada con respecto al polipéptido correspondiente sin el aminoácido de amina aromática no natural; (13) tiene inmunogenicidad modulada con respecto al polipéptido correspondiente sin el aminoácido de amina aromática no natural; (14) tiene actividad biológica modulada con respecto al polipéptido correspondiente sin el aminoácido de amina aromática no natural; (15) es parte de una composición farmacéutica; (16) se obtiene de cultivo celular; (17) se sintetiza químicamente; (18) se usa en procedimientos de cribado de bibliotecas; (19) se usa con matrices; (20) se usa con matrices de proteína; (21) se usa para el análisis de expresión génica; (22) se acopla a al menos un agente; (23) se acopla a una marca; (24) se acopla a un colorante; (25) se acopla a un polímero; (26) se acopla a un compuesto citotóxico; (27) se acopla a un fármaco; (28) se acopla a una segunda proteína o polipéptido o análogo de polipéptido; (29) se acopla a un anticuerpo o fragmento de anticuerpo; (30) se acopla a un hidrato de carbono; (31) se acopla a un polinucleótido; (32) se acopla a un polinucleótido antisentido; (33) se acopla a un sacárido, (34) se acopla a un fluoróforo, (35) se acopla a un grupo químicamente escindible; (36) se acopla a un grupo fotoescindible; (37) se acopla a un agente de transferencia de energía; (38) se acopla a un radionucleótido; (39) puede ser postraduccionalmente modificado; (40) puede ser postraduccionalmente modificado por alquilación reductora; (41) puede ser postraduccionalmente modificado por alquilación reductora en un intervalo de pH entre aproximadamente 4 y aproximadamente 10; (42) puede derivatizarse específicamente en el sitio por alquilación reductora postraduccional; (43) puede ser rápidamente postraduccionalmente modificado por alquilación reductora a temperatura ambiente, (44) puede ser postraduccionalmente modificado por alquilación reductora en condiciones acuosas; (45) puede ser postraduccionalmente modificado por alquilación reductora con condiciones de reacción estequiométricas; (46) puede ser postraduccionalmente modificado por alquilación reductora con condiciones de reacción casi estequiométricas; (47) puede ser postraduccionalmente modificado por alquilación reductora con condiciones de reacción similares a estequiométricas; (48) se usa para tratar un mamífero que padece una enfermedad, trastorno o afección; (49) se usa para tratar un ser humano que padece una enfermedad, trastorno o afección; (50) se usa para diagnosticar un mamífero que padece una enfermedad, trastorno o afección; (51) se usa para diagnosticar un ser humano que padece una enfermedad, trastorno o afección; (52) es parte de composiciones de liberación sostenida; (53) el resto de amina se forma por reducción postraduccional de un resto de amina enmascarado; (54) el resto de amina se forma por reducción postraduccional de un resto imina; (55) el resto de amina se forma por reducción postraduccional de un resto azida; (56) el resto de amina se forma por reducción postraduccional de un resto hidracina; (57) el resto de amina se forma por reducción postraduccional de un resto nitro, (58) se acopla a un pro-fármaco; (59) se obtiene de lisado celular; (60) se traduce ribosómicamente; (61) puede ser postraduccionalmente modificado por alquilación reductora en un intervalo de pH entre aproximadamente 4 y aproximadamente 7; (62) puede ser postraduccionalmente modificado por alquilación reductora en un intervalo de pH entre aproximadamente 4 y aproximadamente 5; (63) puede ser postraduccionalmente modificado por alquilación reductora a un pH de aproximadamente 5; (64) puede ser postraduccionalmente modificado por alquilación reductora a un pH de aproximadamente 4; (65) reacciona rápidamente en reacciones de alquilación reductora; (66) reacciona en menos de aproximadamente 10 horas en reacciones de alquilación reductora; (67) reacciona en menos de aproximadamente 8 horas en reacciones de alquilación reductora; (68) reacciona en menos de aproximadamente 6 horas en reacciones de alquilación reductora; (68) reacciona en menos de aproximadamente 4 horas en reacciones de alquilación reductora; (69) reacciona en menos de aproximadamente 2 horas en reacciones de alquilación reductora; (70) reacciona en menos de aproximadamente 1 hora en reacciones de alquilación reductora, o (71) reacciona en menos de aproximadamente 30 minutos en reacciones de alquilación reductora.

Los polipéptidos de aminoácidos de amina aromática no naturales y/o polipéptidos de aminoácidos de amina aromática no naturales modificados descritos en el presente documento pueden tener al menos dos de las características anteriormente mencionadas. Los polipéptidos de aminoácidos de amina aromática no naturales y/o polipéptidos de aminoácidos de amina aromática no naturales modificados descritos en el presente documento pueden tener al menos tres de las características anteriormente mencionadas. Los polipéptidos de aminoácidos de amina aromática no naturales y/o polipéptidos de aminoácidos de amina aromática no naturales modificados descritos en el presente documento pueden tener al menos cuatro de las características anteriormente mencionadas. Los polipéptidos de aminoácidos de amina aromática no naturales y/o polipéptidos de aminoácidos de amina aromática no naturales modificados descritos en el presente documento pueden tener al menos cinco de las características anteriormente mencionadas.

Los polipéptidos de aminoácidos basados en aldehído no naturales y/o los polipéptidos de aminoácidos basados en aldehído no naturales modificados descritos en el presente documento pueden tener al menos una de las siguientes características: (1) contiene un resto de aldehído protegido o enmascarado; (2) contiene un resto de aldehído desprotegido o desenmascarado; (3) contiene un resto de aldehído desprotegido o desenmascarado que puede reaccionar con una amina aromática; (4) contiene un resto de aldehído desprotegido o desenmascarado que puede reaccionar con una amina heteroaromática; (5) contiene un resto de aldehído desprotegido o desenmascarado que puede reaccionar con una amina aromática o una amina heteroaromática por aminación reductora; (6) se acopla a un polímero soluble en agua; (7) está PEGilado; (8) tiene elevada semivida terapéutica con respecto al polipéptido correspondiente sin el aminoácido basado en aldehído no natural; (9) tiene elevada semivida en suero con respecto

al polipéptido correspondiente sin el aminoácido basado en aldehído no natural; (10) tiene elevado tiempo en circulación con respecto al polipéptido correspondiente sin el aminoácido basado en aldehído no natural; (11) tiene elevada solubilidad en agua con respecto al polipéptido correspondiente sin el aminoácido basado en aldehído no natural; (12) tiene biodisponibilidad potenciada con respecto al polipéptido correspondiente sin el aminoácido basado en aldehído no natural; (13) tiene inmunogenicidad modulada con respecto al polipéptido correspondiente sin el aminoácido basado en aldehído no natural; (14) tiene actividad biológica modulada con respecto al polipéptido correspondiente sin el aminoácido basado en aldehído no natural; (15) es parte de una composición farmacéutica; (16) se obtiene de cultivo celular; (17) se sintetiza químicamente; (18) se usa en procedimientos de cribado de bibliotecas; (19) se usa con matrices; (20) se usa con matrices de proteína; (21) se usa para el análisis de expresión génica; (22) puede acoplarse a un agente mediante aminación reductora; (23) se acopla a una marca; (24) se acopla a un colorante; (25) se acopla a un polímero; (26) se acopla a un compuesto citotóxico; (27) se acopla a un fármaco; (28) se acopla a una segunda proteína o polipéptido o análogo de polipéptido; (29) se acopla a un anticuerpo o fragmento de anticuerpo; (30) se acopla a un hidrato de carbono; (31) se acopla a un polinucleótido; (32) se acopla a un polinucleótido antisentido; (33) se acopla a un sacárido, (34) se acopla a un fluoróforo, (35) se acopla a un grupo químicamente escindible; (36) se acopla a un grupo fotoescindible; (37) se acopla a un agente de transferencia de energía; (38) se acopla a un radionucleótido; (39) puede ser postraduccionalmente modificado; (40) puede ser postraduccionalmente modificado por aminación reductora; (41) puede ser postraduccionalmente modificado por aminación reductora en un intervalo de pH entre aproximadamente 4 y aproximadamente 10; (42) puede derivatizarse específicamente en el sitio por aminación reductora post-traducción; (43) puede ser rápidamente postraduccionalmente modificado por aminación reductora a temperatura ambiente, (44) puede ser postraduccionalmente modificado por aminación reductora en condiciones acuosas; (45) puede ser postraduccionalmente modificado por aminación reductora con condiciones de reacción estequiométricas; (46) puede ser postraduccionalmente modificado por aminación reductora con condiciones de reacción casi estequiométricas; (47) puede ser postraduccionalmente modificado por aminación reductora con condiciones de reacción similares a estequiométricas; (48) se usa para tratar un mamífero que padece una enfermedad, trastorno o afección; (49) se usa para tratar un ser humano que padece una enfermedad, trastorno o afección; (50) se usa para diagnosticar un mamífero que padece una enfermedad, trastorno o afección; (51) se usa para diagnosticar un ser humano que padece una enfermedad, trastorno o afección; (52) es parte de composiciones de liberación sostenida; (53) se acopla a un pro-fármaco; (54) se obtiene de lisado celular; (55) se traduce ribosómicamente; (56) puede ser postraduccionalmente modificado por aminación reductora en un intervalo de pH entre aproximadamente 4 y aproximadamente 7; (57) puede ser postraduccionalmente modificado por aminación reductora en un intervalo de pH entre aproximadamente 4 y aproximadamente 5; (58) puede ser postraduccionalmente modificado por aminación reductora a un pH de aproximadamente 5; (59) puede ser postraduccionalmente modificado por aminación reductora a un pH de aproximadamente 4; (60) reacciona rápidamente en reacciones de aminación reductora; (61) reacciona en menos de aproximadamente 10 horas en reacciones de aminación reductora; (62) reacciona en menos de aproximadamente 8 horas en reacciones de aminación reductora; (63) reacciona en menos de aproximadamente 6 horas en reacciones de aminación reductora; (68) reacciona en menos de aproximadamente 4 horas en reacciones de aminación reductora; (64) reacciona en menos de aproximadamente 2 horas en reacciones de aminación reductora; (65) reacciona en menos de aproximadamente 1 hora en reacciones de aminación reductora, o (66) reacciona en menos de aproximadamente 30 minutos en reacciones de aminación reductora.

Los polipéptidos de aminoácidos basados en aldehído no naturales y/o polipéptidos de aminoácidos basados en aldehído no naturales modificados descritos en el presente documento pueden tener al menos dos de las características anteriormente mencionadas. Los polipéptidos de aminoácidos basados en aldehído no naturales y los polipéptidos de aminoácidos basados en aldehído no naturales modificados descritos en el presente documento pueden tener al menos tres de las características anteriormente mencionadas. Los polipéptidos de aminoácidos basados en aldehído no naturales y/o polipéptidos de aminoácidos basados en aldehído no naturales modificados descritos en el presente documento pueden tener al menos cuatro de las características anteriormente mencionadas. Los polipéptidos de aminoácidos basados en aldehído no naturales y/o polipéptidos de aminoácidos basados en aldehído no naturales modificados descritos en el presente documento pueden tener al menos cinco de las características anteriormente mencionadas.

El procedimiento de la invención es útil para la preparación de un polipéptido de aminoácidos no naturales ligado a un polímero soluble en agua. En algunas realizaciones, el procedimiento comprende poner en contacto un polipéptido aislado que comprende un aminoácido no natural con un polímero soluble en agua que comprende un resto que reacciona con el aminoácido no natural. En algunas realizaciones, el aminoácido no natural incorporado es reactivo hacia un polímero soluble en agua que es de otro modo no reactivo hacia cualquiera de los 20 aminoácidos comunes. En algunas realizaciones, el polímero acuoso comprende un resto de polietilenglicol. El peso molecular del polímero puede estar de un amplio intervalo, que incluye, pero no se limita a, entre aproximadamente 100 Da y aproximadamente 100.000 Da o más. El peso molecular del polímero puede estar entre aproximadamente 100 Da y aproximadamente 100.000 Da, que incluye, pero no se limita a, 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 y 100 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 100 Da y aproximadamente 50.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 100 Da y

aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 1.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 5.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 10.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, la molécula de polietilenglicol es un polímero ramificado. El peso molecular del PEG de cadena ramificada puede estar entre aproximadamente 1.000 Da y aproximadamente 100.000 Da, que incluye, pero no se limita a, 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 y 1.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 1.000 Da y aproximadamente 50.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 1.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 5.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 5.000 Da y aproximadamente 20.000 Da.

En el presente documento también se describen composiciones que comprenden un polipéptido que comprende al menos uno de los aminoácidos no naturales descritos en el presente documento y un vehículo farmacéuticamente aceptable. El aminoácido no natural puede ligarse a un polímero soluble en agua. En el presente documento también se describen composiciones farmacéuticas que comprenden un vehículo farmacéuticamente aceptable y un polipéptido, en el que al menos un aminoácido está sustituido con un aminoácido no natural. El aminoácido no natural puede comprender un resto de sacárido. El polímero soluble en agua puede ligarse al polipéptido mediante un resto de sacárido.

En el presente documento también se describen células que comprenden un polinucleótido que codifica el polipéptido que comprende un codón selector. Las células pueden comprender una ARN sintetasa ortogonal y/o un ARNt ortogonal para sustituir un aminoácido no natural en el polipéptido. Las células pueden ser un cultivo celular, o parte de un organismo multicelular, que incluye anfibios, reptiles, aves y mamíferos. El polipéptido de aminoácidos no naturales puede producirse por expresión del polinucleótido. El polipéptido de aminoácidos no naturales puede producirse *in vitro.* El polipéptido de aminoácidos no naturales puede producirse en lisado celular. El polipéptido de aminoácidos no naturales puede producirse por traducción ribosómica.

En el presente documento también se describen procedimientos de preparación de un polipéptido que comprende un aminoácido no natural. Los procedimientos comprenden cultivar células que comprenden un polinucleótido o polinucleótidos que codifican un polipéptido, una ARN sintetasa ortogonal y/o un ARNt ortogonal en condiciones para permitir la expresión del polipéptido; y purificar el polipéptido de las células y/o medio de cultivo.

En el presente documento también se describen bibliotecas de los aminoácidos no naturales descritos en el presente documento o bibliotecas de los polipéptidos de aminoácidos no naturales descritos en el presente documento, o bibliotecas de los polipéptidos de aminoácidos no naturales modificados descritos en el presente documento, o bibliotecas de combinación de los mismos. En el presente documento también se describen matrices que contienen al menos un aminoácido no natural, al menos un polipéptido de aminoácidos no naturales, y/o al menos un aminoácido no natural modificado. En el presente documento también se describen matrices que contienen al menos un polinucleótido que codifica un polipéptido que comprende un codón selector. Las matrices descritas en el presente documento pueden usarse para cribar para la producción de los polipéptidos de aminoácidos no naturales en un organismo (tanto detectando la transcripción del polinucleótido que codifica el polipéptido como detectando la traducción del polipéptido).

En el presente documento también se describen procedimientos de cribado de bibliotecas descritas en el presente documento para una actividad deseada, o para usar las matrices descritas en el presente documento para cribar las bibliotecas descritas en el presente documento, o para otras bibliotecas de compuestos y/o polipéptidos y/o polinucleótidos para una actividad deseada. En el presente documento también se describe el uso de tales datos de actividad del cribado de bibliotecas para desarrollar y descubrir nuevos agentes terapéuticos, además de los propios agentes terapéuticos.

En el presente documento también se describen procedimientos de aumento de la semivida terapéutica, semivida en suero o tiempo en circulación de un polipéptido. En algunas realizaciones, los procedimientos comprenden sustituir al menos un aminoácido no natural con uno cualquiera o más aminoácidos en un polipéptido que se producen naturalmente y/o enlazar el polipéptido con un polímero soluble en agua.

En el presente documento también se describen procedimientos de tratamiento de un paciente en necesidad de tal tratamiento, con una cantidad eficaz de una composición farmacéutica que comprende un polipéptido que comprende un aminoácido no natural y un vehículo farmacéuticamente aceptable. En algunas realizaciones, el aminoácido no natural está ligado a un polímero soluble en agua.

Debe entenderse que los procedimientos y composiciones descritos en el presente documento no se limitan a la metodología particular, protocolos, líneas celulares, construcciones y reactivos descritos en el presente documento y

como tales pueden variar. También debe entenderse que la terminología usada en el presente documento es con el fin de describir realizaciones particulares solo, y no pretende limitar el alcance de los procedimientos y composiciones descritos en el presente documento, que se limitarán solo por las reivindicaciones adjuntas.

Como se usa en el presente documento y en las reivindicaciones adjuntas, las formas en singular "un", "una", "el" y "la" incluyen referencia en plural, a menos que el contexto indique claramente de otro modo.

A menos que se defina de otro modo, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que comúnmente es entendido por un experto habitual en la materia a la que pertenecen las invenciones descritas en el presente documento. Aunque cualquier procedimiento, dispositivo y material similar o equivalente a aquellos descritos en el presente documento puede usarse en la práctica o prueba de las invenciones descritas en el presente documento, ahora se describen los procedimientos, dispositivos y materiales preferidos.

Las publicaciones tratadas en el presente documento se proporcionan únicamente para su divulgación antes de la fecha de presentación de la presente solicitud. Nada en el presente documento debe interpretarse como una admisión de que los inventores descritos en el presente documento no tienen derecho a anteceder tal divulgación en virtud de la invención anterior o por cualquier otro motivo.

El término "marca de afinidad", como se usa en el presente documento, se refiere a una marca que se une reversiblemente o irreversiblemente a otra molécula, tanto para modificarla, destruirla, como formar un compuesto con ella. A modo de ejemplo, marcas de afinidad incluyen enzimas y sus sustratos, o anticuerpos y sus antígenos.

Los términos "alcoxi", "alquilamino" y "alquiltio" se usan en su sentido convencional, y se refieren a grupos alquilo ligados a moléculas mediante un átomo de oxígeno, un grupo amino, un átomo de azufre, respectivamente.

El término "alquilo", por sí mismo o como parte de otra molécula, significa, a menos que se establezca de otro modo, un radical de hidrocarburo de cadena lineal o ramificada, o cíclico, o combinación de los mismos, que puede estar completamente saturado, mono- o poliinsaturado y puede incluir radicales di- y multivalentes, que tiene el número de átomos de carbono designado (es decir, C_1-C_{10} significa uno a diez carbonos). Ejemplos de radicales de hidrocarburo saturados incluyen, pero no se limitan a, grupos tales como metilo, etilo, n-propilo, isopropilo, n-butilo, t-butilo, isobutilo, sec-butilo, ciclohexilo, (ciclohexil)metilo, ciclopropilmetilo, homólogos e isómeros de, por ejemplo, n-pentilo, n-hexilo, n-heptilo, n-octilo, y similares. Un grupo alquilo insaturado es uno que tiene uno o más dobles enlaces o triples enlaces. Ejemplos de grupos alquilo insaturado incluyen, pero no se limitan a, vinilo, 2-propenilo, crotilo, 2-isopentenilo, 2-(butadienilo), 2,4-pentadienilo, 3-(1,4-pentadienilo), etinilo, 1- y 3-propinilo, 3-butinilo, y los homólogos superiores e isómeros. El término "alquilo", a menos que se indique lo contrario, también pretende incluir aquellos derivados de alquilo definidos más abajo en más detalle, tales como "heteroalquilo", "haloalquilo" y "homoalquilo".

El término "alquileno", por sí mismo o como parte de otra molécula, significa un radical divalente derivado de un alcano, como se ejemplifica por $(-CH_2-)_n$, en el que n puede ser 1 a aproximadamente 24. A modo de ejemplo solo, tales grupos incluyen, pero no se limitan a, grupos que tienen 10 o menos átomos de carbono tales como las estructuras $-CH_2CH_2-$ y $-CH_2CH_2CH_2CH_2-$. Un "alquilo inferior" o "alquileno inferior" es un grupo alquilo o alquileno de cadena más corta, generalmente que tiene ocho o menos átomos de carbono. El término "alquileno", a menos que se indique lo contrario, también pretende incluir aquellos grupos descritos más adelante como "heteroalquileno".

El término "aminoácido" se refiere a aminoácidos que se producen naturalmente y no naturales, además de a análogos de aminoácidos y miméticos de aminoácidos que funcionan de un modo similar a los aminoácidos que se producen naturalmente. Los aminoácidos codificados naturalmente son los 20 aminoácidos comunes (alanina, arginina, asparagina, ácido aspártico, cisteína, glutamina, ácido glutámico, glicina, histidina, isoleucina, leucina, lisina, metionina, fenilalanina, prolina, serina, treonina, triptófano, tirosina y valina) y pirrolisina y selenocisteína. Análogos de aminoácido se refiere a compuestos que tienen la misma estructura química básica que un aminoácido que se produce naturalmente, a modo de ejemplo, un carbono α que está unido a un hidrógeno, un grupo carboxilo, un grupo amino y un grupo R. Tales análogos pueden tener grupos R modificados (a modo de ejemplo, norleucina) o pueden tener esqueletos de péptido modificados, aunque todavía retienen la misma estructura química básica que un aminoácido que se produce naturalmente. Ejemplos no limitantes de análogos de aminoácidos incluyen homoserina, norleucina, sulfóxido de metionina, metionina metilsulfonio.

Los aminoácidos pueden denominarse en el presente documento por tanto su nombre, sus símbolos de tres letras comúnmente conocidos, como por los símbolos de una letra recomendados por la IUPAC-IUB Biochemical Nomenclature Commission. Adicionalmente, los nucleótidos pueden denominarse por sus códigos de una sola letra comúnmente aceptados.

Un "grupo de modificación del extremo amino" se refiere a cualquier molécula que puede unirse a un grupo de amina terminal. A modo de ejemplo, tales grupos de amina terminal pueden estar en el extremo de moléculas poliméricas, en los que tales moléculas poliméricas incluyen, pero no se limitan a, polipéptidos, polinucleótidos y polisacáridos. Los grupos de modificación del extremo incluyen, pero no se limitan a, diversos polímeros solubles en agua,

péptidos o proteínas. A modo de ejemplo solo, los grupos de modificación del extremo incluyen polietilenglicol o albúmina de suero. Los grupos de modificación del extremo pueden usarse para modificar características terapéuticas de la molécula polimérica, que incluyen, pero no se limita a, aumentar la semivida en suero de péptidos.

Por "fragmento de anticuerpo" se indica cualquier forma de un anticuerpo distinta de la forma de longitud completa. Fragmentos de anticuerpos en el presente documento incluyen anticuerpos que son componentes más pequeños que existen dentro de los anticuerpos de longitud completa, y anticuerpos que han sido manipulados. Los fragmentos de anticuerpos incluyen, pero no se limitan a, Fv, Fc, Fab y $(Fab')_2$, Fv monocatenario (scFv), diacuerpos, triacuerpos, tetracuerpos, anticuerpos híbridos bifuncionales, CDR1, CDR2, CDR3, combinaciones de CDR, regiones variables, regiones estructurales, regiones constantes, cadenas pesadas, cadenas ligeras y regiones variables, y moléculas de no anticuerpo de andamiaje alternativo, anticuerpos biespecíficos y similares (Maynard & Georgiou, 2000, Annu. Rev. Biomed. Eng. 2:339-76; Hudson, 1998, Curr. Opin. Biotechnol. 9:395-402). Otra subestructura funcional es un Fv monocatenario (scFv), que comprende las regiones variables de las cadenas pesadas y ligeras de inmunoglobulina, covalentemente conectadas por un conector peptídico (S-z Hu y col., 1996, Cancer Research, 56, 3055-3061). Estas pequeñas proteínas (Mr 25.000) generalmente retienen la especificidad y afinidad por antígeno en un único polipéptido y pueden proporcionar un bloque constructivo conveniente para moléculas mayores específicas de antígeno. A menos que se indique específicamente de otro modo, las afirmaciones y reivindicaciones que usan el término "anticuerpo" o "anticuerpos" incluyen específicamente "fragmento de anticuerpo" y "fragmentos de anticuerpos".

El término "amina aromática", como se usa en el presente documento, se refiere a un resto arilo que contiene un resto amino. Tales restos amino pueden incluir, pero no se limitan a, aminas primarias, aminas secundarias, aminas terciarias, aminas enmascaradas o aminas protegidas. Tales aminas terciarias, aminas enmascaradas o aminas protegidas pueden convertirse en restos de amina primaria o amina secundaria. Adicionalmente, el resto de amina puede incluir un resto similar a amina que tiene características químicas similares a los restos de amina, que incluyen reactividad química.

El término "aromático" o "arilo", como se usan en el presente documento, se refiere a una estructura de anillo cerrada que tiene al menos un anillo que tiene un sistema de electrones pi conjugado e incluye tanto grupos arilo carbocíclico como arilo heterocíclico (o "heteroarilo" o "heteroaromático"). El grupo aromático carbocíclico o heterocíclico puede contener de 5 a 20 átomos de anillo. El término incluye anillos monocíclicos ligados covalentemente o grupos policíclicos de anillos condensados (es decir, anillos que comparten pares adyacentes de átomos de carbono). Un grupo aromático puede estar sin sustituir o sustituido. Ejemplos de grupos "aromáticos" o "arilo" incluyen fenilo, 1-naftilo, 2-naftilo, 4-bifenilo, antracenilo y fenantracenilo. Sustituyentes para cada uno de los sistemas de arilo y heteroarilo anteriormente indicados están seleccionados del grupo de sustituyentes aceptables descrito más adelante.

Por brevedad, el término "aromático" o "arilo", cuando se usa en combinación con otros términos (que incluyen, pero no se limitan a, ariloxi, ariltioxi, aralquilo), incluye tanto anillos de arilo como de heteroarilo como se han definido anteriormente. Así, el término "aralquilo" o "alcarilo" pretende incluir aquellos radicales en los que un grupo arilo está unido a un grupo alquilo (que incluyen, pero no se limitan a, bencilo, fenetilo, piridilmetilo y similares) que incluye aquellos grupos alquilo en los que un átomo de carbono (que incluye, pero no se limita a, un grupo metileno) se ha sustituido con un heteroátomo, a modo de ejemplo solo, con un átomo de oxígeno. Ejemplos de tales grupos arilo incluyen fenoximetilo, 2-piridiloximetilo, 3-(1-naftiloxi)propilo y similares.

Un "polímero bifuncional", también denominado un "conector bifuncional", se refiere a un polímero que comprende dos grupos funcionales que pueden reaccionar específicamente con otros restos para formar enlaces covalentes o no covalentes. Tales restos pueden incluir los grupos laterales sobre aminoácidos o péptidos naturales o no naturales que contienen tales aminoácidos naturales o no naturales. A modo de ejemplo solo, un conector bifuncional puede tener un grupo funcional reactivo con un grupo sobre un primer péptido y otro grupo funcional que es reactivo con un grupo sobre un segundo péptido, formando así un conjugado que incluye el primer péptido, el conector bifuncional y el segundo péptido. Se conocen muchos procedimientos y moléculas de conector para la unión de diversos compuestos a péptidos. Véase, por ejemplo, la solicitud de patente europea nº 188.256; las patentes de EE.UU. nº 4.671.958, 4.659.839, 4.414.148, 4.699.784; 4.680.338; y 4.569.789. Un "polímero multi-funcional" también denominado un "conector multi-funcional" se refiere a un polímero que comprende dos o más grupos funcionales que pueden reaccionar con otros restos. Tales restos pueden incluir los grupos laterales sobre aminoácidos o péptidos naturales o no naturales que contienen tales aminoácidos naturales o no naturales (que incluyen, pero no se limitan a, grupos laterales de aminoácidos) para formar enlaces covalentes o no covalentes. Un polímero bifuncional o polímero multi-funcional puede tener cualquier longitud o peso molecular deseado, y puede seleccionarse para proporcionar una separación o conformación deseada particular entre una o más moléculas ligadas a un compuesto y moléculas a las que se une o el compuesto.

El término "biodisponibilidad", como se usa en el presente documento, se refiere a la tasa y grado al que una sustancia o su resto activo se administra de una forma farmacéutica y está disponible en el sitio de acción o en la circulación general.

El término “molécula biológicamente activa”, “resto biológicamente activo” o “agente biológicamente activo”, cuando se usa en el presente documento, significa cualquier sustancia que puede afectar cualquier propiedad física o bioquímica de un sistema biológico, ruta, molécula, o interacción referente a un organismo, que incluye virus, bacterias, bacteriófago, transposón, prión; insectos, hongos, plantas, animales y seres humanos. En particular, como se usa en el presente documento, moléculas biológicamente activas incluyen, pero no se limitan a, cualquier sustancia prevista para el diagnóstico, cura, mitigación, tratamiento o prevención de enfermedad en seres humanos u otros animales, o para potenciar de otro modo el bienestar físico o mental de seres humanos o animales. Ejemplos de moléculas biológicamente activas incluyen péptidos, proteínas, enzimas, fármacos de molécula pequeña, drogas duras, drogas blandas, hidratos de carbono, átomos inorgánicos o moléculas, colorantes, lípidos, nucleósidos, radionúclidos, oligonucleótidos, toxinas, células, virus, liposomas, micropartículas y micelas. Clases de agentes biológicamente activos que son adecuados para su uso con los procedimientos y composiciones descritos en el presente documento incluyen fármacos, profármacos, radionúclidos, agentes de obtención de imágenes, polímeros, antibióticos, fungicidas, agentes antivirales, agentes antiinflamatorios, agentes antitumorales, agentes cardiovasculares, agentes ansiolíticos, hormonas, factores de crecimiento, agentes esteroideos, toxinas microbianamente derivadas, y similares.

El término “biomaterial”, como se usa en el presente documento, se refiere a un material derivado biológicamente que incluye material obtenido de biorreactores y/o de procedimientos y técnicas recombinantes.

El término “sonda biofísica”, como se usa en el presente documento, se refiere a sondas que pueden detectar o monitorizar cambios estructurales en moléculas. Tales moléculas incluyen proteínas y la “sonda biofísica” puede usarse para detectar o monitorizar la interacción de proteínas con otras macromoléculas. Ejemplos de sondas biofísicas incluyen, pero no se limitan a, marcas de espín, fluoróforos y grupos fotoactivables.

El término “análogo de biotina”, o también denominado “mimético de biotina”, como se usa en el presente documento, es cualquier molécula, distinta de biotina, que se une con alta afinidad a avidina y/o estreptavidina.

El término “grupo de modificación del extremo carboxi” se refiere a cualquier molécula que puede unirse a un grupo carboxi terminal. A modo de ejemplo, tales grupos carboxi terminales pueden estar en el extremo de moléculas poliméricas, en los que tales moléculas poliméricas incluyen polipéptidos, polinucleótidos y polisacáridos. Los grupos de modificación del extremo incluyen diversos polímeros solubles en agua, péptidos o proteínas. A modo de ejemplo solo, los grupos de modificación del extremo incluyen polietilenglicol o albúmina de suero. Los grupos de modificación del extremo pueden usarse para modificar características terapéuticas de la molécula polimérica, que incluyen, pero no se limitan a, aumentar la semivida en suero de péptidos.

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

El término “grupo quimioluminiscente”, 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. A modo de ejemplo solo, el luminol (5-amino-2,3-dihidro-1,4-ftalazindiona) reacciona con oxidantes como 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).

El término “cromóforo”, 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”, como se usa en el presente documento, se refiere a un átomo o molécula esencial para la acción de una gran molécula. Los cofactores incluyen, pero no se limitan a, iones inorgánicos, coenzimas, proteínas, o algún otro factor necesario para la actividad de enzimas. Ejemplos incluyen hemo en hemoglobina, magnesio en clorofila e iones metálicos para proteínas.

“Coplegamiento”, como se usa en el presente documento, se refiere a procesos, reacciones o procedimientos de replegamiento que emplean al menos dos moléculas que interaccionan entre sí y producen la transformación de moléculas no plegadas o inapropiadamente plegadas en moléculas apropiadamente plegadas. A modo de ejemplo solo, el “coplegamiento” emplea al menos dos polipéptidos que interaccionan entre sí y producen la transformación de polipéptidos no plegados o inapropiadamente plegados en polipéptidos nativos apropiadamente plegados. Tales polipéptidos pueden contener aminoácidos naturales y/o al menos un aminoácido no natural.

Una “ventana de comparación”, como se usa en el presente documento, se refiere a un segmento de una cualquiera de las posiciones continuas usadas para comparar una secuencia con una secuencia de referencia del mismo número de posiciones continuas después de alinearse óptimamente las dos secuencias. Tales posiciones continuas incluyen un grupo que consiste en de aproximadamente 20 a aproximadamente 600 unidades secuenciales, que incluyen aproximadamente 50 a aproximadamente 200 unidades secuenciales, y aproximadamente 100 a aproximadamente 150 unidades secuenciales. A modo de ejemplo solo, tales secuencias incluyen polipéptidos y polipéptidos que contienen aminoácidos no naturales, incluyendo las unidades secuenciales aminoácidos naturales y

no naturales. Además, a modo de ejemplo solo, tales secuencias incluyen polinucleótidos con nucleótidos que son las unidades secuenciales correspondientes. Los procedimientos de alineamiento de secuencias para comparación son muy conocidos en la técnica. Puede realizarse alineamiento óptimo de secuencias para comparación, que incluye por el algoritmo de homología local de Smith y Waterman (1970) Adv. Appl. Math. 2:482c, por el algoritmo de alineamiento por homología de Needleman y Wunsch (1970) J. Mol. Biol. 48:443, por el procedimiento de búsqueda de similitud de Pearson y Lipman (1988) Proc. Nat'l. Acad. Sci. USA 85:2444, por implementaciones computerizadas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), o por alineamiento manual e inspección visual (véase, por ejemplo, Ausubel y col., Current Protocols in Molecular Biology (suplemento de 1995)).

A modo de ejemplo, un algoritmo que puede usarse para determinar el porcentaje de identidad de secuencias y similitud de secuencias son los algoritmos BLAST y BLAST 2.0, que se describen en Altschul y col. (1997) Nuc. Acids Res. 25:3389-3402, y Altschul y col. (1990) J. Mol. Biol. 215:403-410, respectivamente. El software para realizar el análisis de BLAST está públicamente disponible mediante el Centro Nacional para Información Biotecnológica. Los parámetros del algoritmo BLAST W, T y X determinan la sensibilidad y velocidad del alineamiento. El programa BLASTN (para secuencias de nucleótidos) usa por defecto una longitud de palabra (W) de 11, una expectativa (E) o 10, M=5, N=-4 y una comparación de ambas hebras. Para secuencias de aminoácidos, el programa BLASTP usa por defecto una longitud de palabra de 3 y expectativa (E) de 10, y la matriz de puntuación BLOSUM62 (véase Henikoff y Henikoff (1992) Proc. Natl. Acad. Sci. USA 89:10915) alineamientos (B) de 50, expectativa (E) de 10, M=5, N=-4, y una comparación de ambas hebras. El algoritmo BLAST normalmente se realiza con el filtro de "baja complejidad" apagado.

El algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin y Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873-5787). Una medida de la similitud proporcionada por el algoritmo BLAST es la probabilidad de suma más pequeña (P(N)), que proporciona una indicación de la probabilidad por la cual se producirá por casualidad una correspondencia entre dos secuencias de nucleótidos o de aminoácidos. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la probabilidad de suma más pequeña en una comparación del ácido nucleico de prueba con el ácido nucleico de referencia es menos de aproximadamente 0,2, más o menos de aproximadamente 0,01, o menos de aproximadamente 0,001.

El término "variantes conservativamente modificadas" se aplica a tanto secuencias de aminoácidos naturales y no naturales como a secuencias de ácidos nucleicos naturales y no naturales, y combinaciones de las mismas. Con respecto a secuencias de ácidos nucleicos particulares, "variantes conservativamente modificadas" se refiere a aquellos ácidos nucleicos naturales y no naturales que codifican secuencias de aminoácidos naturales y no naturales idénticas o esencialmente idénticas, o si el ácido nucleico natural y no natural no codifica una secuencia de aminoácidos naturales y no naturales, a secuencias esencialmente idénticas. A modo de ejemplo, debido a la degeneración del código genético, un gran número de ácidos nucleicos funcionalmente idénticos codifican cualquier proteína dada. Por ejemplo, los codones GCA, GCC, GCG y GCU codifican todos el aminoácido alanina. Así, en cualquier posición en la que se especifique una alanina por un codón, el codón puede cambiarse a cualquiera de los codones correspondientes descritos sin cambiar el polipéptido codificado. Tales variaciones de ácido nucleico son "variaciones silenciosas", que son una especie de variaciones conservativamente modificadas. Así, a modo de ejemplo, cada secuencia de ácidos nucleicos naturales o no naturales en el presente documento que codifica un polipéptido natural o no natural también describe cualquier variación silenciosa posible del ácido nucleico natural o no natural. Un experto habitual en la materia reconocerá que cada codón en un ácido nucleico natural o no natural (excepto AUG, que es generalmente el único codón para metionina, y TGG, que es generalmente el único codón para triptófano) puede modificarse para dar una molécula funcionalmente idéntica. Por consiguiente, cada variación silenciosa de un ácido nucleico natural y no natural que codifica un polipéptido natural y no natural está implícita en cada secuencia descrita.

En cuanto a las secuencias de aminoácidos, sustituciones individuales, deleciones o adiciones a un ácido nucleico, péptido, polipéptido, o secuencia de proteínas que altere, añada o delecione un único aminoácido natural y no natural o un porcentaje pequeño de aminoácidos naturales y no naturales en la secuencia codificada es una "variante conservativamente modificada" en la que la alteración produce la deleción de un aminoácido, adición de un aminoácido, o sustitución de un aminoácido natural y no natural con un aminoácido químicamente similar. Las tablas de sustituciones conservativas que proporcionan aminoácidos naturales funcionalmente similares son muy conocidas en la técnica. Tales variantes conservativamente modificadas son además de y no excluyen variantes polimórficas, homólogos entre especies y alelos de los procedimientos y composiciones descritos en el presente documento.

Las tablas de sustituciones conservativas que proporcionan aminoácidos funcionalmente similares son conocidas para aquellos expertos habituales en la materia. Los ocho siguientes grupos contienen cada uno aminoácidos que son sustituciones conservativas con otro:

1) Alanina (A), glicina (G);
2) Ácido aspártico (D), ácido glutámico (E);
3) Asparagina (N), glutamina (Q);

4) Arginina (R), lisina (K);
5) Isoleucina (I), leucina (L), metionina (M), valina (V);
6) Fenilalanina (F), tirosina (Y), triptófano (W);
7) Serina (S), treonina (T); y
8) Cisteína (C), metionina (M)

(véase, por ejemplo, Creighton, *Proteins: Structures and Molecular properties* (W H Freeman & Co.; 2ª edición (diciembre de 1993)

Los términos "cicloalquilo" y "heterocicloalquilo", por sí mismos o en combinación con otros términos, representan, a menos que se establezca de otro modo, versiones cíclicas de "alquilo" y "heteroalquilo", respectivamente. Así, un cicloalquilo o heterocicloalquilo incluye enlaces de anillo saturados, parcialmente insaturados y completamente insaturados. Adicionalmente, para heterocicloalquilo, un heteroátomo puede ocupar la posición en la que el heterociclo se une al resto de la molécula. El heteroátomo puede incluir oxígeno, nitrógeno o azufre. Ejemplos de cicloalquilo incluyen, pero no se limitan a, ciclopentilo, ciclohexilo, 1-ciclohexenilo, 3-ciclohexenilo, cicloheptilo y similares. Ejemplos de heterocicloalquilo incluyen 1-(1,2,5,6-tetrahidropiridilo), 1-piperidinilo, 2-piperidinilo, 3-piperidinilo, 4-morfolinilo, 3-morfolinilo, tetrahidrofuran-2-ilo, tetrahidrofuran-3-ilo, tetrahidrotien-2-ilo, tetrahidrotien-3-ilo, 1-piperazinilo, 2-piperazinilo y similares. Adicionalmente, el término engloba estructuras multicíclicas, que incluyen estructuras de anillo bicíclicas y tricíclicas. Similarmente, el término "heterocicloalquileno", por sí mismo o como parte de otra molécula, significa un radical divalente derivado de heterocicloalquilo, y el término "cicloalquileno", por sí mismo o como parte de otra molécula, significa un radical divalente derivado de cicloalquilo.

El término "ciclodextrina", 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 un cavidad relativamente no polar capaz de acomodar moléculas pequeñas.

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

"Agente desnaturalizante" o "desnaturalizante", como se usa en el presente documento, se refiere a cualquier compuesto o material que producirá un desplegamiento reversible de un polímero. A modo de ejemplo solo, el "agente desnaturalizante" o "desnaturalizantes" puede producir un desplegamiento reversible de una proteína. La fuerza de un agente desnaturalizante o desnaturalizante se determinará tanto por las propiedades como por la concentración del agente desnaturalizante o desnaturalizante particular. A modo de ejemplo, agentes desnaturalizantes o desnaturalizantes incluyen, pero no se limitan a, caótropos, detergentes, disolventes orgánicos miscibles en agua, fosfolípidos, o una combinación de los mismos. Ejemplos de caótropos incluyen urea, guanidina y tiocianato de sodio. Ejemplos de detergentes pueden incluir detergentes fuertes tales como dodecilsulfato de sodio o polioxietilenéteres (por ejemplo, detergentes Tween o 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 ión bipolar que incluyen, pero no se limitan a, sulfobetaínas (Zwittergent), sulfato de 3-(3-colamidopropil)dimetilamonio-1-propano (CHAPS) y sulfonato de 3-(3-colamidopropil)dimetilamonio-2-hidroxi-1-propano (CHAPSO). Ejemplos no limitantes de disolventes orgánicos miscibles en agua incluyen acetonitrilo, alcanoles inferiores (especialmente alcanoles C_2-C_4 tales como etanol o isopropanol), o pueden usarse alcanodioles inferiores (alcanodioles C_2-C_4 tales como etilenglicol) como desnaturalizantes. Ejemplos de fosfolípidos incluyen fosfolípidos que se producen naturalmente tales como fosfatidiletanolamina, fosfatidilcolina, fosfatidilserina y fosfatidilinositol, o derivados o variantes de fosfolípidos sintéticos tales como dihexanoilfosfatidilcolina o diheptanoilfosfatidilcolina.

El término "marca detectable", como se usa en el presente documento, se refiere a una marca que puede ser observable usando técnicas analíticas que incluyen fluorescencia, quimioluminiscencia, resonancia de espín electrónico, espectroscopía de absorbancia ultravioleta/visible, espectrometría de masas, resonancia magnética nuclear, resonancia magnética y procedimientos electroquímicos.

El término "dicarbonilo", como se usa en el presente documento, se refiere a un grupo que contiene al menos dos restos seleccionados del grupo que consiste en -C(O)-, -S(O)-, -$S(O)_2$- y -C(S)-, que incluyen grupos 1,2-dicarbonilo, grupos 1,3-dicarbonilo y grupos 1,4-dicarbonilo, y grupos que contienen al menos un grupo cetona, y/o al menos un grupos aldehído, y/o al menos un grupo éster, y/o al menos un grupo ácido carboxílico, y/o al menos un grupo tioéster. Tales grupos dicarbonilo incluyen dicetonas, cetoaldehídos, cetoácidos, cetoésteres y cetotioésteres. Además, tales grupos pueden ser parte de moléculas lineales, ramificadas o cíclicas.

El término "fármaco", como se usa en el presente documento, se refiere a cualquier sustancia usada en la prevención, diagnóstico, alivio, tratamiento o cura de una enfermedad o afección.

El término "colorante", como se usa en el presente documento, se refiere a una sustancia colorante soluble que contiene un cromóforo.

El término “cantidad eficaz”, como se usa en el presente documento, se refiere a una cantidad suficiente de un agente o un compuesto que se administra que aliviará de algún modo uno o más de los síntomas de la enfermedad o afección que está tratándose. El resultado puede ser la reducción y/o alivio de los signos, síntomas o causas de una enfermedad, o cualquier otra alteración deseada de un sistema biológico. A modo de ejemplo, un agente o un compuesto que se administra incluye, pero no se limita a, un polipéptido de aminoácidos naturales, polipéptido de aminoácidos no naturales, polipéptido de aminoácidos naturales modificado, o polipéptido de aminoácidos no naturales modificado. Composiciones que contienen tales polipéptidos de aminoácidos naturales, polipéptidos de aminoácidos no naturales, polipéptidos de aminoácidos naturales modificados, o polipéptidos de aminoácidos no naturales modificados pueden administrarse para tratamientos profilácticos, de potenciamiento y/o terapéuticos. Una cantidad “eficaz” apropiada en cualquier caso individual puede determinarse usando técnicas, tales como un estudio de aumento de dosis.

El término “grupo con densidad electrónica”, como se usa en el presente documento, se refiere a un grupo que dispersa electrones cuando se irradia con un haz de electrones. Tales grupos incluyen molibdato de amonio, subnitrato de bismuto, yoduro de cadmio, 99 %, carbohidrazida, cloruro férrico hexahidratado, hexametilentetramina, 98,5 %, tricloruro de indio anhidro, nitrato de lantano, acetato de plomo trihidratado, citrato de plomo trihidratado, nitrato de plomo, ácido peryódico, ácido fosfomolíbdico, ácido fosfotúngstico, ferricianuro de potasio, ferrocianuro de potasio, rojo de rutenio, nitrato de plata, proteinato de plata (ensayo de Ag: 8,0-8,5 %) “Fuerte”, tetrafenilporfirina 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.

El término “agente de transferencia de energía”, como se usa en el presente documento, se refiere a una molécula que puede tanto donar como aceptar energía de otra molécula. A modo de ejemplo solo, la transferencia de energía por resonancia de fluorescencia (FRET) es un procedimiento de acoplamiento dipolo-dipolo por el que la energía en estado excitado de una molécula donante de fluorescencia se transfiere no radiativamente a una molécula aceptora no excitada que entonces emite fluorescentemente la energía donada a una longitud de onda más larga.

Los términos “potenciar” o “potenciamiento” significan aumentar o prolongar tanto en potencia como en duración un efecto deseado. A modo de ejemplo, “potenciar” el efecto de agentes terapéuticos se refiere a la capacidad para aumentar o prolongar, tanto en potencia como en duración, el efecto de agentes terapéuticos durante el tratamiento de una enfermedad, trastorno o afección. Una “cantidad eficaz potenciadora”, 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 afección. Cuando se usa en un paciente, cantidades eficaces de este uso dependerán de la gravedad y evolución de la enfermedad, trastorno o afección, terapia previa, el estado de salud del paciente y respuesta a los fármacos, y el criterio del médico práctico.

Como se usa en el presente documento, el término “eucariota” se refiere a organismos que pertenecen al dominio filogenético Eucarya, que incluye, pero no se limita a, animales (que incluyen mamíferos, insectos, reptiles, aves, etc.), ciliados, plantas (que incluyen, pero no se limitan a, monocotiledóneas, dicotiledóneas y algas), hongos, levaduras, flagelados, microsporidios y protistas.

El término “ácido graso”, como se usa en el presente documento, se refiere a ácidos carboxílicos con cadena lateral de hidrocarburo de aproximadamente C_6 o más larga.

El término “fluoróforo”, como se usa en el presente documento, se refiere a una molécula que tras la excitación emite fotones y es así fluorescente.

Los términos “grupo funcional”, “resto activo”, “grupo activante”, “grupo saliente”, “sitio reactivo”, “grupo químicamente reactivo” y “resto químicamente reactivo”, como se usan en el presente documento, se refieren a porciones o unidades de una molécula en la que se producen reacciones químicas. Los términos son en cierto modo sinónimos en las artes químicas y se usan en el presente documento para indicar las porciones de moléculas que realizan alguna función o actividad y son reactivas con otras moléculas.

El término “halógeno” incluye flúor, cloro, yodo y bromo.

El término “haloacilo”, como se usa en el presente documento, se refiere a grupos acilo que contienen restos de halógeno, que incluyen $-C(O)CH_3$, $-C(O)CF_3$, $-C(O)CH_2OCH_3$ y similares.

El término “haloalquilo”, como se usa en el presente documento, se refiere a grupos alquilo que contienen restos de halógeno, que incluyen $-CF_3$ y $-CH_2CF_3$ y similares.

El término “heteroalquilo”, como se usa en el presente documento, se refiere a radicales de hidrocarburo de cadena lineal o ramificada, o cíclico, o combinaciones de los mismos, que consisten en un grupo alquilo y al menos un heteroátomo seleccionado del grupo que consiste en O, N, Si y S, y en el que los átomos de nitrógeno y de azufre pueden oxidarse opcionalmente y el heteroátomo de nitrógeno puede cuaternizarse opcionalmente. El (Los) heteroátomo(s) O, N y S y Si puede(n) colocarse en cualquier posición interior del grupo heteroalquilo o en la

posición en la que el grupo alquilo está unido al resto de la molécula. Ejemplos incluyen $-CH_2-CH_2-O-CH_3$, $-CH_2-CH_2-NH-CH_3$, $-CH_2-CH_2-N(CH_3)-CH_3$, $-CH_2-S-CH_2-CH_3$, $-CH_2-CH_2$, $-S(O)-CH_3$, $-CH_2-CH_2-S(O)_2-CH_3$, $-CH=CH-O-CH_3$, $-Si(CH_3)_3$, $-CH_2-CH=N-OCH_3$ y $-CH=CH-N(CH_3)-CH_3$. Además, hasta dos heteroátomos pueden ser consecutivos, tal como, a modo de ejemplo, $-CH_2-NH-OCH_3$ y $-CH_2-O-Si(CH_3)_3$.

El término "heteroalquileno", como se usa en el presente documento, se refiere a un radical divalente derivado de heteroalquilo, como se ejemplifica por $-CH_2-CH_2-S-CH_2-CH_2-$ y $-CH_2-S-CH_2-CH_2-NH-CH_2-$. Para grupos heteroalquileno, heteroátomos iguales o diferentes también pueden ocupar cualquiera o ambos de los extremos de cadena (incluyendo alquilenoxi, alquilendioxi, alquilenamino, alquilendiamino, aminooxialquileno y similares). Todavía además, para grupos de enlace de alquileno y heteroalquileno, ninguna orientación del grupo de enlace está implicada por la dirección en la que se escribe la fórmula del grupo de enlace. A modo de ejemplo, la fórmula $-C(O)_2R'-$ representa tanto $-C(O)_2R'-$ como $-R'C(O)_2-$.

El término "heteroarilo" o "heteroaromático", como se usa en el presente documento, se refiere a grupos arilo que contienen al menos un heteroátomo seleccionado de N, O y S; en el que los átomos de nitrógeno y de azufre pueden estar opcionalmente oxidados, y el (los) átomo(s) de nitrógeno puede(n) estar opcionalmente cuaternizado(s). Los grupos heteroarilo pueden estar sustituidos o sin sustituir. Un grupo heteroarilo puede unirse al resto de la molécula mediante un heteroátomo. Ejemplos de grupos heteroarilo incluyen 1-pirrolilo, 2-pirrolilo, 3-pirrolilo, 3-pirazolilo, 2-imidazolilo, 4-imidazolilo, pirazinilo, 2-oxazolilo, 4-oxazolilo, 2-fenil-4-oxazolilo, 5-oxazolilo, 3-isoxazolilo, 4-isoxazolilo, 5-isoxazolilo, 2-tiazolilo, 4-tiazolilo, 5-tiazolilo, 2-furilo, 3-furilo, 2-tienilo, 3-tienilo, 2-piridilo, 3-piridilo, 4-piridilo, 2-pirimidilo, 4-pirimidilo, 5-benzotiazolilo, purinilo, 2-bencimidazolilo, 5-indolilo, 1-isoquinolilo, 5-isoquinolilo, 2-quinoxalinilo, 5-quinoxalinilo, 3-quinolilo y 6-quinolilo.

El término "homoalquilo", como se usa en el presente documento, se refiere a grupos alquilo que son grupos de hidrocarburo.

El término "idéntico", como se usa en el presente documento, se refiere a dos o más secuencias o subsecuencias que son iguales. Además, el término "sustancialmente idéntico", 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 comparan y se alinean para correspondencia máxima con respecto a una ventana de comparación, o región designada, como se mide usando un algoritmo de comparación o por alineamiento manual e inspección visual. A modo de ejemplo solo, dos o más secuencias pueden ser "sustancialmente idénticas" si las unidades secuenciales son aproximadamente el 60 % idénticas, aproximadamente el 65 % idénticas, aproximadamente el 70 % idénticas, aproximadamente el 75 % idénticas, aproximadamente el 80 % idénticas, aproximadamente el 85 % idénticas, aproximadamente el 90 % idénticas, o aproximadamente el 95 % idénticas con respecto a una región especificada. Tales porcentajes son para describir la "identidad en porcentaje" de dos o más secuencias. La identidad de una secuencia pueden existir con respecto a una región que tiene al menos aproximadamente 75-100 unidades secuenciales de longitud, con respecto a una región que tiene aproximadamente 50 unidades secuenciales de longitud, o, si no se especifica, a través de la secuencia entera. Esta definición también se refiere al complemento de una secuencia de prueba. A modo de ejemplo solo, dos o más secuencias de polipéptidos son idénticas cuando los residuos de aminoácidos son iguales, mientras que dos o más secuencias de polipéptidos son "sustancialmente idénticas" si los residuos de aminoácidos son aproximadamente el 60 % idénticos, aproximadamente el 65 % idénticos, aproximadamente el 70 % idénticos, aproximadamente el 75 % idénticos, aproximadamente el 80 % idénticos, aproximadamente el 85 % idénticos, aproximadamente el 90 % idénticos, o aproximadamente el 95 % idénticos con respecto a una región especificada. La identidad pueden existir con respecto a una región que tiene al menos aproximadamente 75-100 aminoácidos de longitud, con respecto a una región que tiene aproximadamente 50 aminoácidos en longitud, o, si no se especifica, a través de la secuencia entera de una secuencia de polipéptidos. Además, a modo de ejemplo solo, dos o más secuencias de polinucleótidos son idénticas cuando los residuos de ácidos nucleicos son iguales, mientras que dos o más secuencias de polinucleótidos son "sustancialmente idénticas" si los residuos de ácidos nucleicos son aproximadamente el 60 % idénticos, aproximadamente el 65 % idénticos, aproximadamente el 70 % idénticos, aproximadamente el 75 % idénticos, aproximadamente el 80 % idénticos, aproximadamente el 85 % idénticos, aproximadamente el 90 % idénticos, o aproximadamente el 95 % idénticos con respecto a una región especificada. La identidad puede existir con respecto a una región que tiene al menos aproximadamente 75-100 ácidos nucleicos en longitud, con respecto a una región que tiene aproximadamente 50 ácidos nucleicos en longitud, o, si no se especifica, a través de la secuencia entera de una secuencia de polinucleótidos.

Para la comparación de secuencias, normalmente una secuencia actúa de secuencia de referencia, con la que se comparan las secuencias de prueba. Si se usa un algoritmo de comparación de secuencias, secuencias de prueba y de referencia se entran en un ordenador, se designan coordenadas de subsecuencias, si fuera necesario, y se designan parámetros de programa del algoritmo de secuencias. Pueden usarse parámetros del programa por defecto, o pueden designarse parámetros alternativos. El algoritmo de comparación de secuencias calcula entonces el porcentaje de identidades de secuencia para las secuencias de prueba con respecto a la secuencia de referencia, basándose en los parámetros de programa.

El término "agente intercalante", también denominado "grupo intercalante", como se usa en el presente documento, se refiere a una sustancia química que puede insertarse en el espacio intramolecular de una molécula o el espacio

intermolecular entre moléculas. A modo de ejemplo solo, un agente intercalante o grupo puede ser una molécula que se inserta en las bases apiladas de la hélice doble de ADN.

El término “aislado”, como se usa en el presente documento, se refiere a separar y eliminar un componente de interés de componentes de no interés. Las sustancias aisladas pueden estar en tanto un estado seco como semi-seco, o en disolución, que incluye, pero no se limita a, una disolución acuosa. El componente aislado puede estar en un estado homogéneo o el componente aislado puede ser una parte de una composición farmacéutica que comprende vehículos y/o excipientes farmacéuticamente aceptables adicionales. La pureza y homogeneidad pueden determinarse usando técnicas de química analítica que incluyen, pero no se limitan a, electroforesis en gel de poliacrilamida o cromatografía líquida de alta resolución. Además, cuando un componente de interés se aísla y es la especie predominante presente en una preparación, el componente se describe en el presente documento como sustancialmente purificado. El término “purificado”, como se usa en el presente documento, puede referirse a un componente de interés que es al menos el 85 % puro, al menos el 90 % puro, al menos el 95 % puro, al menos el 99 % o de mayor pureza. A modo de ejemplo solo, los ácidos nucleicos o proteínas están “aislados” cuando tales ácidos nucleicos o proteínas están libres de al menos algunos de los componentes celulares con los que están asociados en el estado natural, o que el ácido nucleico o proteína se ha concentrado a un nivel superior a la concentración de su producción *in vivo* o *in vitro*. Por tanto, a modo de ejemplo, un gen se aísla cuando se separa de marcos de lectura abiertos que flanquean el gen y codifican una proteína distinta del gen de interés.

El término “marca”, como se usa en el presente documento, se refiere a una sustancia que se incorpora en un compuesto y es fácilmente detectada, por lo que puede detectarse y/o monitorizarse su distribución física.

El término “enlace”, como se usa en el presente documento, para referirse a enlaces o resto químico formado a partir de una reacción química entre el grupo funcional de un conector y otra molécula. Tales enlaces pueden incluir enlaces covalentes y enlaces no covalentes, mientras que tales restos químicos pueden incluir ésteres, carbonatos; iminas, ésteres de fosfato, hidrazonas, acetales, ortoésteres, enlaces peptídicos y enlaces de oligonucleótidos. Enlaces hidrolíticamente estables significa que los enlaces son sustancialmente estables en agua y no reaccionan con agua a valores de pH útiles, que incluyen, pero no se limitan a, bajo condiciones fisiológicas durante un periodo de tiempo prolongado, quizás incluso indefinidamente. Enlaces hidrolíticamente inestables o degradables significa que los enlaces son degradables en agua o en disoluciones acuosas, que incluyen, por ejemplo, sangre. Enlaces enzimáticamente inestables o degradables significa que el enlace puede ser degradado por una o más enzimas. A modo de ejemplo solo, el PEG y polímeros relacionados pueden incluir enlaces degradables en el esqueleto del polímero o en el grupo conector entre el esqueleto del polímero y uno o más de los grupos funcionales terminales de la molécula de polímero. Tales enlaces degradables incluyen, pero no se limitan a, enlaces éster formados mediante la reacción de PEG-ácidos carboxílicos o PEG-ácidos carboxílicos activados con grupos alcohol sobre un agente biológicamente activo, en los que tales grupos éster generalmente se hidrolizan bajo condiciones fisiológicas para liberar el agente biológicamente activo. Otros enlaces hidrolíticamente degradables incluyen enlaces carbonato; enlaces imina resultantes de la reacción de una amina y un aldehído; enlaces éster de fosfato formados haciendo reaccionar un alcohol con un grupo fosfato; enlaces hidrazona que son el producto de reacción de una hidrazida y un aldehído; enlaces acetal que son el producto de reacción de un aldehído y un alcohol; enlaces ortoéster que son el producto de reacción de un formiato y un alcohol; enlaces peptídicos formados por un grupo amina, que se incluyen en un extremo de un polímero tal como PEG, y un grupo carboxilo de un péptido; y enlaces de oligonucleótidos formados por un grupo fosforamidito, que incluyen en el extremo de un polímero, y un grupo hidroxilo en 5' de un oligonucleótido.

Los términos “medio” o “medios”, como se usan en el presente documento, se refieren a cualquier medio de cultivo usado para cultivar y recoger células y/o productos expresados y/o secretados por tales células. Tal “medio” o “medios” incluyen disolución, soportes sólidos, semi-sólidos o rígidos que pueden soportar o contener cualquier célula huésped, que incluye, a modo de ejemplo, células huésped bacterianas, células huésped de levadura, células huésped de insecto, células huésped de planta, células huésped eucariotas, células huésped de mamífero, células CHO, células huésped procariotas, *E. coli* o células huésped de *Pseudomonas*, y contenidos celulares. Tal “medio” o “medios” incluye, pero no se limita a, medio o medios en los que la célula huésped se ha cultivado en los que un polipéptido se ha secretado, que incluyen medio tanto antes como después de una etapa de proliferación. Tal “medio” o “medios” también incluye tampones o reactivos que contienen lisados de células huésped, a modo de ejemplo un polipéptido producido intracelularmente y las células huésped se lisan o rompen para liberar el polipéptido.

El término “metabolito”, como se usa en el presente documento, se refiere a un derivado de un polipéptido de aminoácidos naturales, un polipéptido de aminoácidos no naturales, un polipéptido de aminoácidos naturales modificado, o un polipéptido de aminoácidos no naturales modificado que se forma cuando se metaboliza el polipéptido de aminoácidos naturales, polipéptido de aminoácidos no naturales, polipéptido de aminoácidos naturales modificado o polipéptido de aminoácidos no naturales modificado. El término “metabolito activo” se refiere a un derivado biológicamente activo de un polipéptido de aminoácidos naturales, un polipéptido de aminoácidos no naturales, un polipéptido de aminoácidos naturales modificado o un polipéptido de aminoácidos no naturales modificado que se forma cuando se metaboliza el polipéptido de aminoácidos naturales, polipéptido de aminoácidos no naturales, polipéptido de aminoácidos naturales modificado o polipéptido de aminoácidos no naturales

modificado.

El término “metabolizado”, como se usa en el presente documento, se refiere a la suma de los procesos por los que una sustancia particular es cambiada por un organismo. Tales procedimientos incluyen reacciones de hidrólisis y reacciones catalizadas por enzimas. Información adicional sobre el metabolismo puede obtenerse de The Pharmacological Basis of Therapeutics, 9ª edición, McGraw-Hill (1996). A modo de ejemplo solo, metabolitos de polipéptidos de aminoácidos naturales, polipéptidos de aminoácidos no naturales, polipéptidos de aminoácidos naturales modificados o polipéptidos de aminoácidos no naturales modificados pueden identificarse tanto por administración de polipéptidos de aminoácidos naturales, polipéptidos de aminoácidos no naturales, polipéptidos de aminoácidos naturales modificados o polipéptidos de aminoácidos no naturales modificados a un huésped y análisis de muestras de tejido del huésped, como por incubación de polipéptidos de aminoácidos naturales, polipéptidos de aminoácidos no naturales, polipéptidos de aminoácidos naturales modificados o polipéptidos de aminoácidos no naturales modificados con células hepáticas *in vitro* y análisis de los compuestos resultantes.

El término “quelante de metal”, como se usa en el presente documento, se refiere a una molécula que forma un complejo metálico con iones metálicos. A modo de ejemplo, tales moléculas pueden formar dos o más enlaces de coordinación con un ión metálico central y pueden formar estructuras de anillo.

El término “resto que contiene metal”, como se usa en el presente documento, se refiere a un grupo que contiene un ión, átomo o partícula metálica. Tales restos incluyen cisplatino, iones metálicos quelados (tales como níquel, hierro y platino) y nanopartículas metálicas (tales como níquel, hierro y platino).

El término “resto que incorpora un átomo pesado”, como se usa en el presente documento, se refiere a un grupo que incorpora un ión de átomo que es normalmente más pesado que el carbono. Tales iones o átomos incluyen silicio, tungsteno, oro, plomo y uranio.

El término “modificado”, como se usa en el presente documento, se refiere a la presencia de un cambio en un aminoácido natural, un aminoácido no natural, un polipéptido de aminoácidos naturales o un polipéptido de aminoácidos no naturales. Tales cambios, o modificaciones, pueden obtenerse por modificaciones post-síntesis de aminoácidos naturales, aminoácidos no naturales, polipéptidos de aminoácidos naturales o polipéptidos de aminoácidos no naturales, o por modificación postraduccional de aminoácidos naturales, aminoácidos no naturales, polipéptidos de aminoácidos naturales o polipéptidos de aminoácidos no naturales. La forma “(modificado)” significa que el aminoácido natural, aminoácido no natural, polipéptido de aminoácidos naturales o polipéptido de aminoácidos no naturales que se tratan están opcionalmente modificados, es decir, el aminoácido natural, aminoácido no natural, polipéptido de aminoácidos naturales o polipéptido de aminoácidos no naturales en discusión pueden modificarse o no modificarse.

Como se usa en el presente documento, el término “semivida en suero modulada” se refiere a cambios positivos o negativos en la semivida circulante 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, aminoácido natural, aminoácido no natural, polipéptido de aminoácidos naturales o polipéptido de aminoácidos no naturales. A modo de ejemplo, la semivida en suero se mide tomando muestras de sangre en diversos momentos de tiempo después de la administración de la molécula biológicamente activa o molécula biológicamente activa modificada, y determinando la concentración de esa 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 un aumento en la semivida en suero, que puede permitir unas pautas de dosificación mejoradas o evitar efectos tóxicos. Tal aumento en suero puede ser al menos aproximadamente dos veces, al menos aproximadamente tres veces, al menos aproximadamente cinco veces, o al menos aproximadamente diez veces.

El término “semivida terapéutica modulada”, como se usa en el presente documento, se refiere a cambios positivos o negativos en la semivida de la cantidad terapéuticamente eficaz de una molécula biológicamente activa modificada, con respecto a su forma no modificada. A modo de ejemplo, las moléculas biológicamente activas modificadas incluyen aminoácido natural, aminoácido no natural, polipéptido de aminoácidos naturales o polipéptido de aminoácidos no naturales. A modo de ejemplo, la semivida terapéutica se mide midiendo propiedades farmacocinéticas y/o farmacodinámicas de la molécula en diversos momentos de tiempo después de la administración. Una semivida terapéutica elevada puede permitir una pauta de dosificación beneficiosa particular, una dosis total beneficiosa particular, o evita un efecto no deseado. A modo de ejemplo, el aumento de la semivida terapéutica pueden resultar de evada potencia, elevada o disminuida unión de la molécula modificada a su diana, un aumento o disminución en otro parámetro o mecanismo de acción de la molécula no modificada, o un aumento o disminución de la degradación de las moléculas por enzimas tales como, a modo de ejemplo solo, proteasas.

El término “nanopartícula”, 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.

El término “casi estequiométrica”, como se usa en el presente documento, se refiere a que la relación de los moles de compuestos que participan en una reacción química es aproximadamente 0,75 a aproximadamente 1,5 con

respecto al número de hidrógenos sobre la cadena lateral de amina aromática del polipéptido de aminoácidos no naturales. A modo de ejemplo, una amina primaria aromática tiene 2 hidrógenos, mientras que una amina secundaria tiene un hidrógeno.

Como se usa en el presente documento, el término “no eucariota” se refiere a organismos no eucariotas. A modo de ejemplo, un organismo no eucariota puede pertenecer a las eubacterias (que incluyen, pero no se limitan a, *Escherichia coli, Thermus thermophilus* o *Bacillus stearothermophilus, Pseudomonas fluorescens, Pseudomonas aeruginosa, Pseudomonas putida*), dominio filogenético, o las arqueas, que incluyen *Methanococcus jannaschii, Methanobacterium thermoautotrophicum, Archaeoglobus fulgidus, Pyrococcus furiosus, Pyrococcus horikoshii*, *Aeuropirum pernix* o *Halobacterium* tal como *Haloferax volcanii* y especies de *Halobacterium NRC-1,* o dominio filogenético.

Un “aminoácido no natural” se refiere a un aminoácido que no es uno de los 20 aminoácidos comunes o pirrolisina o selenocisteína. Otros términos que pueden usarse sinónimamente con el término “aminoácido no natural” son “aminoácido no naturalmente codificado”, “aminoácido no natural”, “aminoácido que no se produce naturalmente” y versiones con diversos guiones y sin guiones de los mismos. El término “aminoácido no natural” incluye aminoácidos que se producen naturalmente por modificación de un aminoácido naturalmente codificado (incluyendo los 20 aminoácidos comunes o pirrolisina y selenocisteína), pero no se incorporan por sí mismos en una cadena de polipéptido en crecimiento por el complejo de traducción. Ejemplos de aminoácidos que se producen naturalmente que no son naturalmente codificados incluyen *N*-acetilglucosaminil-L-serina, *N*-acetilglucosaminil-L-treonina y O-fosfotirosina. Adicionalmente, el término “aminoácido no natural” incluye aminoácidos que no se producen naturalmente y pueden obtenerse sintéticamente o pueden obtenerse por modificación de aminoácidos no naturales.

El término “ácido nucleico”, como se usa en el presente documento, se refiere a desoxirribonucleótidos, desoxirribonucleósidos, ribonucleósidos o ribonucleótidos y polímeros de los mismos en tanto forma mono- como bicatenaria. A modo de ejemplo solo, tales ácidos nucleicos y polímeros de ácidos nucleicos incluyen (i) análogos de nucleótidos naturales que tienen propiedades de unión similares como un ácido nucleico de referencia y se metabolizan de un modo similar a nucleótidos que se producen naturalmente; (ii) análogos de oligonucleótidos que incluyen PNA (ácido péptidonucleico), análogos de ADN usados en tecnología antisentido (fosforotioatos, fosforoamidatos y similares); (iii) variantes conservativamente modificadas de los mismos (que incluyen sustituciones de codones degenerados) y secuencias complementarias y secuencias explícitamente indicadas. A modo de ejemplo, las sustituciones de codones degenerados pueden lograrse generando secuencias en las que la tercera posición de uno o más codones seleccionados (o todos) está sustituida con residuos de bases mixtas y/o de desoxiinosina (Batzer y col., Nucleic Acid Res. 19:5081 (1991); Ohtsuka y col., J. Biol. Chem. 260:2605-2608 (1985); y Rossolini y col., Mol. Cell. Probes 8:91-98 (1994)).

El término “agente de oxidación”, como se usa en el presente documento, se refiere a un compuesto o material que puede eliminar un electrón de un compuesto que se oxida. Agentes de oxidación a modo de ejemplo incluyen glutatión oxidado, cistina, cistamina, ditiotreitol oxidado, eritreitol oxidado y oxígeno. Son adecuados una amplia variedad de agentes de oxidación para su uso en los procedimientos y composiciones descritos en el presente documento.

El término “marca de fotoafinidad”, 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 que la marca tiene una afinidad. A modo de ejemplo solo, un enlace tal puede ser covalente o no covalente.

El término “resto fotoenjaulado”, como se usa en el presente documento, se refiere a un grupo que, tras la iluminación a ciertas longitudes de onda, se une covalentemente o no covalentemente a otros iones o moléculas.

El término “grupo fotoescindible”, como se usa en el presente documento, se refiere a un grupo que se rompe tras la exposición a la luz.

El término “fotorreticulante”, 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.

El término “resto fotoisomerizable”, como se usa en el presente documento, se refiere a un grupo en el que, tras la iluminación con luz, cambia de una forma isomérica a otra.

El término “polialquilenglicol”, como se usa en el presente documento, se refiere a polioléteres poliméricos lineales o ramificados. Tales polialquilenglicoles incluyen polietilenglicol, polipropilenglicol, polibutilenglicol y derivados de los mismos. Otras realizaciones a modo de ejemplo se enumeran, por ejemplo, en catálogos de proveedores comerciales tales como el catálogo de Shearwater Corporation “Polyethylene Glycol and Derivatives for Biomedical Applications” (2001). A modo de ejemplo solo, tales polioléteres poliméricos tienen pesos moleculares promedio entre aproximadamente 0,1 kDa y aproximadamente 100 kDa. A modo de ejemplo, tales polioléteres poliméricos incluyen entre aproximadamente 100 Da y aproximadamente 100.000 Da o más. El peso molecular del polímero

puede estar entre aproximadamente 100 Da y aproximadamente 100.000 Da, que incluye, pero no se limita a, 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 y 100 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 100 Da y aproximadamente 50.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 100 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 1.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 5.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 10.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, la molécula de poli(etilenglicol) es un polímero ramificado. El peso molecular del PEG de cadena ramificada puede estar entre aproximadamente 1.000 Da y aproximadamente 100.000 Da 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 y 1.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 1.000 Da y aproximadamente 50.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 1.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 5.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 5.000 Da y aproximadamente 20.000 Da.

El término “polímero”, como se usa en el presente documento, se refiere a una molécula compuesta de subunidades de repetición. Tales moléculas incluyen, pero no se limitan a, polipéptidos, polinucleótidos, o polisacáridos o polialquilenglicoles.

Los términos “polipéptido”, “péptido” y “proteína” se usan indistintamente en el presente documento para referirse a un polímero de residuos de aminoácidos. Es decir, una descripción referida a un polipéptido se aplica igualmente a una descripción de un péptido y una descripción de una proteína, y viceversa. Los términos se aplican a polímeros de aminoácidos que se producen naturalmente, además de a polímeros de aminoácidos en los que uno o más residuos de aminoácidos es un aminoácido no natural. Adicionalmente, tales “polipéptidos”, “péptidos” y “proteínas” incluyen cadenas de aminoácidos de cualquier longitud, que incluyen proteínas de longitud completa, en las que los residuos de aminoácidos están ligados por enlaces peptídicos covalentes.

El término “postraduccionalmente modificado” se refiere a cualquier modificación de un aminoácido natural o no natural que se produce después de que un aminoácido tal se haya incorporado traduccionalmente en una cadena de polipéptidos. Tales modificaciones incluyen modificaciones *in vivo* co-traduccionales, modificaciones *in vitro* co-traduccionales (tales como en un sistema de traducción sin células), modificaciones *in vivo* postraduccionales y modificaciones *in vitro* postraduccionales.

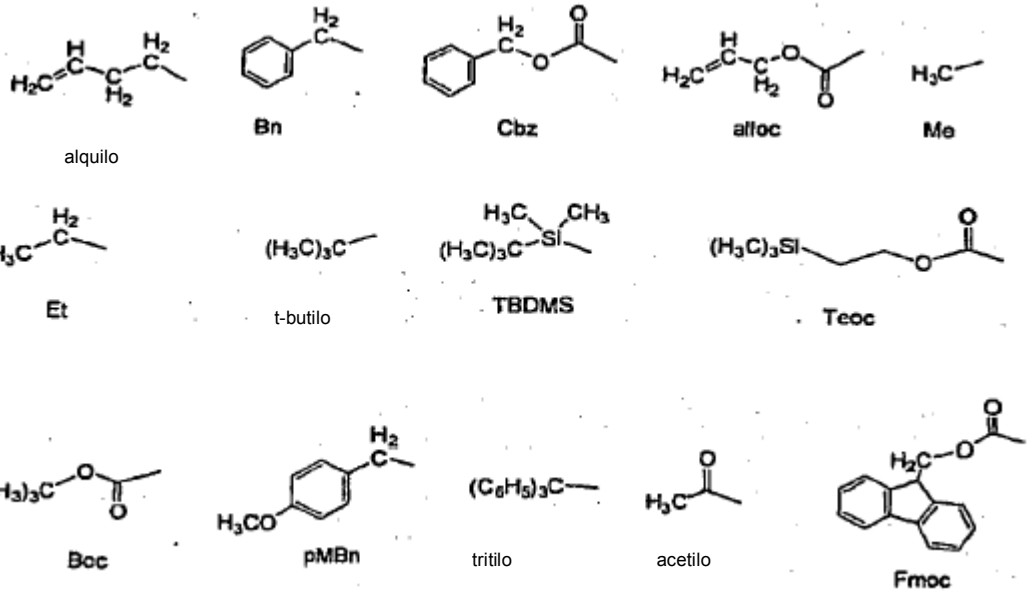
El término “profármaco”, como se usa en el presente documento, se refiere a un agente que se convierte en el fármaco parental *in vivo* o *in vitro.* Los beneficios de tales profármacos incluyen (i) facilidad de administración en comparación con el fármaco parental; (ii) el profármaco puede estar biodisponible por administración por vía oral mientras que el parental no; y (iii) el profármaco también puede tener solubilidad mejorada en composiciones farmacéuticas en comparación con el fármaco parental. Un profármaco incluye un derivado farmacológicamente activo, o de actividad reducida, derivado de un fármaco activo. Los profármacos pueden diseñarse para modular la cantidad de un fármaco o molécula biológicamente activa que llega a un sitio de acción deseado mediante la manipulación de las propiedades de un fármaco, tales como propiedades fisicoquímicas, biofarmacéuticas o farmacocinéticas. Los profármacos se convierten en el fármaco activo dentro del cuerpo mediante reacciones enzimáticas o no enzimáticas. Los profármacos pueden proporcionar propiedades fisicoquímicas mejoradas tales como mejor solubilidad, características de administración potenciadas, tales como elegir específicamente como diana una célula, tejido, órgano o ligando particular, y valor terapéutico mejorado del fármaco.

El término “cantidad profilácticamente eficaz”, como se usa en el presente documento, se refiere a esa cantidad de una composición que contiene al menos un polipéptido de aminoácidos no naturales o al menos un polipéptido de aminoácidos no naturales modificado profilácticamente aplicado a un paciente que aliviará de algún modo uno o más de los síntomas de una enfermedad, afección o trastorno que está tratándose. En tales aplicaciones profilácticas, tales cantidades pueden depender del estado de salud del paciente, peso y similares. Se considera perfectamente dentro de la habilidad de la materia el determinar tales cantidades profilácticamente eficaces por experimentación rutinaria, que incluye un ensayo clínico de aumento de dosis.

El término “protegido”, como se usa en el presente documento, se refiere a la presencia de un “grupo protector” o resto que previene la reacción del grupo funcional químicamente reactivo bajo ciertas condiciones de reacción. El grupo protector variará dependiendo del tipo de grupo químicamente reactivo que se proteja. A modo de ejemplo solo, (i) si el grupo químicamente reactivo es una amina o una hidrazida, el grupo protector puede seleccionarse de 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.

A modo de ejemplo solo, los grupos de bloqueo/protectores también pueden seleccionarse de:



Adicionalmente, los grupos protectores incluyen, pero no se limitan a, grupos fotolábiles tales como Nvoc y MeNvoc y otros grupos protectores conocidos en la técnica. Otros grupos protectores se describen en Greene y Wuts, Protective Groups in Organic Synthesis, 3ª ed., John Wiley & Sons, New York, NY, 1999.

El término "resto radiactivo", como se usa en el presente documento, se refiere a un grupo cuyos núcleos desprenden espontáneamente radiación nuclear, tal como partículas alfa, beta o gamma; en el que las partículas alfa son núcleos de helio, las partículas beta son electrones y las partículas gamma son fotones de alta energía.

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

El término "célula huésped recombinante", también denominada "célula huésped", se refiere a una célula que incluye un polinucleótido exógeno, en el que los procedimientos usados para insertar el polinucleótido exógeno en una célula incluyen captación directa, transducción, apareamiento f, u otros procedimientos conocidos en la técnica para crear células huésped recombinantes. A modo de ejemplo solo, tal polinucleótido exógeno puede ser un vector no integrado, que incluye, pero no se limita a, un plásmido, o puede estar integrado en el genoma del huésped.

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

El término "agente reductor", como se usa en el presente documento, se refiere a un compuesto o material que puede añadir un electrón a un compuesto que se reduce. Agentes reductores a modo de ejemplo incluyen ditiotreitol (DTT), 2-mercaptoetanol, ditioeritritol, cisteína, cisteamina (2-aminoetantiol) y glutatión reducido. Tales agentes reductores pueden usarse, a modo de ejemplo solo, para mantener grupos sulfhidrilo en el estado reducido y para reducir enlaces disulfuro intra- o intermoleculares.

"Replegamiento", como se usa en el presente documento, describe cualquier proceso, reacción o procedimiento que transforme un estado inapropiadamente plegado o desplegado en una conformación nativa o apropiadamente plegada. A modo de ejemplo solo, el replegamiento transforma los polipéptidos que contienen enlace disulfuro de un estado inapropiadamente plegado o desplegado a una conformación nativa o apropiadamente plegada con respecto a enlaces disulfuro. Tales polipéptidos que contienen enlace disulfuro pueden ser polipéptidos de aminoácidos naturales o polipéptidos de aminoácidos no naturales.

El término "resina", como se usa en el presente documento, se refiere a perlas de polímero insoluble de alto peso molecular. A modo de ejemplo solo, tales perlas pueden usarse como soportes para la síntesis de péptidos en fase

sólida, o sitios para la unión de moléculas antes de la purificación.

El término “sacárido”, como se usa en el presente documento, se refiere a una serie de hidratos de carbono que incluyen, pero no se limitan a, azúcares, monosacáridos, oligosacáridos y polisacáridos.

El término “se hibrida selectivamente con” o “se hibrida específicamente con”, como se usa en el presente documento, se refiere a la unión, duplexado o hibridación de una molécula con una secuencia de nucleótidos particular bajo condiciones de hibridación rigurosas cuando esa secuencia está presente en una mezcla compleja que incluye, pero no se limita a, ADN o ARN celular total o de bibliotecas.

El término “marca de espín”, como se usa en el presente documento, se refiere a moléculas que contienen un átomo o un grupo de átomos que presenta un espín electrónico no apareado (es decir, un grupo paramagnético estable) que puede detectarse por espectroscopía de resonancia de espín electrónico y puede unirse a otra molécula. Tales moléculas de marca de espín incluyen radicales nitrilo y nitróxidos, y puede ser marcas de un único espín o marcas de doble espín.

El término “estequiométrica”, como se usa en el presente documento, se refiere a que la relación de los moles de compuestos que participan en una reacción química es aproximadamente 0,9 a aproximadamente 1,1 con respecto al número de hidrógenos sobre la cadena lateral de amina aromática del polipéptido de aminoácidos no naturales. A modo de ejemplo, una amina primaria aromática tiene 2 hidrógenos, mientras que una amina secundaria tiene un hidrógeno.

El término “similar a estequiométrica”, como se usa en el presente documento, se refiere a una reacción química que se vuelve estequiométrica o casi estequiométrica tras los cambios en las condiciones de reacción o en presencia de aditivos. Tales cambios en las condiciones de reacción incluyen un aumento en la temperatura o cambio en el pH. Tales aditivos incluyen acelerantes.

El término “condiciones de hibridación rigurosas” se refiere a la hibridación de secuencias de ADN, ARN, PNA u otro mimético de ácido nucleico, o combinaciones de los mismos, en condiciones de fuerza iónica baja y temperatura alta. A modo de ejemplo, bajo condiciones rigurosas una sonda se hibridará con su subsecuencia diana en una mezcla compleja de ácido nucleico (que incluye ADN o ARN celular total o de biblioteca), pero no se hibridará con otras secuencias en la mezcla compleja. Condiciones rigurosas son dependientes de secuencia y serán diferentes en diferentes circunstancias. A modo de ejemplo, secuencias más largas se hibridan específicamente a mayores temperaturas. Condiciones de hibridación rigurosas incluyen, (i) aproximadamente 5-10 ºC por debajo del punto de fusión térmico (T_m) para la secuencia específica a una fuerza iónica y pH definidos; (ii) la concentración de sales es aproximadamente 0,01 M a aproximadamente 1,0 M a aproximadamente pH 7,0 a aproximadamente pH 8,3 y la temperatura es al menos aproximadamente 30 ºC para sondas cortas (que incluyen 10 a 50 nucleótidos) y al menos aproximadamente 60 ºC para sondas largas (que incluyen más de 50 nucleótidos); (iii) la adición de agentes desestabilizantes que incluyen formamida, (iv) 50 % de formamida, 5X SSC y 1 % de SDS, incubar a 42 ºC, o 5X SSC, 1 % de SDS, incubar a 65 ºC, con lavado en 0,2X SSC y 0,1 % de SDS a 65 ºC durante entre aproximadamente 5 minutos y aproximadamente 120 minutos. A modo de ejemplo solo, la detección de hibridación selectiva o específica incluye una señal positiva de al menos dos veces el ruido de fondo. Una amplia guía para la hibridación de ácidos nucleicos se encuentra en Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes, “Overview of principles of hybridization and the strategy of nucleic acid assays” (1993).

El término “sujeto”, como se usa en el presente documento, se refiere a un animal que es el objetivo del tratamiento, observación o experimento. A modo de ejemplo solo, un sujeto puede ser un mamífero que incluye un ser humano.

El término “sustancialmente purificado”, como se usa en el presente documento, se refiere a un componente de interés que puede estar sustancialmente o esencialmente libre de otros componentes que normalmente acompañan a o interaccionan con el componente de interés antes de la purificación. A modo de ejemplo solo, un componente de interés puede estar “sustancialmente purificado” cuando la preparación del componente de interés contiene menos de aproximadamente el 30 %, menos de aproximadamente el 25 %, menos de aproximadamente el 20 %, menos de aproximadamente el 15 %, menos de aproximadamente el 10 %, menos de aproximadamente el 5 %, menos de aproximadamente el 4 %, menos de aproximadamente el 3 %, menos de aproximadamente el 2 %, o menos de aproximadamente el 1 % (en peso seco) de componentes contaminantes. Así, un componente “sustancialmente purificado” de interés puede tener un nivel de pureza de aproximadamente el 70 %, aproximadamente el 75 %, aproximadamente el 80 %, aproximadamente el 85 %, aproximadamente el 90 %, aproximadamente el 95 %, aproximadamente el 96 %, aproximadamente el 97 %, aproximadamente el 98 %, aproximadamente el 99 % o mayor. A modo de ejemplo solo, un polipéptido de aminoácidos naturales o un polipéptido de aminoácidos no naturales puede purificarse a partir de una célula nativa, o célula huésped en el caso de polipéptidos de aminoácidos naturales o polipéptidos de aminoácidos no naturales recombinantemente producidos. A modo de ejemplo, una preparación de un polipéptido de aminoácidos naturales o un polipéptido de aminoácidos no naturales puede estar “sustancialmente purificada” cuando la preparación contiene menos de aproximadamente el 30 %, menos de aproximadamente el 25 %, menos de aproximadamente el 20 %, menos de aproximadamente el 15 %, menos de

aproximadamente el 10 %, menos de aproximadamente el 5 %, menos de aproximadamente el 4 %, menos de aproximadamente el 3 %, menos de aproximadamente el 2 %, o menos de aproximadamente el 1 % (en peso seco) de material contaminante. A modo de ejemplo, cuando un polipéptido de aminoácidos naturales o un polipéptido de aminoácidos no naturales se produce recombinantemente por células huésped, el polipéptido de aminoácidos naturales o polipéptido de aminoácidos no naturales puede estar presente a aproximadamente el 30 %, aproximadamente el 25 %, aproximadamente el 20 %, aproximadamente el 15 %, aproximadamente el 10 %, aproximadamente el 5 %, aproximadamente el 4 %, aproximadamente el 3 %, aproximadamente el 2 %, o aproximadamente el 1 % o menos del peso seco de las células. A modo de ejemplo, cuando un polipéptido de aminoácidos naturales o un polipéptido de aminoácidos no naturales se produce recombinantemente por células huésped, el polipéptido de aminoácidos naturales o polipéptido de aminoácidos no naturales puede estar 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 del peso seco de las células. A modo de ejemplo, los polipéptidos de aminoácidos naturales o polipéptidos de aminoácidos no naturales “sustancialmente purificados” pueden tener un nivel de pureza de aproximadamente el 30 %, aproximadamente el 35 %, aproximadamente el 40 %, aproximadamente el 45 %, aproximadamente el 50 %, aproximadamente el 55 %, aproximadamente el 60 %, aproximadamente el 65 %, aproximadamente el 70 %, aproximadamente el 75 %, aproximadamente el 80 %, aproximadamente el 85 %, aproximadamente el 90 %, aproximadamente el 95 %, aproximadamente el 99 % o mayor como se ha determinado por procedimientos apropiados, que incluyen análisis de SDS/PAGE, RP-HPLC, SEC y electroforesis capilar.

El término “sustituyentes”, también denominado “sustituyentes no interferentes”, se refiere a grupos que pueden usarse para sustituir otro grupo en una molécula. Tales grupos incluyen halógeno, alquilo C_1-C_{10}, alquenilo C_2-C_{10}, alquinilo C_2-C_{10}, alcoxi C_1-C_{10}, aralquilo C_5-C_{12}, cicloalquilo C_3-C_{12}, cicloalquenilo C_4-C_{12}, fenilo, fenilo sustituido, toluoílo, xilenilo, bifenilo, alcoxi C_2-C_{12}-alquilo, alcoxi C_5-C_{12}-arilo, ariloxi C_5-C_{12}-alquilo, oxiarilo C_7-C_{12}, alquil C_1-C_6-sulfinilo, alquil C_1-C_{10}-sulfonilo, -$(CH_2)_m$-O-(alquilo C_1-C_{10}) en el que m es de 1 a 8, arilo, arilo sustituido, alcoxi sustituido, fluoroalquilo, radical heterocíclico, radical heterocíclico sustituido, nitroalquilo, -NO_2, -CN, -NRC(O)-(alquilo C_1-C_{10}), -C(O)-(alquilo C_1-C_{10}), alquil C_2-C_{10}-tioalquilo, -C(O)O-(alquilo C_1-C_{10}), -OH, -SO_2, =S, -COOH, -NR_2, carbonilo, -C(O)-(alquil C_1-C_{10})-CF_3, -C(O)-CF_3, -C(O)NR_2, -(aril C_1-C_{10})-S-(arilo C_6-C_{10}), -C(O)-(arilo C_6-C_{10}), -$(CH_2)_m$-O-$(CH_2)_m$-O-(alquilo C_1-C_{10}) en la que cada m es de 1 a 8, -C(O)NR_2, -C(S)NR_2, -SO_2NR_2 - NRC(O)NR_2, -NRC(S)NR_2, sales de los mismos, y similares. Cada grupo R en la lista precedente incluye H, alquilo o alquilo sustituido, arilo o arilo sustituido, o alcarilo. Cuando los grupos de sustituyentes se especifican por sus fórmulas químicas convencionales, escritas de izquierda a derecha, engloban igualmente los sustituyentes químicamente idénticos que resultarían de escribir la estructura de derecha a izquierda, por ejemplo, -CH_2O- es equivalente a -OCH_2-.

A modo de ejemplo solo, sustituyentes para radicales alquilo y heteroalquilo (incluyendo aquellos grupos denominados alquileno, alquenilo, heteroalquileno, heteroalquenilo, alquinilo, cicloalquilo, heterocicloalquilo, cicloalquenilo y heterocicloalquenilo) incluyen, pero no se limitan a: -OR, =O, =NR, =N-OR, -NR_2, -SR, -halógeno, -SiR_3, -OC(O)R, -C(O)R, -CO_2R, -$CONR_2$, -$OC(O)NR_2$, -NRC(O)R, -NR-C(O)NR_2, -$NR(O)_2R$, -NR-C(NR_2)=NR, -S(O)R, -$S(O)_2R$, -$S(O)_2NR_2$, -$NRSO_2R$, -CN y -NO_2. Cada grupo R en la lista precedente incluye hidrógeno, heteroalquilo sustituido o sin sustituir, arilo sustituido o sin sustituir, que incluye arilo sustituido con 1-3 halógenos, alquilo sustituido o sin sustituir, grupos alcoxi o tioalcoxi, o grupos aralquilo. Cuando dos grupos R están unidos al mismo átomo de nitrógeno, pueden combinarse con el átomo de nitrógeno para formar un anillo de 5, 6 ó 7 miembros. Por ejemplo, -NR_2 pretende incluir 1-pirrolidinilo y 4-morfolinilo.

A modo de ejemplo, sustituyentes para grupos arilo y heteroarilo incluyen -OR, =O, =NR, =N-OR, -NR_2, -SR, -halógeno, -SiR_3, -OC(O)R, -C(O)R, -CO_2R, -$CONR_2$, -$OC(O)NR_2$, -NRC(O)R, -NR-C(O)NR_2, -$NR(O)_2R$, -NR-C(NR_2)=NR, -S(O)R, -$S(O)_2R$, -$S(O)_2NR_2$, -$NRSO_2R$, -CN, -NO_2, -R, -N_3, -$CH(Ph)_2$, fluoroalcoxi (C_1-C_4) y fluoroalquilo (C_1-C_4), en un número que oscila de cero al número total de valencias libres en el sistema de anillos aromáticos; y si cada grupo R en la lista precedente incluye hidrógeno, alquilo, heteroalquilo, arilo y heteroarilo.

El término “cantidad terapéuticamente eficaz”, como se usa en el presente documento, se refiere a la cantidad de una composición que contiene al menos un polipéptido de aminoácidos no naturales y/o al menos un polipéptido de aminoácidos no naturales modificado administrada a un paciente que ya padece una enfermedad, afección o trastorno, suficiente para curar o al menos detener parcialmente, o aliviar de algún modo uno o más de los síntomas de la enfermedad, trastorno o afección que está tratándose. La eficacia de tales composiciones depende de las condiciones que incluyen la gravedad y evolución de la enfermedad, trastorno o afección, terapia previa, el estado de salud del paciente y respuesta a los fármacos, y el criterio del médico práctico. A modo de ejemplo solo, pueden determinarse cantidades terapéuticamente eficaces por experimentación rutinaria, que incluye un ensayo clínico de aumento de dosis.

El término “tioalcoxi”, como se usa en el presente documento, se refiere a grupos alquilo que contienen azufre ligados a moléculas mediante un átomo de oxígeno.

El término “punto de fusión térmico” o T_m es la temperatura (bajo fuerza iónica, pH y concentración nucleica definida)

a la que el 50 % de las sondas complementarias a una diana se hibridan con la secuencia diana en equilibrio.

El término “resto tóxico”, como se usa en el presente documento, se refiere a un compuesto que puede producir daño o muerte.

Los términos “tratan”, “tratar” o “tratamiento”, como se usan en el presente documento, incluyen aliviar, abatir o mejorar una enfermedad o síntomas de afección, prevenir síntomas adicionales, mejorar o prevenir las causas metabólicas subyacentes de los síntomas, inhibir la enfermedad o afección, por ejemplo, detener el desarrollo de la enfermedad o afección, aliviar la enfermedad o afección, causar regresión de la enfermedad o afección, aliviar una afección producida por la enfermedad o afección, o parar los síntomas de la enfermedad o afección. Los términos “tratan”, “tratar” 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 sea soluble en disolventes acuosos. Tales polímeros solubles en agua incluyen polietilenglicol, polietilenglicol-propionaldehído, monoalcoxi C_1-C_{10} o derivados de ariloxi del mismo (descritos en la patente de EE.UU. nº 5.252.714), monometoxi-polietilenglicol, polivinilpirrolidona, poli(alcohol vinílico), poliaminoácidos, éter divinílico-anhídrido maleico, *N*-(2-hidroxipropil)-metacrilamida, dextrano, derivados de dextrano que incluyen sulfato de dextrano, polipropilenglicol, copolímero de poli(óxido de propileno)/óxido de etileno, poliol polioxietilado, heparina, fragmentos de heparina, polisacáridos, oligosacáridos, glicanos, celulosa y derivados de celulosa, que incluyen, pero no se limitan a, metilcelulosa y carboximetilcelulosa, albúmina de suero, almidón y derivados de almidón, polipéptidos, polialquilenglicol y derivados del mismo, copolímeros de polialquilenglicoles y derivados del mismo, polivinil etil éteres y alfa-beta-poli[(2-hidroxietil)-DL-aspartamida, y similares, o mezclas de los mismos. A modo de ejemplo, el enlace de tales polímeros solubles en agua a un polipéptido de aminoácidos naturales o un polipéptido no natural puede producir cambios que incluyen semivida en suero elevada o modulada, semivida terapéutica elevada o modulada con respecto a la forma no modificada, inmunogenicidad modulada, características de asociación física modulada que incluyen agregación y formación de multímeros, unión a receptores alterada, unión alterada a uno o más componentes de unión, y dimerización o multimerización de receptores alterada. Además, tales polímeros solubles en agua pueden o pueden no tener su propia actividad biológica.

A menos que se indique lo contrario, se emplean procedimientos convencionales de espectroscopía de masas, RMN, HPLC, química de proteínas, bioquímica, técnicas de ADN recombinante y farmacología, dentro de la habilidad de la materia.

Los compuestos (incluyendo aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados, y reactivos para producir los compuestos anteriormente mencionados) presentados en el presente documento incluyen compuestos isotópicamente marcados, que son idénticos a aquellos citados en las diversas fórmulas y estructuras presentadas en el presente documento, excepto por el hecho de que uno o más átomos están sustituidos con un átomo que tiene una masa atómica o número másico diferente de la masa atómica o número másico normalmente encontrado en la naturaleza. Ejemplos de isótopos que pueden incorporarse en los presentes compuestos incluyen isótopos de hidrógeno, carbono, nitrógeno, oxígeno, flúor y cloro, tales como ^{2}H, ^{3}H, ^{13}C, ^{14}C, ^{15}N, ^{18}O, ^{17}O, ^{35}S, ^{18}F, ^{36}Cl, respectivamente. Ciertos compuestos isotópicamente marcados descritos en el presente documento, por ejemplo, aquellos en los que se incorporan los isótopos radiactivos tales como ^{3}H y ^{14}C, son útiles en ensayos de distribución en tejido de fármaco y/o sustrato. Además, la sustitución con isótopos tales como deuterio, es decir, ^{2}H, puede proporcionar ciertas ventajas terapéuticas que resultan de mayor estabilidad metabólica, por ejemplo, elevada semivida *in vivo* o requisitos de dosificación reducidos.

Algunos de los compuestos en el presente documento (que incluyen aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados, y reactivos para producir los compuestos anteriormente mencionados) tienen átomos de carbono asimétricos y pueden, por tanto, existir como enantiómeros o diaestereómeros. Pueden separarse mezclas diaestereoméricas en sus diaestereómeros individuales basándose en sus diferencias fisicoquímicas mediante procedimientos conocidos, por ejemplo, por cromatografía y/o cristalización fraccionada. Los enantiómeros pueden separarse convirtiendo la mezcla enantiomérica en una mezcla diaestereomérica mediante reacción con un compuesto ópticamente activo apropiado (por ejemplo, alcohol), separando los diaestereómeros y convirtiendo (por ejemplo, hidrolizando) los diaestereómeros individuales en los enantiómeros puros correspondientes. Todos aquellos isómeros, que incluyen diaestereómeros, enantiómeros y mezclas de los mismos, se consideran parte de las composiciones descritas en el presente documento.

En realizaciones adicionales u otras realizaciones, los compuestos descritos en el presente documento (que incluyen aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados, y reactivos para producir los compuestos anteriormente mencionados) se usan en forma de pro-fármacos. En realizaciones adicionales u otras realizaciones, los compuestos descritos en el presente documento (que incluyen aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados, y reactivos para producir los compuestos anteriormente mencionados) son metabolizados tras la administración a un organismo en necesidad de producir un metabolito que a continuación se usa para

producir un efecto deseado, que incluye un efecto terapéutico deseado. En otras realizaciones o realizaciones adicionales, son metabolitos activos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados o no modificados.

Los procedimientos y formulaciones descritos en el presente documento incluyen el uso de *N*-óxidos, formas cristalinas (también conocidas como polimorfos), o sales farmacéuticamente aceptables de aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados. En ciertas realizaciones, los aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados pueden existir como tautómeros. Todos los tautómeros están incluidos dentro del alcance de los aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados presentados en el presente documento. Además, los aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento pueden existir en formas no solvatadas, además de solvatadas, con disolventes farmacéuticamente aceptables tales como agua, etanol y similares. Las formas solvatadas de los aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados presentados en el presente documento también se considera que se desvelan en el presente documento.

Algunos de los compuestos en el presente documento (que incluyen aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados, y reactivos para producir los compuestos anteriormente mencionados) pueden existir en varias formas tautómeras. Todas aquellas formas tautómeras se consideran como parte de las composiciones descritas en el presente documento. Por tanto, por ejemplo, todas las formas de enol-ceto de cualquier compuesto (incluyendo aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados, y reactivos para producir los compuestos anteriormente mencionados) en el presente documento se consideran como parte de las composiciones descritas en el presente documento.

Algunos de los compuestos en el presente documento (que incluyen aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados, y reactivos para producir cualquiera de los compuestos anteriormente mencionados) son ácidos y pueden formar una sal con un catión farmacéuticamente aceptable. Algunos de los compuestos en el presente documento (que incluyen aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados, y reactivos para producir los compuestos anteriormente mencionados) pueden ser básicos y, por consiguiente, pueden formar una sal con un anión farmacéuticamente aceptable. Todas aquellas sales, que incluyen di-sales, están dentro del alcance de las composiciones descritas en el presente documento y pueden prepararse mediante procedimientos convencionales. Por ejemplo, las sales pueden prepararse poniendo en contacto las entidades ácidas y básicas en tanto un medio acuoso, no acuoso como parcialmente acuoso. Las sales se recuperan usando al menos una de las siguientes técnicas: filtración, precipitación con un no disolvente seguido de filtración, evaporación del disolvente, o, en el caso de disoluciones acuosas, liofilización.

Las sales, por ejemplo, incluyen: (1) sales de adición de ácido, formadas con ácidos inorgánicos tales como ácido clorhídrico, ácido bromhídrico, ácido sulfúrico, ácido nítrico, ácido fosfórico y similares; o formadas con ácidos orgánicos tales como ácido acético, ácido propiónico, ácido hexanoico, ácido ciclopentanopropiónico, ácido glicólico, ácido pirúvico, ácido láctico, ácido malónico, ácido succínico, ácido málico, ácido maleico, ácido fumárico, ácido tartárico, ácido cítrico, ácido benzoico, ácido 3-(4-hidroxibenzoil)benzoico, ácido cinámico, ácido mandélico, ácido metanosulfónico, ácido etanosulfónico, ácido 1,2-etanodisulfónico, ácido 2-hidroxietanosulfónico, ácido bencenosulfónico, ácido 2-naftalenosulfónico, ácido 4-metilbiciclo-[2.2.2]oct-2-eno-1-carboxílico, ácido glucoheptónico, ácido 4,4'-metilenbis-(3-hidroxi-2-eno-1-carboxílico), ácido 3-fenilpropiónico, ácido trimetilacético, ácido terc-butilacético, ácido laurilsulfúrico, ácido glucónico, ácido glutámico, ácido hidroxinaftoico, ácido salicílico, ácido esteárico, ácido mucónico y similares; (2) sales formadas cuando un protón ácido presente en el compuesto parental tanto está sustituido con un ión metálico, por ejemplo, un ión de metal alcalino, un ión de metal alcalinotérreo, o un ión aluminio, como está coordinado con una base orgánica. Bases orgánicas aceptables incluyen etanolamina, dietanolamina, trietanolamina, trometamina, N-metilglucamina y similares. Bases inorgánicas aceptables incluyen hidróxido de aluminio, hidróxido de calcio, hidróxido potásico, carbonato sódico, hidróxido sódico y similares.

Debe entenderse que una referencia a una sal incluye las formas de adición de disolvente o formas cristalinas de las mismas, particularmente solvatos o polimorfos. Los solvatos contienen tanto cantidades estequiométricas como no estequiométricas de un disolvente, y se forman frecuentemente durante el procedimiento de cristalización. Los hidratos se forman cuando el disolvente es agua, o los alcoholatos se forman cuando el disolvente es alcohol. Los polimorfos incluyen las diferentes disposiciones de empaquetamiento cristalino de la misma composición elemental de un compuesto. Los polimorfos tienen normalmente diferentes patrones de difracción de rayos X, espectros de infrarrojos, puntos de fusión, densidad, dureza, forma del cristal, propiedades ópticas y eléctricas, estabilidad y solubilidad. Diversos factores tales como el disolvente de recristalización, tasa de cristalización y temperatura de almacenamiento pueden hacer que domine una forma monocristalina.

BREVE DESCRIPCIÓN DE LAS FIGURAS

Puede obtenerse un mejor entendimiento de las características y ventajas de los presentes procedimientos y composiciones por referencia a la siguiente descripción detallada que expone realizaciones ilustrativas, en las que se utilizan los principios de los procedimientos, composiciones, dispositivos y aparatos de los presentes inventores, y los dibujos adjuntos que ilustran realizaciones de la invención, además de aspectos ilustrativos para ayudar en el trabajo de la invención.

La FIG. 1 ilustra un aspecto de procedimientos de selección y diseño de un polipéptido que va a modificarse usando los procedimientos, composiciones y técnicas descritos en el presente documento.

La FIG. 2 presenta ejemplos del tipos de aminoácidos no naturales que pueden usarse en el procedimiento de la invención. Tales aminoácidos no naturales pueden usarse en o incorporarse en cualquiera de los procedimientos, composiciones, técnicas y estrategias para preparar, purificar, caracterizar y usar aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento.

La FIG. 3 presenta ejemplos ilustrativos de los tipos de aminoácidos no naturales descritos en el presente documento. Tales aminoácidos no naturales pueden usarse en o incorporarse en cualquiera de los procedimientos, composiciones, técnicas y estrategias para preparar, purificar, caracterizar y usar aminoácidos no naturales; polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento. Para la FIG. 3 solo, X representa un halógeno.

La FIG. 4 presenta ejemplos ilustrativos de los tipos de aminoácidos no naturales descritos en el presente documento. Tales aminoácidos no naturales pueden usarse en o incorporarse en cualquiera de los procedimientos, composiciones, técnicas y estrategias para preparar, purificar, caracterizar y usar aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento.

La FIG. 5 presenta ejemplos ilustrativos de los tipos de aminoácidos no naturales descritos en el presente documento. Tales aminoácidos no naturales pueden usarse en o incorporarse en cualquiera de los procedimientos, composiciones, técnicas y estrategias para preparar, purificar, caracterizar y usar aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento.

La FIG. 6 presenta ejemplos ilustrativos de los tipos de aminoácidos no naturales descritos en el presente documento. Tales aminoácidos no naturales pueden usarse en o incorporarse en cualquiera de los procedimientos, composiciones, técnicas y estrategias para preparar, purificar, caracterizar y usar aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento. Para la FIG. 6 solo, X representa un halógeno.

La FIG. 7 presenta ejemplos ilustrativos de los tipos de aminoácidos no naturales descritos en el presente documento. Tales aminoácidos no naturales pueden usarse en o incorporarse en cualquiera de los procedimientos, composiciones, técnicas y estrategias para preparar, purificar, caracterizar y usar aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento.

La FIG. 8 presenta ejemplos de los tipos de aminoácidos no naturales descritos en el presente documento y que son realizaciones de la invención. Tales aminoácidos no naturales pueden usarse en o incorporarse en cualquiera de los procedimientos, composiciones, técnicas y estrategias para preparar, purificar, caracterizar y usar aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento.

La FIG. 9 presenta ejemplos ilustrativos de los tipos de aminoácidos no naturales descritos en el presente documento. Tales aminoácidos no naturales pueden usarse en o incorporarse en cualquiera de los procedimientos, composiciones, técnicas y estrategias para preparar, purificar, caracterizar y usar aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento. Para la FIG. 9 solo, X representa un halógeno.

La FIG. 10 presenta ejemplos ilustrativos de los tipos de aminoácidos no naturales descritos en el presente documento. Tales aminoácidos no naturales pueden usarse en o incorporarse en cualquiera de los procedimientos, composiciones, técnicas y estrategias para preparar, purificar, caracterizar y usar aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento.

La FIG. 11 presenta un ejemplo ilustrativo de la metodología sintética usada para preparar los aminoácidos no naturales de la invención y como se describen en el presente documento. Tales aminoácidos no naturales pueden usarse en o incorporarse en cualquiera de los procedimientos, composiciones, técnicas y estrategias para preparar, purificar, caracterizar y usar aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento.

La FIG. 12 presenta un ejemplo ilustrativo de la metodología sintética usada para preparar los aminoácidos no naturales de la invención y como se describen en el presente documento. Tales aminoácidos no naturales pueden usarse en o incorporarse en cualquiera de los procedimientos, composiciones, técnicas y estrategias para preparar, purificar, caracterizar y usar aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento.

La FIG. 13 presenta un ejemplo ilustrativo de la metodología sintética usada para preparar los aminoácidos no naturales de la invención y como se describen en el presente documento. Tales aminoácidos no naturales

pueden usarse en o incorporarse en cualquiera de los procedimientos, composiciones, técnicas y estrategias para preparar, purificar, caracterizar y usar aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento.

La FIG. 14 presenta ejemplos ilustrativos de aminoácidos no naturales descritos en el presente documento que contienen restos de amina enmascarados y/o protegidos que pueden convertirse en restos de amina desenmascarados y/o desprotegidos. Tales aminoácidos no naturales pueden usarse en o incorporarse en cualquiera de los procedimientos, composiciones, técnicas y estrategias para preparar, purificar, caracterizar y usar aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento.

La FIG. 15 presenta ejemplos ilustrativos de aminoácidos no naturales descritos en el presente documento que contienen cadenas laterales de amina enmascarada y/o protegida para la aminación reductora con un reactivo que contiene aldehído. Tales aminoácidos no naturales pueden usarse en o incorporarse en cualquiera de los procedimientos, composiciones, técnicas y estrategias para preparar, purificar, caracterizar y usar aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento.

La FIG. 16 presenta ejemplos ilustrativos de formación de polipéptidos con aminoácidos no naturales que contienen aminas aromáticas por alquilaciones reductoras con reactivos que contienen aldehído.

La FIG. 17 presenta ejemplos ilustrativos de aminoácidos no naturales que contienen aldehído protegido aromático descritos en el presente documento. Tales aminoácidos no naturales pueden usarse en o incorporarse en cualquiera de los procedimientos, composiciones, técnicas y estrategias para preparar, purificar, caracterizar y usar aminoácidos no naturales, polipéptidos de aminoácidos no naturales y polipéptidos de aminoácidos no naturales modificados descritos en el presente documento.

La FIG. 18 presenta ejemplos ilustrativos de la formación de polipéptidos con aminoácidos no naturales que contienen aldehído por desprotección/desenmascaramiento postraduccional de un precursor protegido/enmascarado, seguido de aminaciones reductoras con reactivos que contienen aminas aromáticas.

La FIG. 19 presenta ejemplos ilustrativos de alquilación reductora y aminación reductora selectivas de sitio de polipéptidos de aminoácidos no naturales.

La FIG 20 presenta ejemplos ilustrativos de las alquilaciones reductoras de urotensina-II reducida (UT-II-SH) con propionaldehído (I) y benzaldehído (II) y los cromatogramas de HPLC correspondientes.

La FIG 21 presenta ejemplos ilustrativos de alquilaciones reductoras de urotensina-II (UT-II) con propionaldehído (I) y benzaldehído (II) y los cromatogramas de HPLC correspondientes.

La FIG. 22 presenta ejemplos ilustrativos de alquilaciones reductoras del péptido XT-8 con propionaldehído (I) y benzaldehído (II), isobutaldehído (III) y pivalaldehído (IV) y los cromatogramas de HPLC correspondientes.

La FIG. 23 presenta ejemplos ilustrativos de alquilaciones reductoras del péptido SXT-9 con propionaldehído (I), benzaldehído (II), isobutaldehído (III) y pivalaldehído (IV) y los cromatogramas de HPLC correspondientes.

La FIG. 24 presenta ejemplos ilustrativos de alquilaciones reductoras del péptido HXT-9 con propionaldehído (I), benzaldehído (II), isobutaldehído (III) y pivalaldehído (IV) y los cromatogramas de HPLC correspondientes.

La FIG. 25 presenta ejemplos ilustrativos de alquilaciones reductoras del péptido WXT-9 con propionaldehído (I), benzaldehído (II) e isobutaldehído (III) y los cromatogramas de HPLC correspondientes.

La FIG. 26 presenta ejemplos ilustrativos de alquilaciones reductoras del péptido NXT-9 con propionaldehído (I) y benzaldehído (II) y los cromatogramas de HPLC correspondientes.

La FIG. 27 presenta ejemplos ilustrativos de alquilaciones reductoras del péptido RXT-10 con propionaldehído (I) y benzaldehído (II) y los cromatogramas de HPLC correspondientes.

La FIG. 28 presenta ejemplos ilustrativos de alquilaciones reductoras del péptido AXT-11 con propionaldehído (I) y benzaldehído (II) y los cromatogramas de HPLC correspondientes.

La FIG. 29 presenta ejemplos ilustrativos de alquilaciones reductoras del péptido AXT-11 con diferentes aldehídos (I-III) y los cromatogramas de HPLC correspondientes.

La FIG. 30 presenta ejemplos ilustrativos de alquilaciones reductoras del péptido AXT-11 con diferentes aldehídos (IV-VII) y los cromatogramas de HPLC correspondientes.

La FIG. 31 presenta ejemplos ilustrativos de reacciones competitivas del péptido AXT-11 con un aldehído y una 1,3-dicetona (I-III) y los cromatogramas de HPLC correspondientes.

La FIG. 32 presenta ejemplos ilustrativos de reacciones competitivas del péptido NXT-9 con un aldehído y una 1,2-dicetona (I-III) y los cromatogramas de HPLC correspondientes.

La FIG. 33 presenta ejemplos ilustrativos de reacciones competitivas del péptido MXT-9 con diferentes aldehídos y cetonas (I-V) y los cromatogramas de HPLC correspondientes.

La FIG. 34 presenta ejemplos ilustrativos de reducción del péptido MXT-9-N3 a MXT-9NH2 (I), seguido de alquilaciones reductoras con propionaldehído (II) y benzaldehído (III) y los cromatogramas de HPLC correspondientes.

La FIG. 35 presenta ejemplos ilustrativos y comparación de PEGilación de péptidos por PEGilación del extremo N y PEGilación de aminoácidos no naturales que contienen aminas aromáticas descritos en el presente documento.

La FIG. 36 presenta ejemplos ilustrativos de reactivos de PEG-aldehído usados para PEGilar aminoácidos no naturales que contienen aminas aromáticas incorporados en péptidos.

La FIG. 37 presenta ejemplos ilustrativos de PEGilación de MT-9 por alquilación reductora del extremo N.

La FIG. 38 presenta ejemplos ilustrativos de PEGilación de MXT-9 por alquilación reductora de aminoácido no natural que contiene amina aromática incorporado en MXT-9.

La FIG. 39 presenta ejemplos ilustrativos de PEGilación de hGH e IFNα por alquilación reductora de aminoácido no natural que contiene amina aromática incorporado en hGH e IFNα.
La FIG. 40 presenta una imagen ilustrativa de un gel de electroforesis de diversas PEGilaciones de hGH.

DESCRIPCIÓN DETALLADA DE LA INVENCIÓN

I. Introducción

Recientemente se ha informado de una tecnología completamente nueva en las ciencias de las proteínas que promete vencer muchas de las limitaciones asociadas a las modificaciones de proteínas específicas de sitio. Específicamente, se han añadido nuevos componentes a la maquinaria biosintética de proteínas del procariota *Escherichia coli (E. coli*) (por ejemplo, L. Wang, y col., (2001), Science 292:498-500) y el eucariota *Saccharomyces cerevisiae* (*S. cerevisiae)* (por ejemplo, J. Chin y col., Science 301:964-7 (2003)), que ha permitido la incorporación de aminoácidos no naturales en proteínas *in vivo.* Varios aminoácidos nuevos con novedosas propiedades químicas, físicas o biológicas, que incluyen marcas de fotoafinidad y aminoácidos fotoisomerizables, aminoácidos fotorreticulantes (véanse, por ejemplo, Chin, J. W., y col. (2002) Proc. Natl. Acad. Sci. U. S. A. 99:11020-11024; y, Chin, J. W., y col., (2002) J. Am. Chem. Soc. 124:9026-9027), aminoácidos que contienen átomos pesados, cetoaminoácidos y aminoácidos glucosilados se han incorporado eficazmente y con alta fidelidad en proteínas en *E. coli* y en levadura en respuesta al codón ámbar, TAG, usando esta metodología. Véanse, por ejemplo, J. W. Chin y col., (2002), Journal of the American Chemical Society 124:9026-9027; J. W. Chin, & P. G. Schultz, (2002), ChemBioChem 3(11):1135-1137; J. W. Chin, y col., (2002), PNAS United States of America 99:11020-11024; y, L. Wang, & P. G. Schultz, (2002), Chem. Comm., 1:1-11. Estos estudios han demostrado que es posible introducir selectiva y rutinariamente grupos funcionales químicos que no se encuentran en proteínas, que son químicamente inertes a todos los grupos funcionales encontrados en los 20 aminoácidos comunes genéticamente codificados y que pueden usarse para reaccionar eficaz y selectivamente para formar enlaces covalentes estables.

II. Visión general

La FIG. 1 presenta una visión general de las composiciones, procedimientos y técnicas que se describen en el presente documento. En un nivel, en el presente documento se describen las herramientas (procedimientos, composiciones, técnicas) para crear y usar un polipéptido que comprende al menos un aminoácido no natural o aminoácido no natural modificado con una amina aromática o una amina heteroaromática. El grupo heteroaromático incluye furanos, pirroles, tiofenos, piridinas, quinolinas, isoquinolinas, imidazol, tiazoles, pirimidinas, piridazinas, pirazinas, benzotiazoles, tiazolopiridinas, oxazoles, benzoxazoles, oxazolopiridinas, tiazolopirimidinas, oxazolopirimidinas, benzoxazinas y benzotiazinas. Tales aminoácidos no naturales pueden contener adicionalmente funcionalidad, que incluye una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos.

Como se muestra en la FIG. 1, en un aspecto son procedimientos de selección y diseño de un polipéptido que va a modificarse usando los procedimientos, composiciones y técnicas descritos en el presente documento. El nuevo polipéptido puede diseñarse *de novo*, que incluye a modo de ejemplo solo, como parte de un procedimiento de selección de alto rendimiento (en cuyo caso pueden diseñarse, sintetizarse, caracterizarse y/o probarse numerosos polipéptidos) o basándose en los intereses del investigador. El nuevo polipéptido también puede diseñarse basándose en la estructura de un polipéptido conocido o parcialmente caracterizado. A modo de ejemplo solo, la superfamilia de genes de la hormona de crecimiento (véase más adelante) ha sido objeto de intenso estudio por la comunidad científica; puede diseñarse un nuevo polipéptido basándose en la estructura de un miembro o miembros de esta superfamilia de genes. Los principios para seleccionar qué aminoácido(s) sustituir y/o modificar se describen por separado en el presente documento. La elección de qué modificación emplear también se describe en el presente documento, y puede usarse para cumplir la necesidad del investigador o usuario final. Tales necesidades pueden incluir, pero no se limitan a, manipular la eficacia terapéutica del polipéptido, mejorar el perfil de seguridad

del polipéptido, ajustar la farmacocinética, farmacológica y/o farmacodinámica del polipéptido, tal como, a modo de ejemplo solo, aumentar la solubilidad en agua, biodisponibilidad, aumentar la semivida en suero, aumentar la semivida terapéutica, modular la inmunogenicidad, modular la actividad biológica, o prologar el tiempo en circulación. Además, tales modificaciones incluyen, a modo de ejemplo solo, proporcionar funcionalidad adicional al polipéptido, incorporar una etiqueta, marca o señal detectable en el polipéptido, facilitar las propiedades de aislamiento del polipéptido, y cualquier combinación de las modificaciones anteriormente mencionadas.

En el presente documento también se describen aminoácidos no naturales que se han modificado o pueden modificarse para contener una amina aromática o una amina heteroaromática. El grupo heteroaromático incluye furanos, pirroles, tiofenos, piridinas, quinolinas, isoquinolinas, imidazol, tiazoles, pirimidinas, piridazinas, pirazinas, benzotiazoles, tiazolopiridinas, oxazoles, benzoxazoles, oxazolopiridinas, tiazolopirimidinas, oxazolopirimidinas, benzoxazinas y benzotiazinas. En este aspecto están incluidos procedimientos de producción, purificación, caracterización y uso de tales aminoácidos no naturales. En otro aspecto descrito en el presente documento están procedimientos, estrategias y técnicas para incorporar al menos uno de tales aminoácidos no naturales en un polipéptido. También están incluidos en este aspecto procedimientos de producción, purificación, caracterización y uso de tales polipéptidos que contienen al menos uno de tales aminoácidos no naturales. También están incluidos en este aspecto composiciones y procedimientos de producción, purificación, caracterización y uso de oligonucleótidos (que incluyen ADN y ARN) que pueden usarse para producir, al menos en parte, un polipéptido que contiene al menos un aminoácido no natural. También están incluidos en este aspecto composiciones y procedimientos de producción, purificación, caracterización y uso de células que pueden expresar tales oligonucleótidos que pueden usarse para producir, al menos en parte, un polipéptido que contiene al menos un aminoácido no natural.

Así, en el presente documento se proporcionan y describen polipéptidos que comprenden al menos un aminoácido no natural o aminoácido no natural modificado con una amina aromática o una amina heteroaromática. El grupo heteroaromático incluye furanos, pirroles, tiofenos, piridinas, quinolinas, isoquinolinas, imidazol, tiazoles, pirimidinas, piridazinas, pirazinas, benzotiazoles, tiazolopiridinas, oxazoles, benzoxazoles, oxazolopiridinas, tiazolopirimidinas, oxazolopirimidinas, benzoxazinas y benzotiazinas. En ciertas realizaciones, polipéptidos con al menos un aminoácido no natural o aminoácido no natural modificado con una amina aromática o una amina heteroaromática incluyen al menos una modificación postraduccional en alguna posición sobre el polipéptido. En tales realizaciones, los aminoácidos no naturales modificados en el heterociclo pueden incluir furanos, pirroles, tiofenos, piridinas, quinolinas, isoquinolinas, imidazol, tiazoles, pirimidinas, piridazinas, pirazinas, benzotiazoles, tiazolopiridinas, oxazoles, benzoxazoles, oxazolopiridinas, tiazolopirimidinas, oxazolopirimidinas, benzoxazinas y benzotiazinas. En algunas realizaciones, la modificación postraduccional se produce mediante la maquinaria celular (por ejemplo, glucosilación, acetilación, acilación, modificación por lípidos, palmitoilación, adición de palmitato, fosforilación, modificación del enlace con glucolípido, y similares), en muchos casos, tales modificaciones postraduccionales basadas en la maquinaria celular se producen en los sitios de aminoácidos que se producen naturalmente sobre el polipéptido, sin embargo, en ciertas realizaciones, las modificaciones postraduccionales basadas en la maquinaria celular se producen sobre el (los) sitio(s) de aminoácido(s) no natural(es) sobre el polipéptido.

En otras realizaciones, la modificación postraduccional no utiliza maquinaria celular, pero la funcionalidad se proporciona en su lugar por la unión de una molécula (que incluye una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica: un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos) que comprende un segundo grupo reactivo con el al menos un aminoácido no natural que comprende un primer grupo reactivo (que incluye aminoácido no natural que contiene un grupo funcional de amina aromática o amina heteroaromática) que utiliza metodología de alquilación reductora. En ciertas realizaciones, tales segundos grupos reactivos pueden ser grupos que contienen carbonilo, que incluyen aldehídos y cetonas. En ciertas realizaciones, la modificación postraduccional se hace *in vivo* en una célula eucariota o en una célula no eucariota. En ciertas realizaciones, la modificación postraduccional se hace *in vitro.* También están incluidos en este aspecto procedimientos de producción, purificación, caracterización y uso de tales polipéptidos que contienen al menos uno de tales aminoácidos no

naturales postraduccionalmente modificados.

También están incluidos dentro del alcance de los procedimientos, composiciones, estrategias y técnicas descritos en el presente documento reactivos que pueden reaccionar con un aminoácido no natural (que contiene tanto un grupo de amina aromática, un grupo de amina heteroaromática, o formas protegidas de los mismos) que es parte de un polipéptido de manera que se produzca cualquiera de las modificaciones postraduccionales anteriormente mencionadas. En general, el aminoácido no natural postraduccionalmente modificado resultante contendrá al menos un aldehído o cetona reductoramente aminado, que puede someterse adicionalmente a reacciones de modificación. También están incluidos en este aspecto procedimientos de producción, purificación, caracterización y uso de tales reactivos capaces de cualquiera de tales modificaciones postraduccionales de tal(es) aminoácido(s) no natural(es).

En ciertas realizaciones, la proteína incluye al menos una modificación postraduccional que se hace *in vivo* por una célula huésped, en la que la modificación postraduccional no se hace normalmente por otro tipo de célula huésped. En ciertas realizaciones, la proteína incluye al menos una modificación postraduccional que se hace *in vivo* por una célula eucariota, en la que la modificación postraduccional normalmente no se hace por una célula no eucariota. Ejemplos de tales modificaciones postraduccionales incluyen glucosilación, acetilación, acilación, modificación por lípidos, palmitoilación, adición de palmitato, fosforilación, modificación del enlace con glucolípido, y similares. En una realización, la modificación postraduccional comprende unión de un oligosacárido a una asparagina por un enlace de GlcNAc-asparagina (que incluye, pero no se limita a, en los que el oligosacárido comprende $(GlcNAc\text{-}Man)_2\text{-}Man\text{-}GlcNAc\text{-}GlcNAc$, y similares). En otra realización, la modificación postraduccional comprende la unión de un oligosacárido (que incluye, pero no se limita a, Gal-GalNAc, Gal-GlcNAc, etc.) a una serina o treonina por un enlace de GalNAc-serina, GalNAc-treonina, GlcNAc-serina o de GlcNAc-treonina. Ejemplos de secuencias señal de la secreción incluyen una secuencia señal de la secreción procariota, una secuencia señal de la secreción eucariota, una secuencia señal de la secreción eucariota optimizada en 5' para expresión bacteriana, una secuencia señal de la secreción novedosa, secuencia señal de la secreción de pectato liasa, secuencia señal de la secreción de Omp A y una secuencia señal de la secreción de fago. Ejemplos de secuencias señal de la secreción incluyen STII (procariota), Fd GIII y M13 (fago), Bgl2 (levadura), y la secuencia señal bla derivada de un transposón. En ciertas realizaciones, una proteína o polipéptido puede comprender una secuencia de secreción o de localización, una marca de epítope, una marca FLAG, una marca de polihistidina, una fusión de GST y/o similares. También están incluidos en este aspecto procedimientos de producción, purificación, caracterización y uso de tales polipéptidos que contienen al menos un modificación postraduccional tal. En otras realizaciones, el polipéptido de aminoácidos no naturales glucosilado se produce en una forma no glucosilada. Una forma no glucosilada tal de un aminoácido no natural glucosilado puede producirse mediante procedimientos que incluyen eliminación química o enzimática de grupos de oligosacáridos de un polipéptido de aminoácidos no naturales glucosilado aislado o sustancialmente purificado o no purificado; producción del aminoácido no natural en un huésped que no glucosila un polipéptido de aminoácidos no naturales tal (incluyendo un huésped tal procariotas o eucariotas manipulados o mutados para no glucosilar un polipéptido tal), la introducción de un inhibidor de la glucosilación en el medio de cultivo celular en el que un polipéptido de aminoácidos no naturales tal está siendo producido por un eucariota que normalmente glucosilaría un polipéptido tal, o una combinación de cualquiera de tales procedimientos. En el presente documento también se describen tales formas no glucosiladas de polipéptidos de aminoácidos no naturales normalmente glucosilados (por normalmente glucosilado se indica un polipéptido que se glucosilaría cuando se produjera en condiciones en las que polipéptidos que se producen naturalmente están glucosilados). Por supuesto, tales formas no glucosiladas de polipéptidos de aminoácidos no naturales normalmente glucosilados pueden estar en una forma no purificada, una forma sustancialmente purificada, o en una forma aislada.

En ciertas realizaciones, la proteína incluye al menos una modificación postraduccional que se hace *in vitro,* en la que la modificación postraduccional es estequiométrica, similar a estequiométrica o casi estequiométrica. Ejemplos de tales modificaciones postraduccionales incluyen alquilaciones reductoras de grupos amina aromática o grupos amina heteroaromática con reactivos que contienen carbonilo usando un agente reductor. En ciertas realizaciones, tales modificaciones postraduccionales incluyen alquilaciones reductoras de grupos amina aromática o grupos amina heteroaromática con reactivos que contienen aldehído usando un agente reductor. En ciertas realizaciones, tales modificaciones postraduccionales incluyen alquilaciones reductoras de grupos amina aromática o grupos amina heteroaromática con reactivos que contienen aldehído usando agente reductor de cianoborohidruro de sodio.

El polipéptido que contienen aminoácidos no naturales puede contener al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, o diez o más aminoácidos no naturales que contienen una amina aromática, amina heteroaromática, o formas protegidas de las mismas. Los aminoácidos no naturales pueden ser iguales o diferentes, por ejemplo, puede haber 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más sitios diferentes en la proteína que comprenden 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más aminoácidos no naturales diferentes. En ciertas realizaciones, al menos uno, pero menos de todos, de un aminoácido particular presente en una versión que se produce naturalmente de la proteína está sustituido con un aminoácido no natural. El polipéptido de aminoácidos no naturales también puede contener uno o más aminoácidos no naturales que no son aminoácidos no naturales que contienen una amina aromática, amina heteroaromática, o formas protegidas de las mismas.

Los procedimientos y composiciones proporcionados y descritos en el presente documento incluyen polipéptidos que

comprenden al menos un aminoácido no natural que contiene una amina aromática, amina heteroaromática, o formas protegidas de las mismas. La introducción de al menos un aminoácido no natural tal en un polipéptido puede permitir la aplicación de químicas de conjugación que implican reacciones químicas específicas, que incluyen uno o más aminoácidos no naturales mientras que no reaccionen con los 20 aminoácidos que se producen comúnmente. Tales reacciones químicas específicas que implican tales cadenas laterales del aminoácido incorporadas incluyen metodologías de química de alquilación reductora adecuadas para los grupos funcionales particulares o sustituyentes presentes. Una vez incorporadas, las cadenas laterales del aminoácido pueden entonces modificarse utilizando metodologías de química conocidas para aquellos expertos habituales en la materia por ser adecuadas para los grupos funcionales particulares o sustituyentes presentes en el aminoácido naturalmente codificado.

Los procedimientos y composiciones de aminoácidos no naturales descritos en el presente documento proporcionan conjugados de sustancias que tienen una amplia variedad de grupos funcionales, sustituyentes o restos, con otras sustancias que incluyen una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos.

En otro aspecto de las composiciones, procedimientos, técnicas y estrategias descritos en el presente documento están procedimientos de estudio o uso de cualquiera de los polipéptidos de aminoácidos no naturales modificados o no modificados anteriormente mencionados. Incluidos dentro de este aspecto, a modo de ejemplo solo, están usos terapéuticos, de diagnóstico, basados en ensayos, industriales, cosméticos, de biología vegetal, medioambientales, de producción de energía y/o militares que se beneficiarían de un polipéptido que comprende un polipéptido o proteína de aminoácido no natural modificado o no modificado.

III. Localización de aminoácidos no naturales en polipéptidos

Los procedimientos y composiciones descritos en el presente documento incluyen la incorporación de uno o más aminoácidos no naturales en un polipéptido. Uno o más aminoácidos no naturales pueden incorporarse en una o más posiciones particulares que no alteran la actividad del polipéptido. Esto puede lograrse haciendo sustituciones “conservativas”, que incluyen, pero no se limitan a, sustituir aminoácidos hidrófobos con aminoácidos hidrófobos no naturales o naturales, aminoácidos voluminosos con aminoácidos voluminosos no naturales o naturales, aminoácidos hidrófilos con aminoácidos hidrófilos no naturales o naturales y/o insertar el aminoácido no natural en una localización que no se requiere para actividad.

Puede emplearse una variedad de enfoques bioquímicos y estructurales para seleccionar los sitios deseados para la sustitución con un aminoácido no natural dentro del polipéptido. Cualquier posición de la cadena del polipéptido es adecuada para la selección para incorporar un aminoácido no natural, y la selección puede basarse en el diseño racional o por selección aleatoria para cualquier fin deseado o no particular. La selección de sitios deseados puede ser para producir un polipéptido de aminoácidos no naturales (que puede ser adicionalmente modificado o seguir sin modificar) que tiene cualquier propiedad o actividad deseada, que incluye, pero no se limita a, agonistas, super-agonistas, agonista inversos, antagonistas, moduladores de la unión al receptor, moduladores de la actividad del receptor, moduladores de la unión a componentes de ligantes, moduladores de la actividad de componentes de unión, moduladores de la conformación de componentes de unión, formación de dímeros o multímeros, ningún cambio a la actividad o propiedad en comparación con la molécula nativa, o manipular cualquier propiedad física o química del polipéptido tal como solubilidad, agregación o estabilidad. Por ejemplo, las localizaciones en el polipéptido requerido para la actividad biológica de un polipéptido pueden identificarse usando análisis de mutaciones puntuales, procedimientos de barrido de alanina o de barrido de homólogos conocidos en la técnica. Residuos distintos de aquellos identificados como críticos para la actividad biológica por mutagénesis de barrido de alanina o de homólogos pueden ser buenos candidatos para sustitución con un aminoácido no natural dependiendo de la actividad deseada buscada para el polipéptido. Alternativamente, los sitios identificados como críticos para la actividad biológica también pueden ser buenos candidatos para sustitución con un aminoácido no natural, dependiendo de nuevo de la actividad deseada buscada para el polipéptido. Otra alternativa sería hacer

simplemente sustituciones en serie en cada posición sobre la cadena de polipéptidos con un aminoácido no natural y observar el efecto sobre las actividades del polipéptido. Cualquier medio, técnica o procedimiento para seleccionar una posición para la sustitución con un aminoácido no natural en cualquier polipéptido es adecuado para su uso en los procedimientos, técnicas y composiciones descritos en el presente documento.

La estructura y actividad de mutantes que se producen naturalmente de un polipéptido que contienen deleciones también puede examinarse para determinar regiones de la proteína que es probable que sean tolerantes a la sustitución con un aminoácido no natural. Una vez se han eliminado los residuos que es probable que sean intolerantes a la sustitución con aminoácidos no naturales, el impacto de las sustituciones propuestas en cada una de las posiciones restantes puede examinarse a partir de la estructura tridimensional del polipéptido relevante, y cualquier ligando o proteína de unión asociado. Están disponibles estructuras cristalográficas de rayos X y de RMN de muchos polipéptidos en Protein Data Bank (PDB, www.rcsb.org), una base de datos centralizada que contiene datos estructurales tridimensionales de moléculas de proteínas grandes y ácidos nucleicos. Además, pueden hacerse modelos investigando la estructura secundaria y terciaria de polipéptidos, si no están disponibles datos estructurales tridimensionales. Así, puede obtenerse fácilmente la identidad de posiciones de aminoácidos que pueden estar sustituidas con aminoácidos no naturales.

Sitios a modo de ejemplo de incorporación de un aminoácido no natural incluyen, pero no se limitan a, aquellos que se excluyen de posibles regiones de unión de receptores o regiones para la unión a proteínas de unión o ligandos, pueden estar completamente o parcialmente expuestos a disolvente, tener interacciones mínimas o no de enlaces de hidrógeno con residuos cercanos, pueden exponerse mínimamente a residuos reactivos cercanos, y pueden estar en regiones que son altamente flexibles como se ha predicho por la estructura cristalina tridimensional de un polipéptido particular con su receptor asociado, ligando o proteínas de unión.

Una amplia variedad de aminoácidos no naturales puede sustituirse en, o incorporarse en, una posición dada en un polipéptido. En general, un aminoácido no natural particular se selecciona para incorporación basándose en un examen de la estructura cristalina tridimensional de un polipéptido con su ligando asociado, receptor y/o proteínas de unión, una preferencia de sustituciones conservativas (es decir, aminoácidos no naturales basados en arilo, tales como p-acetilfenilalanina o sustituir Phe, Tyr o Trp por O-propargiltirosina), y la química de conjugación específica que se desea introducir en la proteína de polipéptido.

En una realización, el procedimiento incluye adicionalmente incorporar en la proteína el aminoácido no natural, en la que el aminoácido no natural comprende un primer grupo reactivo; y poner en contacto la proteína con una molécula (que incluye una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos) que comprende un segundo grupo reactivo. En ciertas realizaciones, el primer grupo reactivo es un resto de amina sobre una amina aromática y el segundo grupo reactivo es un aldehído, en el que el grupo de amina se alquila reductoramente tras el contacto con el aldehído en presencia de un agente reductor, tal como cianoborohidruro de sodio. En otras realizaciones, el primer grupo reactivo es un resto de amina sobre una amina heteroaromática y el segundo grupo reactivo es un aldehído, en el que el grupo de amina se alquila reductoramente tras el contacto con el aldehído en presencia de un agente reductor, tal como cianoborohidruro de sodio.

En algunos casos, la(s) sustitución (sustituciones) o incorporación (incorporaciones) de aminoácidos no naturales se combinarán con otras adiciones, sustituciones o deleciones dentro del polipéptido para afectar otros rasgos biológicos. En algunos casos, las otras adiciones, sustituciones o deleciones pueden aumentar la estabilidad (que incluyen la resistencia a degradación proteolítica) del polipéptido o aumentar la afinidad del polipéptido por su receptor, ligando y/o proteínas de unión apropiadas. En algunos casos, las otras adiciones, sustituciones o deleciones pueden aumentar la solubilidad (que incluye, pero no se limita a, cuando se expresan en *E. coli* u otras células huésped) del polipéptido. En algunas realizaciones, los sitios se seleccionan para sustitución con un aminoácido naturalmente codificado o no natural, además de otro sitio para la incorporación de un aminoácido no

natural con el fin de aumentar la solubilidad del polipéptido tras la expresión en *E. coli* u otras células huésped recombinantes. En algunas realizaciones, los polipéptidos comprenden otra adición, sustitución o deleción que modula la afinidad por el ligando asociado, proteínas de unión y/o receptor, modula (que incluye, pero no se limita a, aumenta o disminuye) la dimerización de receptores, estabiliza dímeros de receptores, modula la semivida en circulación, modula la liberación o biodisponibilidad, facilita la purificación, o mejora o altera una vía de administración particular. Similarmente, el polipéptido puede comprender secuencias de escisión química o por enzimas, secuencias de escisión por proteasas, grupo reactivos, dominios de unión a anticuerpo (que incluyen, pero no se limitan a, FLAG o poli-His) u otras secuencias basadas en afinidad (que incluyen, pero no se limitan a, FLAG, poli-His, GST, etc.) o moléculas ligadas (que incluyen, pero no se limitan a, biotina) que mejoran la detección (que incluye, pero no se limitan a, GFP), purificación, transporte a través de tejidos o membranas celulares, liberación o activación de profármacos, reducción de tamaño, u otros rasgos del polipéptido.

IV. Familia de supergenes de la hormona de crecimiento como ejemplo

Los procedimientos, composiciones, estrategias y técnicas descritos en el presente documento no se limitan a un tipo, clase o familia particular de polipéptidos o proteínas. De hecho, prácticamente cualquier polipéptido puede diseñarse o modificarse para incluir al menos un aminoácido no natural modificado o no modificado descrito en el presente documento. A modo de ejemplo solo, el polipéptido puede ser homólogo a una proteína terapéutica seleccionada del grupo que consiste en: alfa-1 antitripsina, angiostatina, factor antihemolítico, anticuerpo, fragmento de anticuerpo, apolipoproteína, apoproteína, factor natriurético auricular, polipéptido natriurético auricular, péptido auricular, quimiocina C-X-C, T39765, NAP-2, ENA-78, gro-a, gro-b, gro-c, IP-10, GCP-2, NAP-4, SDF-1, PF4, MIG, calcitonina, ligando c-kit, citocina, quimiocina CC, proteína 1 quimioatrayente de monocitos, proteína 2 quimioatrayente de monocitos, proteína 3 quimioatrayente de monocitos, proteína 1 alfa inflamatoria de monocitos, proteína 1 beta inflamatoria de monocitos, RANTES, 1309, R83915, R91733, HCC1, T58847, D31065, T64262, CD40, ligando CD40, ligando c-kit, colágeno, factor estimulante de colonias (CSF), factor 5a del complemento, inhibidor del complemento, receptor 1 del complemento, citocina, péptido 78 activante de neutrófilos epiteliales, MIP-16, MCP-1, factor de crecimiento epidérmico (EGF), péptido activante de neutrófilos epiteliales, eritropoyetina (EPO), toxina exfoliante, factor IX, factor VII, factor VIII, factor X, factor de crecimiento de fibroblastos (FGF), fibrinógeno, fibronectina, proteína de cuatro haces helicoidales, G-CSF, glp-1, GM-CSF, glucocerebrosidasa, gonadotropina, factor de crecimiento, receptor de factores de crecimiento, grf, proteína hedgehog, hemoglobina, factor de crecimiento de hepatocitos (hGF), hirudina, hormona de crecimiento humana (hGH), albúmina de suero humano, ICAM-1, receptor de ICAM-1, LFA-1, receptor de LFA-1, insulina, factor de crecimiento similar a la insulina (IGF), IGF-I, IGF-II, interferón (IFN), IFN-alfa, IFN-beta, IFN-gamma, cualquier molécula similar a interferón o miembro de la familia de IFN, interleucina (IL), IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, factor de crecimiento de queratinocitos (KGF), lactoferrina, factor inhibidor de la leucemia, luciferasa, neurturina, factor inhibidor de neutrófilos (NIF), oncostatina M, proteína osteogénica, producto oncogénico, paracitonina, hormona paratiroidea, PD-ECSF, PDGF, hormona peptídica, pleiotropina, proteína A, proteína G, pth, endotoxina A pirogénica, exotoxina B pirogénica, exotoxina C pirogénica, pyy, relaxina, renina, SCF, proteína biosintética pequeña, receptor I del complemento soluble, I-CAM 1 soluble, receptor de interleucina soluble, receptor de TNF soluble, somatomedina, somatostatina, somatotropina, estreptocinasa, superantígenos, enterotoxina estafilocócica, SEA, SEB, SEC1, SEC2, SEC3, SED, SEE, receptor de hormonas esteroideas, superóxido dismutasa, toxina del síndrome del choque tóxico, timosina alfa 1, activador tisular del plasminógeno, factor de crecimiento tumoral (TGF), factor de necrosis tumoral, factor de necrosis tumoral alfa, factor de necrosis tumoral beta, receptor del factor de necrosis tumoral (TNFR), proteína VLA-4, proteína VCAM-1, factor de crecimiento endotelial vascular (VEGF), urocinasa, 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 corticosterona.

Así, la siguiente descripción de la familia de supergenes de la hormona de crecimiento (GH) se proporciona para fines ilustrativos. Además, referencia a polipéptidos de GH en la presente solicitud pretende usar el término genérico como un ejemplo de cualquier miembro de la familia de supergenes de GH. Así, se entiende que las modificaciones y químicas descritas en el presente documento con referencia a polipéptidos o proteína de GH pueden ser igualmente aplicadas a cualquier miembro de la familia de supergenes de GH, que incluye aquellos específicamente enumerados en el presente documento.

Las siguientes proteínas incluyen aquellas codificadas por genes de la familia de supergenes de la hormona de crecimiento (GH) (Bazan, F., Immunology Today 11: 350-354 (1990); Bazan, J. F. Science 257: 410-413 (1992); Mott, H. R. y Campbell, I. D., Current Opinion in Structural Biology 5: 114-121 (1995); Silvennoinen, O. y Ihle, J. N., SIGNALLING BY THE HEMATOPOIETIC CYTOKINE RECEPTORS (1996)): hormona de crecimiento, prolactina, lactógeno placentario, eritropoyetina (EPO), trombopoyetina (TPO), interleucina-2 (IL-2), IL-3, IL-4, IL-5, IL-6, IL-7, IL-9, IL-10, IL-11, IL-12 (subunidad p35), IL-13, IL-15, oncostatina M, factor neurotrófico ciliar, factor inhibidor de la leucemia, interferón alfa, interferón beta, interferón épsilon, interferón gamma, interferón omega, interferón tau, factor estimulante de colonias de granulocitos (G-CSF), factor estimulante de colonias de granulocitos-macrófagos (GM-CSF), factor estimulante de colonias de macrófagos (M-CSF) y cardiotrofina-1 (CT-1) ("familia de supergenes de GH"). Se tiene previsto que miembros adicionales de esta familia de genes se identifiquen en el futuro mediante la clonación y secuenciación de genes. Miembros de la superfamilia de genes de GH tienen estructuras secundarias y terciarias similares, a pesar del hecho de que generalmente tienen identidad limitada de secuencia de aminoácidos o

de ADN. Las características estructurales compartidas permiten identificar fácilmente nuevos miembros de la familia de genes y aplicar similarmente los procedimientos y composiciones de aminoácidos no naturales descritos en el presente documento.

Estructuras de varias citocinas, que incluyen G-CSF (Zink y col., FEBS Lett. 314:435 (1992); Zink y col., Biochemistry 33:8453 (1994); Hill y col., Proc.Natl.Acad.Sci.USA 90:5167 (1993)), GM-CSF (Diederichs, K., y col. Science 154: 1779-1782 (1991); Walteran y col., J. Mol. Biol. 224:1075-1085 (1992)), IL-2 (Bazan, J. F. y McKay, D. B. Science 257: 410-413 (1992)), IL-4 (Redfield y col., Biochemistry 30: 11029-11035 (1991); Powers y col., Science 256:1673-1677 (1992)) e IL-5 (Milburn y col., Nature 363: 172-176 (1993)) han sido determinadas por estudios de difracción de rayos X y de RMN y muestran una notable conservación con la estructura de GH, a pesar de una falta de secuencias de homología primaria significativas. Se considera que IFN es un miembro de esta familia basándose en modelado y otros estudios (Lee y col., J. Interferon Cytokine Res. 15:341 (1995); Murgolo y col., Proteins 17:62 (1993); Radhakrishnan y col., Structure 4:1453 (1996); Klaus y col., J. Mol. Biol. 274:661 (1997)). Un gran número de citocinas y factores de crecimiento adicionales que incluyen factor neurotrófico ciliar (CNTF), factor inhibidor de la leucemia (LIF), trombopoyetina (TPO), oncostatina M, factor estimulante de colonias de macrófagos (M-CSF), IL-3, IL-6, IL-7, IL-9, IL-12, IL-13, IL-15 y factor estimulante de colonias de granulocitos (G-CSF), además de los IFN tales como interferón alfa, beta, omega, tau, épsilon y gamma, pertenecen a esta familia (revisado en Mott y Campbell, Current Opinion in Structural Biology 5: 114-121 (1995); Silvennoinen y Ihle (1996) SIGNALLING BY THE HEMATOPOIETIC CYTOKINE RECEPTORS). Todas las citocinas y factores de crecimiento anteriores se consideran ahora que comprenden una gran familia de genes.

Además de compartir estructuras secundarias y terciarias similares, los miembros de esta familia comparten la propiedad de que deben oligomerizar receptores de la superficie celular para activar rutas de señalización intracelulares. Algunos miembros de la familia de GH, que incluyen, pero no se limitan a, GH y EPO, se unen a un único tipo de receptor y hacen que forme homodímeros. Otros miembros de la familia, que incluyen, pero no se limitan a, IL-2, IL-4 e IL-6, se unen a más de un tipo de receptor y hacen que los receptores formen heterodímeros o agregados de mayor orden (Davis y col., (1993) Science 260: 1805-1808; Paonessa y col., 1995) EMBO J. 14: 1942-1951; Mott y Campbell, Current Opinion in Structural Biology 5: 114-121 (1995)). Estudios de mutagénesis han mostrado que, al igual que GH, estas otras citocinas y factores de crecimiento contienen múltiples sitios de unión al receptor, normalmente dos, y se unen a sus receptores relacionados secuencialmente (Mott y Campbell, Current Opinion in Structural Biology 5: 114-121 (1995); Matthews y col., (1996) Proc. Natl. Acad. Sci. USA 93: 9471-9476). Al igual que GH, los sitios de unión al receptor primario para estos otros miembros de la familia se producen principalmente en las cuatro hélices alfa y el bucle A-B. Los aminoácidos específicos en los haces helicoidales que participan en la unión al receptor se diferencian entre los miembros de la familia. La mayoría de los receptores de la superficie celular que interaccionan con miembros de la superfamilia de genes de GH están estructuralmente relacionados y comprenden una segunda familia de multi-genes grande. Véase, por ejemplo, la patente de EE.UU. nº 6.608.183.

Una conclusión general obtenida de estudios de mutaciones de diversos miembros de la superfamilia de genes de GH es que los bucles que se unen a las hélices alfa generalmente tienden a no participar en la unión al receptor. En particular, el bucle B-C corto parece ser no esencial para la unión al receptor en la mayoría de, si no en todos, los miembros de la familia. Por este motivo, el bucle B-C puede estar sustituido con aminoácidos no naturales como se describe en el presente documento en miembros de la familia de supergenes de GH. El bucle A-B, el bucle C-D (y el bucle D-E de miembros similares a interferón/ IL-10 de la superfamilia de GH) también pueden estar sustituidos con un aminoácido no natural. Aminoácidos proximales a la hélice A y distales a la hélice final también tendieron a no participar en la unión al receptor y también pueden ser sitios para introducir aminoácidos no naturales. En algunas realizaciones, un aminoácido no natural está sustituido en cualquier posición dentro de una estructura de bucle que incluye, pero no se limita a, los 1, 2, 3, 4, 5, 6, 7 primeros o más aminoácidos del bucle A-B, B-C, C-D o D-E. En algunas realizaciones, un aminoácido no natural está sustituido dentro de los 1, 2, 3, 4, 5, 6, 7 últimos o más aminoácidos del bucle A-B, B-C, C-D o D-E.

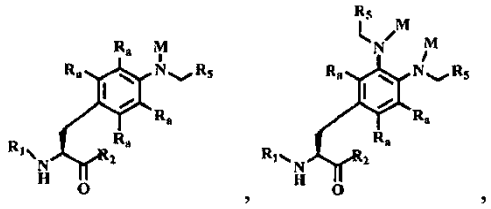
Ciertos miembros de la familia de GH, que incluyen, pero no se limitan a, EPO, IL-2, IL-3, IL-4, IL-6, GM-CSF, TPO, IL-10, IL-12 p35, IL-13, IL-15 e interferón beta contienen azúcares ligados a N y/o ligados a O. Los sitios de glucosilación en las proteínas se producen casi exclusivamente en las regiones de bucle y no en los haces helicoidales alfa. Debido a que las regiones de bucle generalmente no participan en la unión al receptor y debido a que son sitios para la unión covalente de grupos azúcar, pueden ser sitios útiles para introducir sustituciones no naturales de aminoácidos en las proteínas. Los aminoácidos que comprenden los sitios de glucosilación ligados en N y O en las proteínas pueden ser sitios para sustituciones de aminoácidos no naturales debido a que estos aminoácidos están expuestos a la superficie. Por tanto, la proteína natural puede tolerar grupos de azúcar voluminosos unidos a las proteínas en estos sitios y los sitios de glucosilación tienden a localizarse lejos de los sitios de unión al receptor.

Es probable que miembros adicionales de la familia de genes de GH sean descubiertos en el futuro. Pueden identificarse nuevos miembros de la superfamilia de genes de GH mediante análisis de estructuras secundarias y terciarias asistidos por ordenador de las secuencias de proteínas predichas, y por técnicas de selección diseñadas para identificar moléculas que se unen a una diana particular. Los miembros de la superfamilia de genes de GH

normalmente poseen cuatro o cinco hélices anfipáticas unidas por aminoácidos no helicoidales (las regiones de bucle). Las proteínas pueden contener una secuencia señal hidrófoba en su extremo N para promover la secreción de la célula. Tales miembros posteriormente descubiertos de la superfamilia de genes de GH también están incluidos dentro de los procedimientos y composiciones descritos en el presente documento.

V. Aminoácidos no naturales

En un aspecto, la presente invención proporciona un compuesto o sal del mismo seleccionado del grupo que consiste en:



en el que cada R_a está seleccionado independientemente del grupo que consiste en H, halógeno, alquilo, $-NO_2$, -CN, alquilo sustituido, $-N(R')_2$, $-C(O)_kR'$, $-C(O)N(R')_2$, -OR' y $-S(O)_kR'$, en las que k es 1, 2 ó 3;

M es H o $-CH_2R_5$;

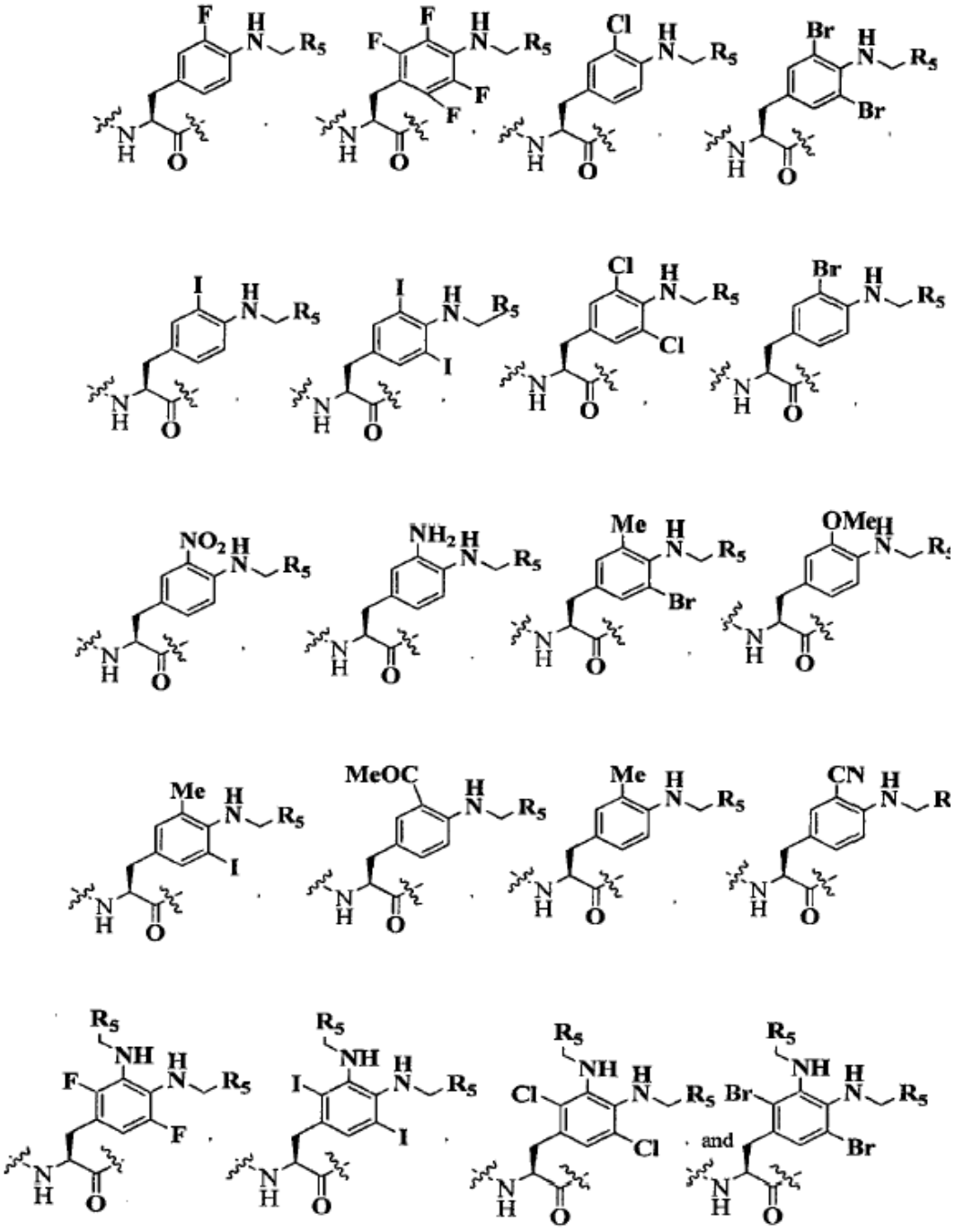
R_1 es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido;

R_2 es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;

R_5 es L-X, en la que X es un polímero que comprende alquilo, alquenilo, alquinilo, alcoxi, alquilalcoxi, poli(óxido de alquileno), arilo, heteroarilo, alcarilo o aralquilo; y L es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno, alquenileno, -O-, -O-(alquilen)-, -(alquilen)-O-, -C(O)-, -C(O)-(alquilen)-, -(alquilen)-C(O)-, -C(O)N(R')-, -C(O)N(R')-(alquilen)-, -(alquilen)-C(O)N(R'), -OC(O)N(R')-, -OC(O)N(R')-(alquilen)-, -(alquilen)-OC(O)N(R')-, -N(R')C(O)-, -NR'C(O)-(alquilen)-, -(alquilen)-NR'C(O)-, -S-, -S-(alquilen)-, $-S(O)_k$- en la que k es 1, 2 ó 3, $-S(O)_k$(alquilen)-, -C(S)-, -C(S)-(alquilen)-, -CSN(R')-, -CSN(R')-(alquilen)-, -N(R')CO-(alquilen)-, -N(R')C(O)O-, -(alquilen)-O-N=CR'-, -(alquilen)-C(O)NR'-(alquilen)-, -(alquilen)-$S(O)_k$-(alquilen o alquilen sustituido)-S-, -(alquilen)-S-S-, $-S(O)_kN(R')$-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, $-N(R')S(O)_kN(R')$-, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, $-C(R')_2$-N=N- y $-C(R')_2$-N(R')-N(R')-; y

cada R' es independientemente H, alquilo o alquilo sustituido.

en el que el compuesto está seleccionado del grupo:



Los restos químicos mediante aminoácidos no naturales que pueden incorporarse en proteínas ofrecen una variedad de ventajas y manipulaciones de la proteína. Por ejemplo, un aminoácido no natural de átomo pesado, por ejemplo, puede ser útil para poner en fase datos de estructuras de rayos X. La introducción específica del sitio de átomos pesados que usan aminoácidos no naturales también proporciona selectividad y flexibilidad en la elección de posiciones para átomos pesados. Aminoácidos no naturales fotorreactivos (que incluyen aminoácidos con benzofenona y cadenas laterales de arilazidas (que incluyen, pero no se limitan a, fenilazida)), por ejemplo, permiten la eficaz fotorreticulación *in vivo* e *in vitro* de proteína. Ejemplos de aminoácidos no naturales fotorreactivos incluyen p-azido-fenilalanina y p-benzoil-fenilalanina. La proteína con los aminoácidos no naturales fotorreactivos pueden entonces reticularse a voluntad por excitación del control temporal que proporciona grupos fotorreactivos. En un ejemplo, el grupo metilo de un aminoácido no natural puede sustituirse con un grupo isotópicamente marcado, que incluye metilo, como sonda de estructura y dinámica local, incluyendo el uso de resonancia magnética nuclear y espectroscopía vibracional.

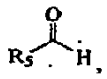
Otros ejemplos de aminoácidos no naturales que contienen un resto de amina aromática se muestran en las FIG. 2-10. Tales aminoácidos no naturales también pueden estar en forma de una sal, o pueden incorporarse en un polipéptido de aminoácidos no naturales, polímero, polisacárido, o un polinucleótido.

Ejemplos de síntesis de compuestos se muestran en las FIG. 11-13. Tales ejemplos muestran las síntesis de aminoácidos no naturales que contienen restos de amina aromática, en los que el resto de amina creado es una amina primaria o secundaria. Tales aminoácidos no naturales pueden estar en forma de una sal, o pueden incorporarse en un polipéptido de aminoácidos no naturales, polímero, polisacárido, o un polinucleótido y opcionalmente alquilarse reductoramente.

En la FIG. 14 se muestra una formación de un aminoácido no natural que contiene amina aromática, en la que, a modo de ejemplo solo, aminoácidos no naturales pueden formarse mediante reducción de restos de amina protegida o enmascarada sobre el resto aromático de un aminoácido no natural. Tales restos de amina protegida o enmascarada incluyen, pero no se limitan a, iminas, hidracinas, nitro, o sustituyentes de azida. Los agentes reductores usados para reducir tales restos de amina protegidos o enmascarados incluyen, pero no se limitan a TCEP, Na_2S, $Na_2S_2O_4$, $LiAlH_4$, $NaBH_4$ o $NaBCNH_3$.

Otros ejemplos de tales aminoácidos no naturales que contienen restos de amina protegidos o enmascarados se muestran en la FIG. 15

Los compuestos de la invención puede formarse por la alquilación reductora de una arilamina con reactivos que contienen carbonilo tales como, a modo de ejemplo, cetonas, ésteres, tioésteres y aldehídos. A modo de ejemplo, tales reactivos que contienen aldehído pueden tener la estructura correspondiente a

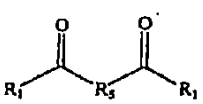


en la que R_5 es como se ha definido previamente.

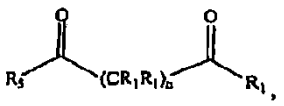
Síntesis a modo de ejemplo de polipéptidos de aminoácidos no naturales que contienen aminoácidos se presentan en la FIG. 15, en la que los restos de amina enmascarada de aminoácidos no naturales contenidos en polipéptidos se reducen inicialmente dando aminoácidos no naturales que contienen restos de amina aromática incorporados en polipéptidos de aminoácidos no naturales. Tales restos de amina aromática se alquilan a continuación de forma reductora con reactivos que contienen carbonilo descritos anteriormente dando polipéptidos que contienen aminoácidos no naturales. Tales reacciones también pueden aplicarse a aminoácidos no naturales incorporados en polímeros sintéticos, polisacáridos o polinucleótidos. Adicionalmente, tales reacciones pueden aplicarse a aminoácidos no naturales no incorporados. A modo de ejemplo, el agente reductor usado para reducir restos de amina enmascarada incluye, pero no se limita a, TCEP, Na_2S, $Na_2S_2O_4$, $LiAlH_4$, B_2H_6 y $NaBH_4$. A modo de ejemplo solo, la alquilación reductora puede producirse en tampones acuosos con un pH de aproximadamente 4 a aproximadamente 7 y usando un agente reductor suave, tal como, a modo de ejemplo solo, cianoborohidruro de sodio ($NaBCNH_3$). Además, pueden usarse otros agentes reductores para la alquilación reductora que incluyen, pero no se limitan a, TCEP, Na_2S, $Na_2S_2O_4$, $LiAlH_4$, B_2H_6 y $NaBH_4$.

En la FIG. 16 se presentan síntesis a modo de ejemplo no limitantes de polipéptidos de aminoácidos no naturales que contienen aminoácidos por alquilación reductora de restos de amina secundaria aromática, contenidos en aminoácidos no naturales, con reactivos que contienen carbonilo descritos anteriormente. Tales alquilaciones reductoras dan polipéptidos que contienen aminoácidos no naturales con restos de amina aromática terciaria. Tales reacciones también pueden aplicarse a aminoácidos no naturales incorporados en polímeros sintéticos, polisacáridos o polinucleótidos. Adicionalmente, tales reacciones pueden aplicarse a aminoácidos no naturales no incorporados. A modo de ejemplo solo, la alquilación reductora puede producirse en tampones acuosos con un pH de aproximadamente 4 a aproximadamente 7 y usando un agente reductor suave, tal como, a modo de ejemplo solo, cianoborohidruro de sodio ($NaBCNH_3$). Además, pueden usarse otros agentes reductores para la alquilación reductora que incluyen, pero no se limitan a, TCEP, Na_2S, $Na_2S_2O_4$, $LiAlH_4$, B_2H_6 y $NaBH_4$.

Los compuestos descritos en el presente documento también pueden formarse por la alquilación reductora de compuestos con reactivos que contienen al menos dos restos carbonilo, que incluyen, pero no se limitan a, dicetonas, cetoaldehídos y dialdehídos. A modo de ejemplo, tales reactivos pueden tener la estructura correspondiente a



o



en la que R_1 y R_5 son como se han definido previamente.

En la FIG. 19 se muestra una representación no limitante de la formación de aminoácidos no naturales descritos en el presente documento por alquilación reductora del resto de amina aromática de aminoácidos no naturales descritos en el presente documento, y una representación no limitante de la formación de aminoácidos no naturales descritos en el presente documento por aminación reductora de un resto de aldehído desprotegido sobre aminoácidos no naturales descritos en el presente documento. En la FIG. 19 se representa la especificidad y selectividad de la reacción de alquilación reductora, en la que solo el resto de amina aromática del aminoácido no natural se alquila reductoramente mientras que otros restos de amina, tioles y enlaces disulfuro no se reducen o no reaccionan bajo las condiciones de reacción usadas. También se representa en la FIG. 19 la especificidad y selectividad de la reacción de aminación reductora, en la que solo el resto de aldehído desprotegido de aminoácido no natural se amina reductoramente mientras que otros restos de carboxilato, tioles y enlaces disulfuro no se reducen o no reaccionan bajo las condiciones de reacción usadas. Tales derivatizaciones específicas del sitio selectivas, reacciones de alquilación reductora o de aminación reductora pueden permitir la modificación de polipéptidos/proteínas para diseñar agonistas y/o antagonistas, PEGilación específica del sitio de polipéptidos/proteínas, conjugación específica del sitio de polipéptidos/proteínas, diseño de profármacos, glucosilación de polipéptidos/proteínas, dimerización de polipéptidos/proteínas, conjugados de fármacos de molécula pequeña de polipéptidos/proteínas y conjugados de fármacos de molécula pequeña de anticuerpos.

La modificación de aminoácidos no naturales descritos en el presente documento usando reacciones de alquilación reductora o de aminación reductora tiene cualquiera o todas de las siguientes ventajas. Primera, las aminas aromáticas pueden alquilarse reductoramente con compuestos que contienen carbonilo, que incluyen aldehídos y cetonas, en un intervalo de pH de aproximadamente 4 a aproximadamente 10 (y en ciertas realizaciones en un intervalo de pH de aproximadamente 4 a aproximadamente 7) para generar enlaces de amina sustituida, que incluye amina secundaria y terciaria. Segunda, bajo estas condiciones de reacción, la química es selectiva para aminoácidos no naturales, ya que las cadenas laterales de aminoácidos que se producen naturalmente son no reactivas. Esto permite la derivatización específica del sitio de polipéptidos que han incorporado aminoácidos no naturales que contienen restos de amina aromática o restos de aldehído protegidos, que incluyen, a modo de ejemplo, proteínas recombinantes. Tales polipéptidos y proteínas derivatizados pueden así prepararse como productos homogéneos definidos. Tercera, las condiciones suaves necesarias para efectuar la reacción de un resto de amina aromática sobre un aminoácido, que se ha incorporado en un polipéptido, con un reactivo que contiene aldehído generalmente no destruyen irreversiblemente la estructura terciaria del polipéptido (exceptuando, por supuesto, si el fin de la reacción es destruir tal estructura terciaria). Similarmente, las condiciones suaves necesarias para efectuar la reacción de un resto de aldehído sobre un aminoácido, que se ha incorporado en un polipéptido y desprotegido, con un reactivo que contiene amina aromática, generalmente no destruyen irreversiblemente la estructura terciaria del polipéptido (exceptuando, por supuesto, si el fin de la reacción es destruir tal estructura terciaria). Cuarta, la reacción se produce rápidamente a temperatura ambiente, que permite el uso de muchos tipos de polipéptidos o reactivos que de otro modo serían inestables a mayores temperaturas. Quinta, la reacción se produce fácilmente a condiciones acuosas, permitiendo de nuevo el uso de polipéptidos y reactivos incompatibles (a cualquier grado) con disoluciones no acuosas. Sexta, la reacción se produce fácilmente incluso cuando la relación de polipéptido o aminoácido con respecto a reactivo sea estequiométrica, similar a estequiométrica o casi estequiométrica, de manera que es innecesario añadir exceso de reactivo o polipéptido para obtener una cantidad útil de producto de reacción. Séptima, la amina resultante puede producirse regioselectivamente y/o regioespecíficamente, dependiendo del diseño de la amina y porciones de carbonilo de los reactantes. Finalmente, la alquilación reductora de aminas aromáticas con reactivos que contienen aldehído, y la aminación reductora de aldehídos con reactivos que contienen aminas aromáticas, genera enlaces de amina, que incluyen amina secundaria y terciaria, que son estables en condiciones biológicas.

En la FIG. 20 se muestran ejemplos de la especificidad y selectividad de la alquilación reductora de restos de amina aromática sobre residuos de aminoácidos próximos a un residuo de cisteína. La alquilación reductora de urotensina-II (UT-II-SH) reducida con propionaldehído (I) o benzaldehído (II) demuestra que predominantemente se forma un producto a partir de las reacciones respectivas, y que el grupo tiol libre no reaccionó. También se muestran en la FIG. 20 la velocidad de las reacciones, que fue aproximadamente 2 horas, y la relación estequiométrica o casi estequiométrica entre la urotensina-II (UT-II-SH) reducida y los reactivos que contienen aldehído respectivos. Adicionalmente, se presentan datos de EM para UT-II-SH, I y II.

En la FIG. 21 se muestran ejemplos de la especificidad y selectividad de la alquilación reductora de restos de amina aromática sobre residuos de aminoácidos próximos a un residuo de cisteína. La alquilación reductora de urotensina-II (UT-II) con propionaldehído (I) o benzaldehído (II) demuestra que se formó predominantemente un producto a partir de las reacciones respectivas, y que no reaccionó el enlace disulfuro. También se muestran en la FIG. 21 la velocidad de las reacciones, que fue aproximadamente 2 horas, y la relación estequiométrica o casi estequiométrica

entre la urotensina-II (UT-II) y los reactivos que contienen aldehído respectivos. Adicionalmente, se presentan datos de EM para UT-II-SH, I y II.

Otros ejemplos de la especificidad y selectividad de las reacciones de alquilación reductora de restos de amina aromática sobre residuos de aminoácidos terminales se muestran en la FIG. 22, en la que la alquilación reductora del péptido XT-8 con propionaldehído (I), benzaldehído (II), isobutaldehído (III) y pivalaldehído (IV) demuestra que se formó predominantemente un producto a partir de las reacciones respectivas, y que no reaccionó el enlace disulfuro. También se muestran en la FIG. 22 la velocidad de las reacciones, que fue aproximadamente 2 horas, y la relación estequiométrica o casi estequiométrica entre XT-8 y los reactivos que contienen aldehído respectivos. Adicionalmente, se presentan los datos de EM para XT-8, I, II, III y IV.

Otros ejemplos de la especificidad y selectividad de las reacciones de alquilación reductora de restos de amina aromática sobre residuos de aminoácidos que están próximos a diversos aminoácidos del extremo N se muestran en las FIG. 23-27. En la FIG. 23, la alquilación reductora del péptido SXT-9 (serina del extremo N) con propionaldehído (I), benzaldehído (II), isobutaldehído (III) y pivalaldehído (IV) muestra la velocidad de las reacciones, que fue aproximadamente 2 horas, y la relación estequiométrica o casi estequiométrica entre SXT-9 y los reactivos que contienen aldehído respectivos. Además, la FIG. 23 demuestra que se formó predominantemente un producto a partir de las reacciones respectivas, que no reaccionó el enlace disulfuro, y que el residuo de Ser de N tiene efecto mínimo sobre la selectividad. Adicionalmente, se presentan los datos de EM para SXT-9, I, II, III y IV.

En la Fig. 24, la alquilación reductora del péptido HXT-9 (histidina del extremo N) con propionaldehído (I), benzaldehído (II), isobutaldehído (III) y pivalaldehído (IV) muestra la velocidad de las reacciones, que fue aproximadamente 2 horas, y la relación estequiométrica o casi estequiométrica entre HXT-9 y los reactivos que contienen aldehído respectivos. Además, la FIG. 24 demuestra que se formó predominantemente un producto a partir de las reacciones respectivas, que no reaccionó el enlace disulfuro, y que el residuo de His de N tiene efecto mínimo sobre la selectividad. Por tanto, se muestran datos de EM para HXT-9, I, II, III y IV.

En la Fig. 25, la alquilación reductora del péptido WXT-9 (triptófano del extremo N) con propionaldehído (I), benzaldehído (II), isobutaldehído (III) y pivalaldehído (IV) muestra la velocidad de las reacciones, que fue aproximadamente 2 horas, y la relación estequiométrica o casi estequiométrica entre WXT-9 y los reactivos que contienen aldehído respectivos. Además, la FIG. 25 demuestra que se formó predominantemente un producto a partir de las reacciones respectivas, que no reaccionó el enlace disulfuro, y que el residuo de Trp de N tiene efecto mínimo sobre la selectividad. Por tanto, se muestran datos de EM para WXT-9, I, II y III.

En la Fig. 26, la alquilación reductora del péptido NXT-9 (asparagina del extremo N) con propionaldehído (I), benzaldehído (II), isobutaldehído (III) y pivalaldehído (IV) muestra la velocidad de las reacciones, que fue aproximadamente 2 horas, y la relación estequiométrica o casi estequiométrica entre NXT-9 y los reactivos que contienen aldehído respectivos. Además, la FIG. 26 demuestra que se formó predominantemente un producto a partir de las reacciones respectivas, que no reaccionó el enlace disulfuro, y que el residuo de Asn de N tiene efecto mínimo sobre la selectividad. Por tanto, se muestran datos de EM para NXT-9, I y II.

En la Fig. 27, la alquilación reductora del péptido RXT-10 (arginina del extremo N) con propionaldehído (I) y benzaldehído (II) muestra la velocidad de las reacciones, que fue aproximadamente 2 horas, y la relación estequiométrica o casi estequiométrica entre RXT-10 y los reactivos que contienen aldehído respectivos. Además, la FIG. 27 demuestra que se formó predominantemente un producto a partir de las reacciones respectivas, que no reaccionó el enlace disulfuro, y que el residuo de Arg de N tiene efecto mínimo sobre la selectividad. Por tanto, se muestran datos de EM para RXT-10, I y II.

Otro ejemplo de la especificidad y selectividad de las reacciones de alquilación reductora de restos de amina aromática sobre residuos de aminoácidos en un péptido se muestra en las FIG. 28, en las que la alquilación reductora del péptido AXT-11 con propionaldehído (I) y benzaldehído (II) muestra la velocidad de las reacciones, que fue aproximadamente 2 horas, y la relación estequiométrica o casi estequiométrica entre AXT-11 y los reactivos que contienen aldehído respectivos. Además, la FIG. 28 demuestra que se formó predominantemente un producto a partir de las reacciones respectivas, que no reaccionó el enlace disulfuro, y que el residuo de Ala de N tiene efecto mínimo sobre la selectividad. También se presentan datos de EM para AXT-11, I y II.

Otros ejemplos de la especificidad y selectividad de las reacciones de alquilación reductora de restos de amina aromática sobre residuos de aminoácidos en un péptido con diversos reactivos que contienen aldehído se muestran en las FIG. 29-30. En la FIG. 29, la alquilación reductora del péptido AXT-11 con diversos reactivos que contienen aldehído muestra la velocidad de las reacciones, que fue aproximadamente 5 horas, y la relación estequiométrica o casi estequiométrica entre AXT-11 y los reactivos que contienen aldehído respectivos. Además, la FIG. 29 demuestra que se formó predominantemente un producto a partir de las reacciones respectivas, y que no reaccionó el enlace disulfuro. Por tanto, se presentan datos de EM para AXT-11, I, II y III. En la FIG. 30, la alquilación reductora del péptido AXT-11 con diversos reactivos que contienen aldehído heteroaromático muestra la velocidad de las reacciones, que fue aproximadamente 5 horas, y la relación estequiométrica o casi estequiométrica entre AXT-11 y los reactivos que contienen aldehído respectivos. Además, la FIG. 30 demuestra que se formó predominantemente

un producto a partir de las reacciones respectivas, y que no reaccionó el enlace disulfuro. Por tanto, se presentan datos de EM para AXT-11, IV, V, VI.

Las FIG. 31-33 muestran ejemplos de la especificidad y selectividad de las reacciones de alquilación reductora de restos de amina aromática sobre residuos de aminoácidos en un péptido, en las que los restos de amina aromática se hacen reaccionar con tanto un reactivo que contiene aldehído, un reactivo que contiene cetona, como una mezcla de los mismos. En la FIG. 31, la alquilación reductora del péptido AXT-11 con un reactivo que contiene aldehído, un reactivo que contiene cetona, o una mezcla de los mismos, muestra la velocidad de las reacciones, que fue aproximadamente 2 horas, y la relación estequiométrica o casi estequiométrica entre AXT-11 y los reactivos que contienen aldehído respectivos. Además, la FIG. 31 demuestra que bajo las condiciones de reacción usadas, solo el reactivo que contiene aldehído reaccionó con el resto de amina aromática y solo se formó predominantemente un producto. De nuevo, no reaccionó el enlace disulfuro. En la FIG. 32, la alquilación reductora del péptido NXT-9 con un reactivo que contiene aldehído, un reactivo que contiene cetona, o una mezcla de los mismos, muestra la velocidad de las reacciones, que fue aproximadamente 2 horas, y la relación estequiométrica o casi estequiométrica entre NXT-10 y los reactivos que contienen aldehído respectivos. Además, la FIG. 32 demuestra que bajo las condiciones de reacción usadas, solo el reactivo que contiene aldehído reaccionó con el resto de amina aromática y solo se formó predominantemente un producto. De nuevo, no reaccionó el enlace disulfuro. En la FIG. 33, la alquilación reductora del péptido MXT-9 con un reactivo que contiene aldehído, un reactivo que contiene cetona, o una mezcla de los mismos, demuestra que bajo las condiciones de reacción usadas, solo el reactivo que contiene aldehído reaccionó con el resto de amina aromática y solo se formó predominantemente un producto. De nuevo, no reaccionó el enlace disulfuro.

Un ejemplo de la desprotección (o desenmascaramiento) de un resto de amina aromática protegida (o enmascarada) sobre un aminoácido en un péptido, seguido de alquilación reductora del resto de amina aromática resultante con diversos reactivos que contienen aldehído, se muestra en la FIG. 34. La reducción de la azida sobre el péptido MXT-9-N3 con TCEP dio el resto de amina primaria aromática sobre el péptido MXT-9NH2 (I). Las posteriores alquilaciones reductoras de MXT-9NH2 (I) con propionaldehído (II) o benzaldehído (III) dan a continuación los péptidos alquilados correspondientes. La FIG. 34 demuestra la capacidad para formar restos de amina aromática después de la incorporación de un aminoácido no natural que contiene un grupo de amina protegida o enmascarada. También se muestran la velocidad de las reacciones, que fue aproximadamente 1 hora, la relación estequiométrica o casi estequiométrica entre MXT-9NH2 (I) y los reactivos que contienen aldehído respectivos, solo se formó predominantemente un producto, y no reaccionaron los grupos tiol.

D. Captación celular de aminoácidos no naturales

La captación de aminoácidos no naturales por una célula es una cuestión que se considera cuando se diseñan y seleccionan aminoácidos no naturales, incluyendo para la incorporación en una proteína. Por ejemplo, la alta densidad de carga de los α-aminoácidos sugiere que es poco probable que estos compuestos sean permeables a la célula. Aminoácidos naturales son capturados en la célula eucariota mediante un conjunto de sistemas de transporte basados en proteínas. Puede hacerse una selección rápida que evalúa qué aminoácidos no naturales, si los hay, son capturados por las células. Véanse, por ejemplo, los ensayos de toxicidad en, por ejemplo, la publicación de patente de EE.UU. nº 2004/198637 titulada “Protein Arrays”; y Liu, D.R. & Schultz, P.G. (1999) Progress toward the evolution of an organism with an expanded genetic code. PNAS United States 96:4780-4785. Aunque la captación se analiza fácilmente con diversos ensayos, una alternativa a diseñar aminoácidos no naturales que son susceptibles a rutas de captación celular es proporcionar rutas biosintéticas para crear aminoácidos *in vivo.*

E. Biosíntesis de aminoácidos no naturales

Ya existen muchas rutas biosintéticas en células para la producción de aminoácidos y otros compuestos. Aunque un proceso biosintético para un aminoácido no natural particular puede no existir en la naturaleza, que incluye, pero no se limita a, en una célula, los procedimientos y composiciones descritos en el presente documento proporcionan tales procedimientos. Por ejemplo, rutas biosintéticas para aminoácidos no naturales se generan opcionalmente en célula huésped añadiendo nuevas enzimas o modificando rutas de células huésped existentes. Nuevas enzimas adicionales son opcionalmente enzimas que se producen naturalmente o enzimas artificialmente desarrolladas. Por ejemplo, la biosíntesis de *p*-aminofenilalanina (como se presenta en un ejemplo en el documento WO 2002/085923 titulado “In vivo incorporation of unnatural amino acids”) se basa en la adición de una combinación de enzimas conocidas de otros organismos. Los genes para estas enzimas pueden introducirse en una célula eucariota transformando la célula con un plásmido que comprende los genes. Los genes, cuando se expresan en la célula, proporcionan una ruta enzimática para sintetizar el compuesto deseado. Ejemplos de los tipos de enzimas que son opcionalmente añadidos se proporcionan en los ejemplos más adelante. Se encuentran secuencias de enzimas adicionales, por ejemplo, en Genbank. Enzimas artificialmente desarrolladas también se añaden opcionalmente en una célula del mismo modo. De este modo, la maquinaria celular y los recursos de una célula se manipulan para producir aminoácidos no naturales.

Están disponibles una variedad de procedimientos de producción de enzimas novedosas para su uso en rutas biosintéticas o para la evolución de rutas existentes. Por ejemplo, la recombinación recurrente, que incluye como se

ha desarrollado por Maxygen, Inc. (disponible en la malla mundial en www.maxygen.com), se usa opcionalmente para desarrollar enzimas y rutas novedosas. Véanse, por ejemplo, Stemmer (1994), Rapid evolution of a protein in vitro by DNA shuffling, Nature 370(4):389-391; y Stemmer, (1994), DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution, Proc. Natl. Acad. Sci. USA., 91:10747-10751. Similarmente DesignPath™, desarrollado por Genencor (disponible en la malla mundial en genencor.com) se usa opcionalmente para la ingeniería de rutas metabólicas, que incluye manipular una ruta para crear O-metil-L-tirosina en una célula. Esta tecnología reconstruye rutas existentes en organismos huésped usando una combinación de nuevos genes, que incluyen, pero no se limita a, aquellos identificados mediante genómica funcional, y evolución y diseño molecular. Diversa Corporation (disponible en la malla mundial en diversa.com) también proporciona tecnología para cribar rápidamente bibliotecas de genes y rutas de genes, que incluye crear nuevas rutas.

Normalmente, el aminoácido no natural producido con una ruta biosintética manipulada como se describe en el presente documento se produce en una concentración suficiente para la eficaz biosíntesis de proteínas, que incluye una cantidad celular natural, pero no a un grado tal como para afectar la concentración de los otros aminoácidos o agotar recursos celulares. Concentraciones típicas producidas *in vivo* de este modo son aproximadamente 10 mM a aproximadamente 0,05 mM. Una vez se transforma una célula con un plásmido que comprende los genes usados para producir enzimas deseadas para una ruta específica y se genera un aminoácido no natural, opcionalmente se usan selecciones *in vivo* para optimizar adicionalmente la producción del aminoácido no natural para tanto la síntesis de proteínas ribosómicas como el crecimiento celular.

F. Metodología sintética adicional

Los aminoácidos no naturales descritos en el presente documento pueden sintetizarse usando metodologías descritas en la materia o usando las técnicas descritas en el presente documento o por una combinación de las mismas. Como ayuda, la siguiente tabla proporciona diversos electrófilos y nucleófilos de partida que pueden combinarse para crear un grupo funcional deseado.

Tabla: Ejemplos de enlaces covalentes y precursores de los mismos

Producto Enlace Covalente	Electrófilo	Nucleófilo
Carboxamidas	Ésteres activados	Aminas / anilinas
Carboxamidas	Acidas de acilo	Aminas / anilinas
Carboxamidas	Haluros de acilo	Aminas / anilinas
Ësteres	Haluros de acilo	Alcoholes / Fenoles
Ésteres	Nitrilos de acilo	Alcoholes / Fenoles
Carboxamidas	Nitrilos de acilo	Aminas / anilinas
Iminas	Aldehidos	Aminas / anilinas
Hidrazonas	Aldehídos o cetonas	Hidracinas
Oximas	Aldehídos o cetonas	Hidroxilaminas
Aminas de alquilo	Haluros de alquilo	Aminas / anilinas
Ésteres	Haluros de alquilo	Acidos carboxilicos
Tioéteres	Haluros de alquilo	Tioles
Ëteres	Haluros de alquilo	Alcoholes / Fenoles
Tioéteres	Sulfatos de alquilo	Tioles
Ésteres	Sulfatos de alquilo	Acidos carboxilicos
Ëteres	Sulfatos de alquilo	Alcoholes / Fenoles
Ésteres	Anhídridos	Alcoholes / Fenoles
Carboxamidas	Anhídridos	Aminas / anilinas
Tiofenoles	Haluros de acilo	Tioles
Aminas de arilo	Haluros de acilo	Aminas
Tioeteres	Acidinas	Tioles
Ésteres de boronato	Boronato	Glicoles
Carboxamidas	Acidos carboxilicos	Aminas / anilinas
Ësteres	Acidos carboxilicos	Alcoholes
N-acilureas o anhídridos	Carbodiimidas	Acidos carboxilicos
Ésteres	Diazoalcanos	Acidos carboxilicos
Tioéteres	Epoxidos	Tioles
Tioéteres	Haloacetamidas	Tioles
Ammotriacinas	Halotriacinas	Aminas / anilinas
Éteres de triacinilo	Halotriacinas	Alcoholes / Fenoles
Aminidas	Esteres de imido	Aminas / anilinas
Ureas	Isocianatos	Aminas / anilinas
Uretanos	Isocianatos	Alcoholes / Fenoles
Tioureas	Isotiocianatos	Aminas / anilinas

Tioeteres	Maleimidas	Tioles
Esteres de fosfito	Fosforamiditos	Alcoholes
Eteres de sirilo	Haluros de sirilo	Alcoholes
Aminas de alquilo	Ésteres de sulfonato	Aminas / anilinas
Tioeteres	Ésteres de sulfonato	Tioles
Esteres	Ésteres de sulfonato	Acidos carboxilicos
Eteres	Ésteres de sulfonato	Alcoholes
Sulfonamidas	Haluros de sulfonilo	Aminas / anilinas
Ésteres de sulfonato	Haluros de sulfonilo	Alcoholes / Fenoles

En general, los electrófilos de carbono son susceptibles al ataque por nucleófilos complementarios, que incluyen nucleófilos de carbono, en los que un nucleófilo atacante lleva un par de electrones al electrófilo de carbono con el fin de formar un nuevo enlace entre el nucleófilo y el electrófilo de carbono.

Nucleófilos de carbono adecuados incluyen Grignard de alquilo, alquenilo, arilo y alquinilo, organolitio, organocinc, reactivos de alquil-, alquenil-, aril- y alquinil-estaño (organoestannanos), reactivos de alquil-, alquenil-, aril- y alquinil-borano (organoboranos y organoboronatos); estos nucleófilos de carbono tienen la ventaja de ser cinéticamente estables en agua o disolventes orgánicos polares. Otros nucleófilos de carbono incluyen iluros de fósforo, reactivos de enol y enolato; estos nucleófilos de carbono tienen la ventaja de ser relativamente fáciles de generar a partir de precursores muy conocidos para aquellos expertos en la materia de la química orgánica sintética. Los nucleófilos de carbono, cuando se usan conjuntamente con electrófilos de carbono, engendran nuevos enlaces carbono-carbono entre el nucleófilo de carbono y el electrófilo de carbono.

Nucleófilos de no carbono adecuados para acoplarse a electrófilos de carbono incluyen aminas primarias y secundarias, tioles, tiolatos y tioéteres, alcoholes, alcóxidos, azidas, semicarbazidas, y similares. Estos nucleófilos de no carbono, cuando se usan conjuntamente con electrófilos de carbono, normalmente generan enlaces de heteroátomos (C-X-C), en los que X es un heteroátomo, por ejemplo, oxígeno o nitrógeno.

VI. Polipéptidos con aminoácidos no naturales

Las composiciones y procedimientos descritos en el presente documento proporcionan la incorporación de al menos un aminoácido no natural en un polipéptido. El aminoácido no natural puede estar presente en cualquier localización sobre el polipéptido, que incluye cualquier posición terminal o cualquier posición interna del polipéptido. Preferentemente, el aminoácido no natural no destruye la actividad y/o la estructura terciaria del polipéptido con respecto al polipéptido de aminoácidos que se producen naturalmente homólogo, a menos que tal destrucción de la actividad y/o estructura terciaria sea uno de los fines de incorporar el aminoácido no natural en el polipéptido. Además, la incorporación del aminoácido no natural en el polipéptido puede modificar de algún modo la actividad (por ejemplo, manipular la eficacia terapéutica del polipéptido, mejorar el perfil de seguridad del polipéptido, ajustar la farmacocinética, farmacológica y/o farmacodinámica del polipéptido (por ejemplo, aumentar la solubilidad en agua, biodisponibilidad, aumentar la semivida en suero, aumentar la semivida terapéutica, modular la inmunogenicidad, modular la actividad biológica, o prolongar el tiempo en circulación), proporcionar funcionalidad adicional al polipéptido, incorporar una etiqueta, marca o señal detectable en el polipéptido, facilitar las propiedades de aislamiento del polipéptido, y cualquier combinación de las modificaciones anteriormente mencionadas) y/o estructura terciaria del polipéptido con respecto al polipéptido de aminoácidos que se producen naturalmente homólogo sin causar completamente la destrucción de la estructura terciaria de actividad. Tales modificaciones de la actividad y/o estructura terciaria son frecuentemente uno de los objetivos de efectuar tales incorporaciones, aunque, por supuesto, la incorporación del aminoácido no natural en el polipéptido también puede tener poco efecto sobre la actividad y/o estructura terciaria del polipéptido con respecto al polipéptido de aminoácidos que se producen naturalmente homólogo. Correspondientemente, los polipéptidos de aminoácidos no naturales, composiciones que comprenden polipéptidos de aminoácidos no naturales, procedimientos de preparación de tales polipéptidos y composiciones de polipéptidos, procedimientos de purificación, aislamiento y caracterización de tales polipéptidos y composiciones de polipéptidos, y procedimientos de uso de tales polipéptidos y composiciones de polipéptidos se consideran dentro del alcance de la presente divulgación. Además, los polipéptidos de aminoácidos no naturales descritos en el presente documento también pueden ligarse a otro polipéptido (que incluye, a modo de ejemplo, un polipéptido de aminoácidos no naturales o un polipéptido de aminoácidos que se producen naturalmente).

Los polipéptidos de aminoácidos no naturales descritos en el presente documento pueden producirse biosintéticamente o no biosintéticamente. Por biosintéticamente se indica cualquier procedimiento que utilice un sistema de traducción (celular o no celular), que incluye el uso de al menos uno de los siguientes componentes: un polinucleótido, un codón, un ARNt y un ribosoma. Por no biosintéticamente se indica cualquier procedimiento que no utilice un sistema de traducción: este enfoque puede dividirse adicionalmente en procedimientos que utilizan procedimientos sintéticos de péptidos en estado sólido, procedimientos sintéticos de péptidos en fase sólida, procedimientos que utilizan al menos una enzima y procedimientos que no utilizan al menos una enzima; además, cualquiera de estas sub-divisiones pueden solaparse y muchos procedimientos pueden utilizar una combinación de

estas sub-divisiones.

Los procedimientos, composiciones, estrategias y técnicas descritos en el presente documento no se limitan a un tipo, clase o familia particular de polipéptidos o proteínas. De hecho, prácticamente cualquier polipéptido puede incluir al menos un aminoácido no natural descrito en el presente documento. A modo de ejemplo solo, el polipéptido puede ser homólogo a una proteína terapéutica seleccionada del grupo que consiste en: alfa-1 antitripsina, angiostatina, factor antihemolítico, anticuerpo, apolipoproteína, apoproteína, factor natriurético auricular, polipéptido natriurético auricular, péptido auricular, quimiocina C-X-C, T39765, NAP-2, ENA-78, gro-a, gro-b, gro-c, IP-10, GCP-2, NAP-4, SDF-1, PF4, MIG, calcitonina, ligando c-kit, citocina, quimiocina CC, proteína 1 quimioatrayente de monocitos, proteína 2 quimioatrayente de monocitos, proteína 3 quimioatrayente de monocitos, proteína 1 alfa inflamatoria de monocitos, proteína 1 beta inflamatoria de monocitos, RANTES, 1309, R83915, R91733, HCC1, T58847, D31065, T64262, CD40, ligando CD40, ligando c-kit, colágeno, factor estimulante de colonias (CSF), factor 5a del complemento, inhibidor del complemento, receptor 1 del complemento, citocina, péptido 78 activante de neutrófilos epiteliales, MIP-16, MCP-1, factor de crecimiento epidérmico (EGF), péptido activante de neutrófilos epiteliales, eritropoyetina (EPO), toxina exfoliante, factor IX, factor VII, factor VIII, factor X, factor de crecimiento de fibroblastos (FGF), fibrinógeno, fibronectina, proteína de cuatro haces helicoidales, G-CSF, glp-1, GM-CSF, glucocerebrosidasa, gonadotropina, factor de crecimiento, receptor de factores de crecimiento, grf, proteína hedgehog, hemoglobina, factor de crecimiento de hepatocitos (hGF), hirudina, hormona de crecimiento humana (hGH), albúmina de suero humano, ICAM-1, receptor de ICAM-1, LFA-1, receptor de LFA-1, insulina, factor de crecimiento similar a la insulina (IGF), IGF-I, IGF-II, interferón (IFN), IFN-alfa, IFN-beta, IFN-gamma, cualquier molécula similar a interferón o miembro de la familia de IFN, interleucina (IL), IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, factor de crecimiento de queratinocitos (KGF), lactoferrina, factor inhibidor de la leucemia, luciferasa, neurturina, factor inhibidor de neutrófilos (NIF), oncostatina M, proteína osteogénica, producto oncogénico, paracitonina, hormona paratiroidea, PD-ECSF, PDGF, hormona peptídica, pleiotropina, proteína A, proteína G, pth, endotoxina A pirogénica, exotoxina B pirogénica, exotoxina C pirogénica, pyy, relaxina, renina, SCF, proteína biosintética pequeña, receptor I del complemento soluble, I-CAM 1 soluble, receptor de interleucina soluble, receptor de TNF soluble, somatomedina, somatostatina, somatotropina, estreptocinasa, superantígenos, enterotoxina estafilocócica, SEA, SEB, SEC1, SEC2, SEC3, SED, SEE, receptor de hormonas esteroideas, superóxido dismutasa, toxina del síndrome del choque tóxico, timosina alfa 1, activador tisular del plasminógeno, factor de crecimiento tumoral (TGF), factor de necrosis tumoral, factor de necrosis tumoral alfa, factor de necrosis tumoral beta, receptor del factor de necrosis tumoral (TNFR), proteína VLA-4, proteína VCAM-1, factor de crecimiento endotelial vascular (VEGF), urocinasa, 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 corticosterona. El polipéptido de aminoácidos no naturales también puede ser homólogo a cualquier miembro de polipéptido de la familia de supergenes de la hormona de crecimiento. Ejemplos no limitantes de péptidos/proteínas que pueden incorporar aminoácidos no naturales que contienen un resto de amina aromática se muestran en las FIG. 20-34.

Los polipéptidos de aminoácidos no naturales pueden modificarse adicionalmente como se describe en cualquier parte en la presente divulgación o el polipéptido de aminoácidos no naturales puede usarse sin modificación adicional. La incorporación de un aminoácido no natural en un polipéptido puede hacerse para una variedad de fines, que incluyen, confeccionar cambios en la estructura y/o función de proteínas, cambiar el tamaño, acidez, nucleofilia, enlace de hidrógeno, hidrofobia, accesibilidad de sitios diana de proteasa, elegir como diana un resto (que incluye para una matriz de proteína), etc. Las proteínas que incluyen un aminoácido no natural pueden tener propiedades catalíticas o biofísicas potenciadas o incluso completamente nuevas. A modo de ejemplo solo, las siguientes propiedades se modifican opcionalmente por inclusión de un aminoácido no natural en una proteína: toxicidad, biodistribución, propiedades estructurales, propiedades espectroscópicas, propiedades químicas y/o fotoquímicas, capacidad catalítica, semivida (incluyendo semivida en suero), capacidad para reaccionar con otras moléculas, que incluye covalentemente o no covalentemente, y similares. Las composiciones que incluyen proteínas que incluyen al menos un aminoácido no natural son útiles para incluir terapéuticos novedosos, diagnósticos, enzimas catalíticas, enzimas industriales, proteínas de unión (incluyendo anticuerpos), e incluyendo el estudio de estructura y función de proteínas. Véase, por ejemplo, Dougherty, (2000) Unnatural Amino Acids as Probes of Protein Structure and Function, Current Opinion in Chemical Biology, 4:645-652.

Además, la cadena lateral del (de los) componente(s) de aminoácidos no naturales de un polipéptido puede proporcionar un amplia variedad de funcionalidad adicional al polipéptido; a modo de ejemplo solo, y no como una limitación, la cadena lateral de la porción de aminoácido no natural de un polipéptido puede incluir cualquiera de lo siguiente: una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo

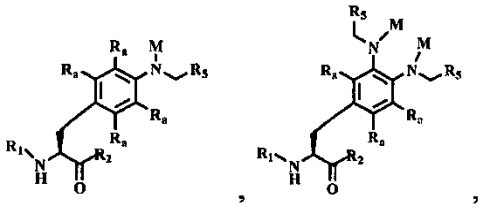
fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos.

En un aspecto, una composición incluye al menos un proteína con al menos uno, que incluye al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, o al menos diez o más aminoácidos no naturales. Los aminoácidos no naturales pueden ser iguales o diferentes, que incluyen, pero no se limitan a, puede haber 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más sitios diferentes en la proteína que comprenden 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más aminoácidos no naturales diferentes. En otro aspecto, una composición incluye una proteína con al menos uno, pero menos de todos, de un aminoácido particular presente en la proteína está sustituido con el aminoácido no natural. Para una proteína dada con más de un aminoácido no natural, los aminoácidos no naturales pueden ser idénticos o diferentes (incluyendo que la proteína puede incluir dos o más tipos diferentes de aminoácidos no naturales, o puede incluir dos de los mismos aminoácidos no naturales). Para una proteína dada con más de dos aminoácidos no naturales, los aminoácidos no naturales pueden ser iguales o diferentes.

Aunque realizaciones de los polipéptidos de aminoácidos no naturales descritos en el presente documento pueden sintetizarse químicamente mediante procedimientos de síntesis de péptidos en fase sólida (por ejemplo, sobre una resina sólida), por procedimientos de síntesis de péptidos en fase de disolución y/o sin la ayuda de enzimas, otras realizaciones de los polipéptidos de aminoácidos no naturales descritos en el presente documento permiten la síntesis mediante una membrana celular, extracto celular, o sistema de lisado o mediante un sistema *in vivo*, es decir, usando la maquinaria celular de una célula procariota o eucariota. En otras realizaciones o realizaciones adicionales, una de las características clave de los polipéptidos de aminoácidos no naturales descritos en el presente documento es que pueden sintetizarse utilizando ribosomas. En otras realizaciones o realizaciones adicionales de los polipéptidos de aminoácidos no naturales descritos en el presente documento, los polipéptidos de aminoácidos no naturales pueden sintetizarse por una combinación de resinas sólidas, sin la ayuda de enzimas, mediante la ayuda de ribosomas, mediante un sistema *in vitro*, mediante un sistema *in vivo*, o cualquier combinación de los mismos.

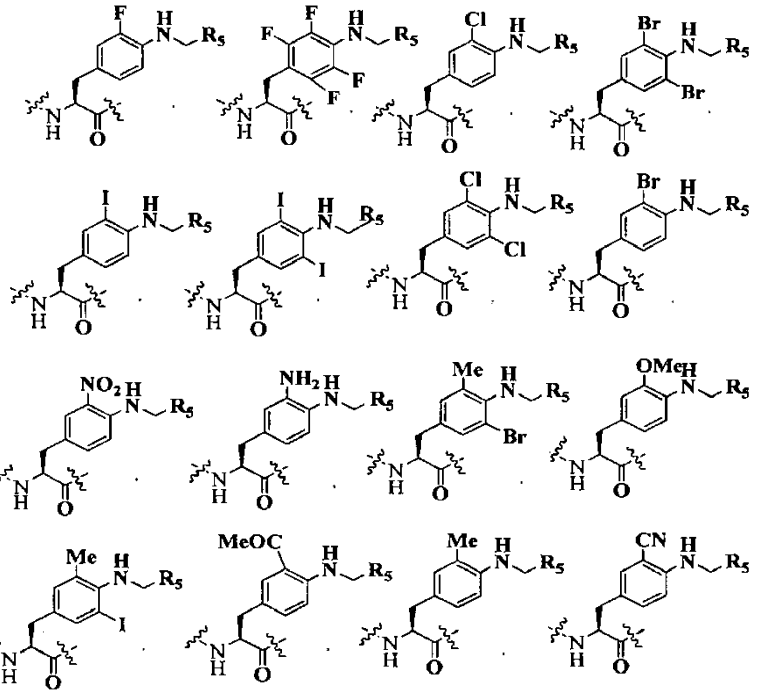
La síntesis de polipéptidos de aminoácidos no naturales mediante ribosomas y/o un sistema *in vivo* tiene distintas ventajas y características de un polipéptido de aminoácidos no naturales sintetizado sobre una resina sólida o sin la ayuda de enzimas. Estas ventajas o características incluyen diferentes perfiles de impurezas, un sistema que utiliza ribosomas y/o un sistema *in vivo* tendrá impurezas que proceden del sistema biológico utilizado, que incluye proteínas de células huésped, porciones de membrana y lípidos, mientras que el perfil de impurezas de un sistema que utiliza una resina sólida y/o sin la ayuda de enzimas incluirá disolventes orgánicos, grupos protectores, materiales de resina, reactivos de acoplamiento y otros productos químicos usados en los procedimientos de síntesis. Además, el patrón isotópico del polipéptido de aminoácidos no naturales sintetizado mediante el uso de ribosomas y/o un sistema *in vivo* reflejará el patrón isotópico de la materia prima utilizada para las células; por otra parte; el patrón isotópico del polipéptido de aminoácidos no naturales sintetizado sobre una resina sólida y/o sin la ayuda de enzimas reflejará el patrón isotópico de los aminoácidos utilizados en la síntesis. Además, el aminoácido no natural sintetizado mediante el uso de ribosomas y/o un sistema *in vivo* estará sustancialmente libre de los D-isómeros de los aminoácidos y/o será capaz de incorporar fácilmente aminoácidos de cisteína internos en la estructura del polipéptido, y/o proporcionará raramente polipéptidos de deleción de aminoácidos internos. Por otra parte, un polipéptido de aminoácidos no naturales sintetizado mediante una resina sólida y/o sin el uso de enzimas tendrá un mayor contenido de D-isómeros de los aminoácidos y/o un menor contenido de aminoácidos de cisteína internos y/o un mayor porcentaje de polipéptidos de deleción de aminoácidos internos. Además, un experto en la materia podrá diferenciar un polipéptido de aminoácidos no naturales sintetizado usando un ribosoma y/o un sistema *in vivo* a partir de un polipéptido de aminoácidos no naturales sintetizado mediante una resina sólida y/o sin el uso de enzimas.

En una cierta realización, la presente invención proporciona un procedimiento de preparación de un compuesto o sal del mismo que contiene al menos un aminoácido no natural seleccionado del grupo que consiste en:



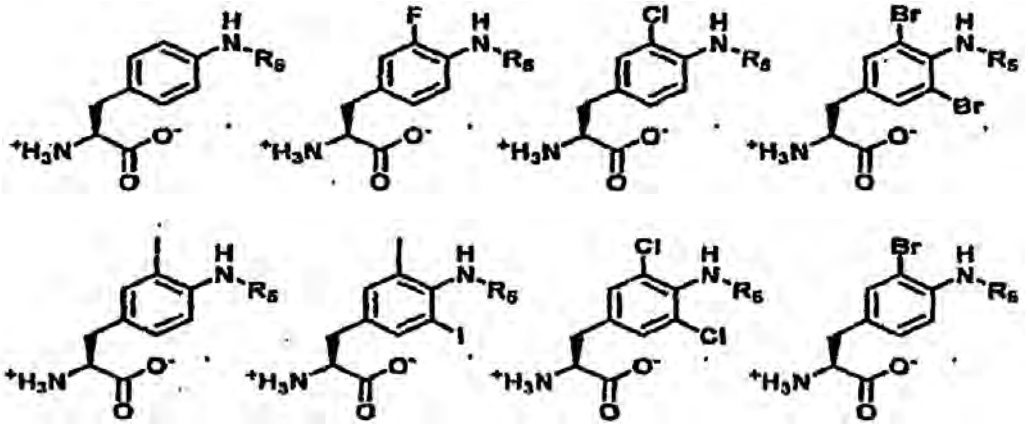
en el que el compuesto se forma por una alquilación reductora de un resto amino aromático sobre al menos un aminoácido no natural con al menos un reactante que comprende al menos un resto de aldehído, en el que R_1, R_2, R_5, R_5 y M son como antes.

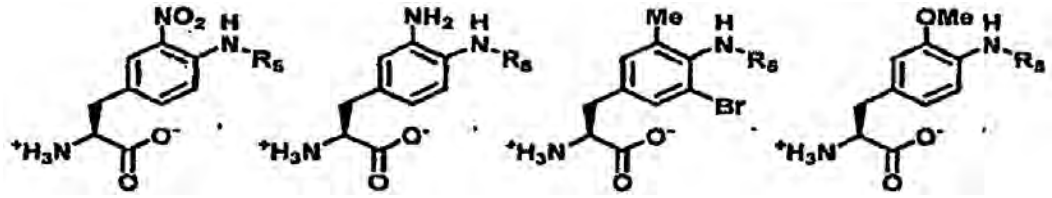
en el que el al menos un aminoácido no natural está seleccionado de un grupo:

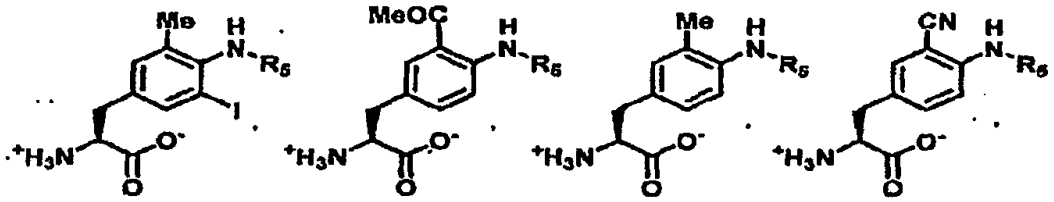


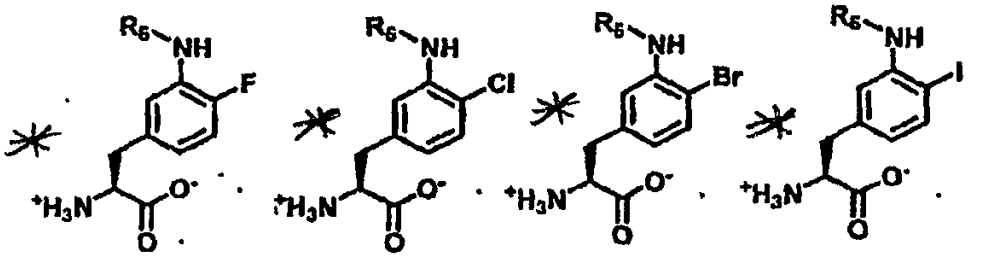
En otras realizaciones es un procedimiento de preparación de un compuesto o sal del mismo en el que al menos un aminoácido no natural está seleccionado de un grupo:

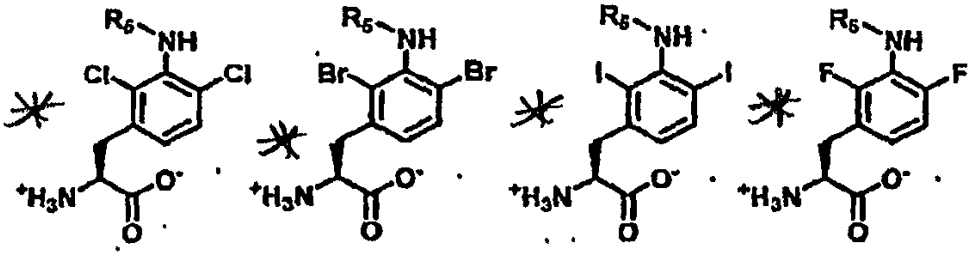
Compuestos marcados con a * no son realizaciones de la invención reivindicada

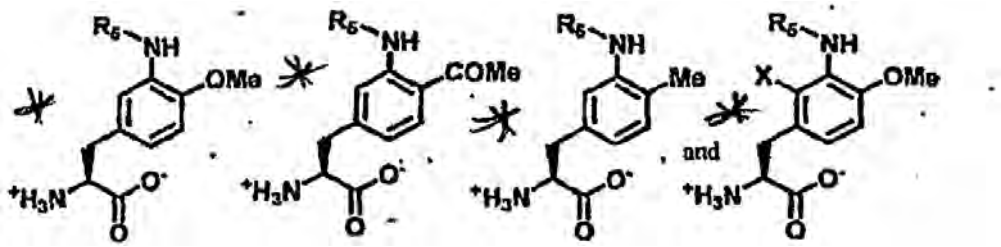




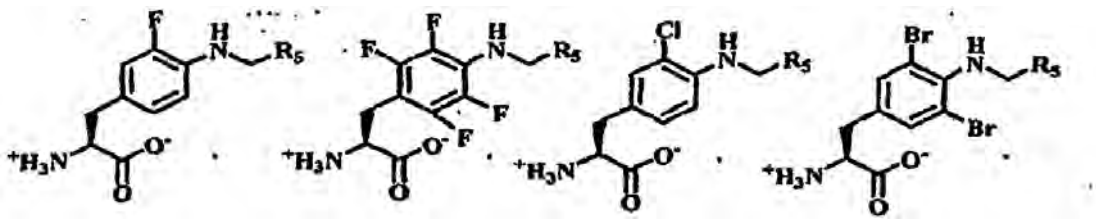


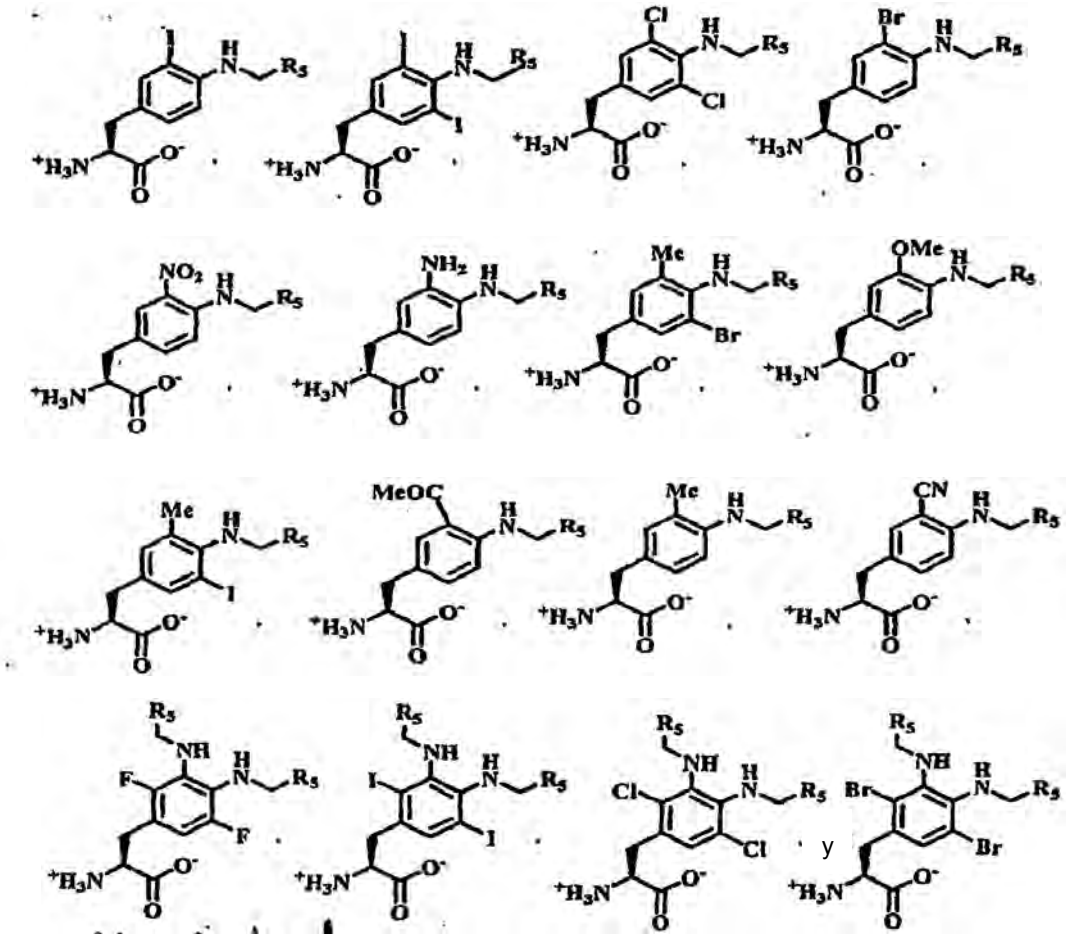






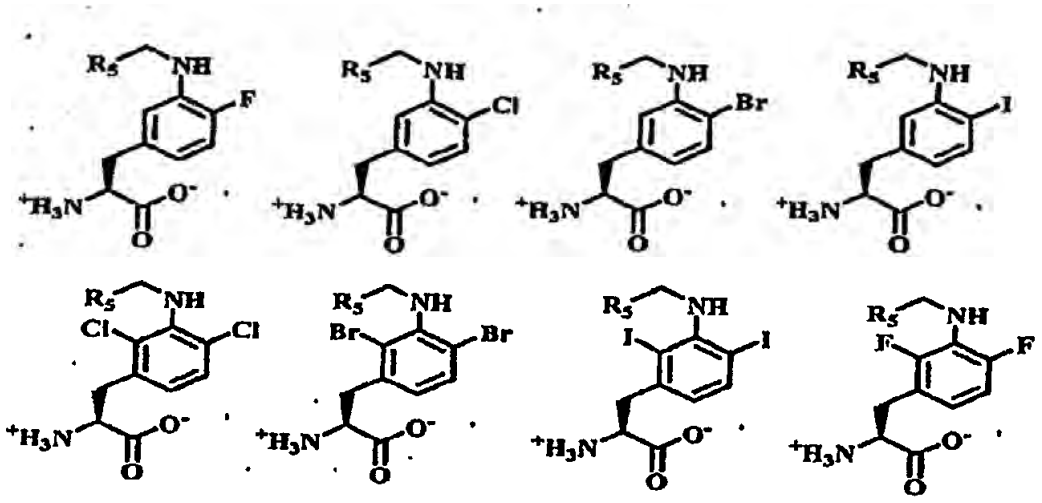
En algunas realizaciones es un procedimiento de preparación de un compuesto o sal del mismo en el que al menos un aminoácido no natural está seleccionado de un grupo:

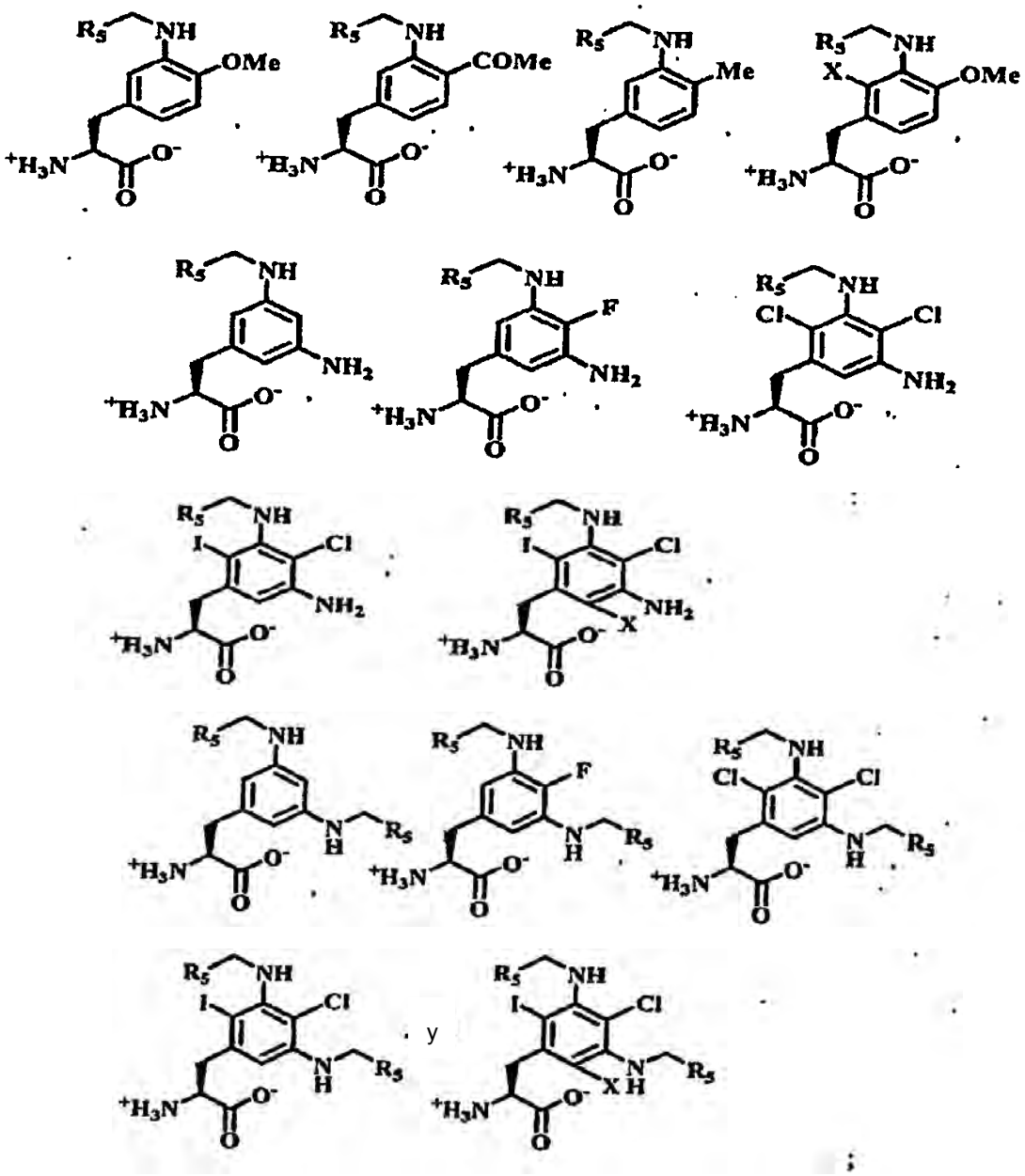




y

También se desvela un procedimiento de preparación de un compuesto o sal del mismo en el que al menos un aminoácido no natural está seleccionado de un grupo (ninguno de los cuales son realizaciones de la invención reivindicada) :





y

en el que X es un halógeno.

VII. Composiciones y procedimientos que comprenden ácidos nucleicos y oligonucleótidos

A. Procedimientos de ácidos nucleicos recombinantes generales para su uso en el presente documento

En numerosas realizaciones de los procedimientos y composiciones descritos en el presente documento, ácidos nucleicos que codifican un polipéptido de interés (incluyendo a modo de ejemplo un polipéptido de GH) se aislarán, clonarán y frecuentemente alterarán usando procedimientos recombinantes. Se usan realizaciones tales que incluyen la expresión de proteínas o durante la generación de variantes, derivados, casetes de expresión, u otras secuencias derivadas de un polipéptido. En algunas realizaciones, las secuencias que codifican los polipéptidos están operativamente ligadas a un promotor heterólogo.

En el presente documento también se describen células que pueden producir polipéptidos de aminoácidos no naturales en los que al menos un aminoácido no natural sobre el polipéptido comprende una cadena lateral que tiene un resto de amina aromática o un resto de aldehído enmascarado o protegido. Las células que biosintetizan al menos un polipéptido de aminoácidos no naturales pueden producirse usando las técnicas, procedimientos, composiciones y estrategias descritos en el presente documento, o variantes de los mismos.

Una secuencia de nucleótidos que codifica un polipéptido que comprende un aminoácido no natural puede sintetizarse basándose en la secuencia de aminoácidos del polipéptido parental, y a continuación cambiando la secuencia de nucleótidos de manera que se efectúe la introducción (es decir, incorporación o sustitución) o eliminación (es decir, deleción o sustitución) del (de los) residuo(s) de aminoácido(s) relevante(s). La secuencia de nucleótidos puede modificarse convenientemente por mutagénesis dirigida al sitio según procedimientos convencionales. Alternativamente, la secuencia de nucleótidos puede prepararse por síntesis química, que incluye usar un sintetizador de oligonucleótidos, en el que se diseñan oligonucleótidos basándose en la secuencia de aminoácidos del polipéptido deseado, y preferentemente seleccionando aquellos codones que son favorecidos en la célula huésped en la que se producirá el polipéptido recombinante. Por ejemplo, varios oligonucleótidos pequeños que codifican porciones del polipéptido deseado pueden sintetizarse y ensamblarse por PCR, ligación o reacción en cadena de ligación. Véanse, por ejemplo, Barany, y col., Proc. Natl. Acad. Sci. 88: 189-193 (1991); documento U.S. 6.521.427.

Los procedimientos y composiciones de aminoácidos no naturales descritos en el presente documento utilizan técnicas rutinarias en el campo de la genética recombinante. Textos básicos que desvelan los procedimientos generales de uso para los procedimientos y composiciones de aminoácidos no naturales descritos en el presente documento incluyen Sambrook y col., Molecular Cloning, A Laboratory Manual (3ª ed. 2001); Kriegler, Gene Transfer and Expression: A Laboratory Manual (1990); y Current Protocols in Molecular Biology (Ausubel y col., eds., 1994).

Textos generales que describen técnicas de biología molecular incluyen Berger y Kimmel, Guide to Molecular Cloning Techniques. Methods in Enzymology volume 152 Academic Press, Inc., San Diego, CA (Berger); Sambrook y col., Molecular Cloning - A Laboratory Manual (2ª ed.) vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989 (“Sambrook”) y Current Protocols in Molecular Biology, F.M. Ausubel y col., eds., Current Protocols, una empresa conjunta entre Greene Publishing Associates, Inc. y John Wiley & Sons, Inc., (complementado durante 1999) (“Ausubel”)). Estos textos describen mutagénesis, el uso de vectores, promotores y muchos otros tópicos relevantes relacionados con, que incluyen la generación de genes o polinucleótidos que incluyen selectores de codones para la producción de proteínas que incluyen aminoácidos no naturales, ARNt ortogonales, sintetasas ortogonales, y pares de los mismos.

Se usan diversos tipos de mutagénesis en los procedimientos y composiciones de aminoácidos no naturales descritos en el presente documento para una variedad de fines, que incluyen, pero no se limitan a, para producir novedosas sintetasas o ARNt, para mutar moléculas de ARNt, para mutar polinucleótidos que codifican sintetasas, para producir bibliotecas de ARNt, para producir bibliotecas de sintetasas, para producir selectores de codones, para insertar selectores de codones que codifican aminoácidos no naturales en una proteína o polipéptido de interés. Incluyen mutagénesis puntuales aleatorias dirigidas al sitio, recombinación homóloga, barajado de ADN u otros procedimientos de mutagénesis recurrentes, construcción quimérica, mutagénesis usando moldes que contienen uracilo, mutagénesis dirigida al oligonucleótido, mutagénesis de ADN modificado con fosforotioato, mutagénesis usando ADN dúplex con huecos o similares, o cualquier combinación de los mismos. Procedimientos adecuados adicionales incluyen reparación de desapareamientos puntuales, mutagénesis usando cepas huésped deficientes en la reparación, restricción-selección y restricción-purificación, mutagénesis por deleción, mutagénesis por síntesis génica total, reparación de roturas bicatenarias y similares. La mutagénesis, que incluye la participación de construcciones quiméricas, también está incluida en los procedimientos y composiciones de aminoácidos no naturales descritos en el presente documento. En una realización, la mutagénesis puede guiarse mediante información conocida de la molécula que se produce naturalmente o molécula que se produce naturalmente alterada o mutada, que incluye, pero no se limita a, secuencia, comparaciones de secuencias, propiedades físicas, estructura cristalina o similares.

Los textos y ejemplos encontrados en el presente documento describen estos procedimientos. Información adicional se encuentra en las siguientes publicaciones y referencias citadas dentro de: Ling y col., Approaches to DNA mutagenesis: an overview, Anal Biochem. 254(2): 157-178 (1997); Dale y col., Oligonucleotide-directed random mutagenesis using the phosphorothioate method, Methods Mol. Biol. 57:369-374 (1996); Smith, In vitro mutagenesis, Ann. Rev. Genet. 19:423-462(1985); Botstein & Shortle, Strategies and applications of in vitro mutagenesis, Science 229:1193-1201(1985); Carter, Site-directed mutagenesis, Biochem. J. 237:1-7 (1986); Kunkel, The efficiency of oligonucleotide directed mutagenesis, in Nucleic Acids & Molecular Biology (Eckstein, F. and Lilley, D.M.J. eds., Springer Verlag, Berlin)) (1987); Kunkel, Rapid and efficient site-specific mutagenesis without phenotypic selection, Proc. Natl. Acad. Sci. USA 82:488-492 (1985); Kunkel y col., Rapid and efficient site-specific mutagenesis without phenotypic selection, Methods in Enzymol. 154, 367-382 (1987); Bass y col., Mutant Trp repressors with new DNA-binding specificities, Science 242:240-245 (1988); Zoller & Smith, Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragments, Nucleic Acids Res. 10:6487-6500 (1982); Zoller & Smith, Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors, Methods in Enzymol. 100:468-500 (1983); Zoller & Smith, Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template, Methods in Enzymol. 154:329-350 (1987); Taylor y col., The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA, Nucl. Acids Res. 13: 8749-8764 (1985); Taylor y col., The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA, Nucl. Acids Res. 13:

8765-8785 (1985); Nakamaye & Eckstein, Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis, Nucl. Acids Res. 14: 9679-9698 (1986); Sayers y col., 5'-3' Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis, Nucl. Acids Res. 16:791-802 (1988); Sayers y col., Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide, (1988) Nucl. Acids Res. 16: 803-814; Kramer y col., The gapped duplex DNA approach to oligonucleotide-directed mutation construction, Nucl. Acids Res. 12: 9441-9456 (1984); Kramer & Fritz Oligonucleotide-directed construction of mutations via gapped duplex DNA, Methods in Enzymol. 154:350-367 (1987); Kramer y col., Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations, Nucl. Acids Res. 16: 7207 (1988); Fritz y col., Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro, Nucl. Acids Res. 16: 6987-6999 (1988); Kramer y col., Point Mismatch Repair, Cell 38:879-887 (1984); Carter y col., Improved oligonucleotide site-directed mutagenesis using M13 vectors, Nucl. Acids Res. 13: 4431-4443 (1985); Carter, Improved oligonucleotide-directed mutagenesis using M13 vectors, Methods in Enzymol. 154: 382-403 (1987); Eghtedarzadeh & Henikoff, Use of oligonucleotides to generate large deletions, Nucl. Acids Res. 14: 5115 (1986); Wells y col., Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin, Phil. Trans. R. Soc. Lond. A 317: 415-423 (1986); Nambiar y col., Total synthesis and cloning of a gene coding for the ribonuclease S protein, Science 223: 1299-1301 (1984); Sakmar and Khorana, Total synthesis and expression of a gene for the alpha subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin), Nucl. Acids Res. 14: 6361-6372 (1988); Wells y col., Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites, Gene 34:315-323 (1985); Grundström y col., Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis, Nucl. Acids Res. 13: 3305-3316 (1985); Mandecki, Oligonucleotide-directed double-strand break repair in plasmids of Escherichia coli: a method for site-specific mutagenesis, Proc. Natl. Acad. Sci. USA, 83:7177-7181 (1986); Arnold, Protein engineering for unusual environments, Current Opinion in Biotechnology 4:450-455 (1993); Sieber, y col., Nature Biotechnology, 19:456-460 (2001). W. P. C. Stemmer, Nature 370, 389-91 (1994); y I. A. Lorimer, I. Pastan, Nucleic Acids Res. 23, 3067-8 (1995). Detalles adicionales sobre muchos de los procedimientos anteriores pueden encontrarse en Methods in Enzymology Volume 154, que también describe controles para la resolución de problemas con diversos procedimientos de mutagénesis.

Los procedimientos y composiciones descritos en el presente documento también incluyen el uso de células huésped eucariotas, células huésped no eucariotas y organismos para la incorporación *in vivo* de un aminoácido no natural mediante pares ortogonales de ARNt/RS. Las células huésped se manipulan genéticamente (que incluye transforman, transducen o transfectan) con los polinucleótidos correspondientes a los polipéptidos descritos en el presente documento o construcciones que incluyen un polinucleótido correspondiente a los polipéptidos descritos en el presente documento, que incluye un vector correspondiente a los polipéptidos descritos en el presente documento, que puede ser, por ejemplo, un vector de clonación o un vector de expresión. Por ejemplo, las regiones codificantes para el ARNt ortogonal, la ARNt sintetasa ortogonal y la proteína que va a derivatizarse están operativamente ligados a elementos de control de la expresión génica que son funcionales en la célula huésped deseada. El vector puede estar, por ejemplo, en forma de un plásmido, un cósmido, un fago, una bacteria, un virus, un polinucleótido desnudo o un polinucleótido conjugado. Los vectores se introducen en células y/o microorganismos mediante procedimientos convencionales que incluyen electroporación (Fromm y col., Proc. Natl. Acad. Sci. USA 82, 5824 (1985), infección por vectores virales, penetración balística a alta velocidad por partículas pequeñas con el ácido nucleico tanto dentro de la matriz de perlas pequeñas o partículas como sobre la superficie (Klein y col., Nature 327, 70-73 (1987)), y/o similares.

Las células huésped manipuladas pueden cultivarse en medios nutritivos convencionales modificados según convenga para actividades tales como, por ejemplo, etapas de cribado, promotores de la activación o transformantes de selección. Estas células pueden cultivarse opcionalmente en organismos transgénicos. Otras referencias útiles, que incluyen el aislamiento y cultivo de células (por ejemplo, para el posterior aislamiento de ácidos nucleicos), incluyen Freshney (1994) Culture of Animal Cells, a Manual of Basic Technique, tercera edición, Wiley-Liss, New York y las referencias citadas en su interior; Payne y col. (1992) Plant Cell and Tissue Culture in Liquid Systems John Wiley & Sons, Inc. New York, NY; Gamborg and Phillips (eds) (1995) Plant Cell, Tissue and Organ Culture; Fundamental Methods Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York) y Atlas and Parks (eds) The Handbook of Microbiological Media (1993) CRC Press, Boca Raton, FL.

Están disponibles varios procedimientos muy conocidos de introducción de ácidos nucleicos diana en células, cualquiera de los cuales puede usarse en procedimientos y composiciones descritos en el presente documento. Éstos incluyen: fusión del células del receptor con protoplastos bacterianos que contienen el ADN, electroporación, bombardeo con proyectiles e infección con vectores virales (tratados adicionalmente más adelante), etc. Pueden usarse células bacterianas para amplificar el número de plásmidos que contienen construcciones de ADN correspondientes a los polipéptidos descritos en el presente documento. Las bacterias se cultivan a fase logarítmica y los plásmidos dentro de las bacterias pueden aislarse mediante una variedad de procedimientos conocidos en la técnica (véase, por ejemplo, Sambrook). Además, están comercialmente disponibles kits para la purificación de plásmidos de bacterias (véase, por ejemplo, EasyPrep™, FlexiPrep™, ambos de Pharmacia Biotech; StrataClean™, de Stratagene; y, QIAprep™ de Qiagen). A continuación, los plásmidos aislados y purificados se manipulan adicionalmente para producir otros plásmidos, usados para transfectar células o incorporarse en vectores relacionados para infectar organismos. Vectores típicos contienen terminadores de la transcripción y de la

traducción, secuencias de iniciación de la transcripción y de la traducción, y promotores útiles para la regulación de la expresión del ácido nucleico diana particular. Los vectores comprenden opcionalmente casetes de expresión genéricos que contiene al menos una secuencia terminadora independiente, secuencias que permiten la replicación del casete en eucariotas, o procariotas, o ambos (que incluyen, pero no se limitan a, vectores lanzadera) y marcadores de selección para tanto sistemas procariotas como eucariotas. Los vectores son adecuados para la replicación e integración en procariotas, eucariotas, o ambos. Véanse Gillam & Smith, Gene 8:81 (1979); Roberts y col., Nature. 328:731 (1987); Schneider, E. y col., Protein Expr. Purif. 6(1):10-14 (1995); Ausubel, Sambrook, Berger (todos arriba). Un catálogo de bacterias y bacteriófagos útiles para la clonación se proporciona, por ejemplo, por la ATCC, por ejemplo, The ATCC Catalogue of bacteria and bacteriophage (1992) Gherna y col. (eds) publicado por la ATCC. Procedimientos básicos adiciones de secuenciación, clonación y otros aspectos de la biología molecular y consideraciones teóricas subyacentes también se encuentran en Watson y col (1992) Recombinant DNA Second Edition Scientific American Books, NY. Además, esencialmente cualquier ácido nucleico (y prácticamente cualquier ácido nucleico marcado, tanto estándar como no estándar) puede solicitarse a la medida o estándar de cualquiera de una variedad de fuentes comerciales, tales como Midland Certified Reagent Company (Midland, TX mcrc.com), The Great American Gene Company (Ramona, CA disponible en la malla mundial en genco.com), ExpressGen Inc. (Chicago, IL disponible en la malla mundial en expressgen.com), Operon Technologies Inc. (Alameda, CA) y muchos otros.

B. Selectores de codones

Los selectores de codones englobados dentro de los procedimientos y composiciones descritos en el presente documento amplían la región estructural de codones genéticos de la maquinaria biosintética de proteínas. Por ejemplo, un selector de codones incluye, pero no se limita a, un único codón de tres bases, un codón de parada, tal como un codón de terminación, que incluye un codón ámbar (UAG), un codón ocre o un codón ópalo (UGA), un codón no natural, un codón de cuatro o más bases, un codón raro, o similares. Hay un amplio intervalo en el número de selectores de codones que pueden introducirse en un gen o polinucleótido deseado que incluye uno o más, dos o más, tres o más, 4, 5, 6, 7, 8, 9, 10 o más en un polinucleótido individual que codifica al menos una porción de un polipéptido de interés.

En una realización, los procedimientos implican el uso de un selector de codones que es un codón de terminación para la incorporación de uno o más aminoácidos no naturales *in vivo.* Por ejemplo, se produce un O-ARNt que reconoce el codón de terminación, que incluye UAG, y se aminoacila por una O-RS con un aminoácido no natural deseado. Este O-ARNt no es reconocido por las aminoacil-ARNt sintetasas del huésped que se producen naturalmente. Puede usarse mutagénesis dirigida al sitio convencional para introducir el codón de terminación, que incluye TAG, en el sitio de interés en un polipéptido de interés. Véase, por ejemplo, Sayers, J.R., y col. (1988), 5-3' Exonuclease in phosphorothioate-based oligonucleotide-directed mutagenesis. Nucleic Acids Res, 16:791-802. Si la O-RS, O-ARNt y el ácido nucleico que codifica el polipéptido de interés se combinan *in vivo*, el aminoácido no natural se incorpora en respuesta al codón UAG para dar un polipéptido que contiene el aminoácido no natural en la posición especificada.

También pueden codificarse aminoácidos no naturales con codones raros. Por ejemplo, cuando se reduce la concentración de arginina en una reacción de síntesis de proteínas *in vitro*, el codón de arginina raro, AGG, ha demostrado ser eficaz para la inserción de Ala por un ARNt sintético acilado con alanina. Véase, por ejemplo, Ma y col., Biochemistry, 32:7939 (1993). En este caso, el ARNt sintético compite con el ARNtArg que se produce naturalmente, que existe como especie minoritaria en *Escherichia coli*. Algunos organismos no usan todos los codones tripletes. Se ha utilizado un codón AGA no asignado en *Micrococcus luteus* para la inserción de aminoácidos en un extracto de transcripción/traducción *in vitro*. Véase, por ejemplo, Kowal y Oliver, Nucl. Acid. Res., 25:4685 (1997). Los componentes de la presente invención pueden generarse para usar estos codones raros *in vivo*.

La incorporación de aminoácidos no naturales *in vivo* puede hacerse sin perturbación significativa de la célula huésped eucariota. Por ejemplo, debido a que la eficiencia de supresión para el codón UAG depende de la competición entre el O-ARNt, que incluye el ARNt supresor ámbar, y un factor de liberación eucariota (que incluye, pero no se limita a, eRF) (que se une a un codón de terminación e inicia la liberación del péptido en crecimiento del ribosoma), la eficiencia de supresión puede modularse por, que incluye aumentar el nivel de expresión de O-ARNt, y/o el ARNt supresor.

Los selectores de codones también comprenden codones prolongados, que incluyen codones de cuatro o más bases, tales como codones de cuatro, cinco, seis o más bases. Ejemplos de codones de cuatro bases incluyen, AGGA, CUAG, UAGA, CCCU y similares. Ejemplos de codones de cinco bases incluyen AGGAC, CCCCU, CCCUC, CUAGA, CUACU, UAGGC y similares. Una característica de los procedimientos y composiciones descritos en el presente documento incluye usar codones prolongados basándose en la supresión del desplazamiento del marco. Pueden insertarse codones de cuatro o más bases, que incluyen uno o múltiples aminoácidos no naturales en la misma proteína. Por ejemplo, en presencia de O-ARNt mutados, que incluyen un ARNt supresor del desplazamiento del marco especial, con bucles anti-codón, por ejemplo, con al menos bucles anti-codón de 8-10 nt, el codón de cuatro o más bases se lee como un único aminoácido. En otras realizaciones, los bucles anti-codón pueden descodificarse, que incluye al menos un codón de cuatro bases, al menos un codón de cinco bases, o al menos un

codón de seis bases o más. Como hay 256 codones de cuatro bases posibles, pueden codificarse múltiples aminoácidos no naturales en la misma célula usando un codón de cuatro o más bases. Véanse Anderson y col., (2002) Exploring the Limits of Codon and Anticodon Size, Chemistry and Biology, 9:237-244; Magliery, (2001) Expanding the Genetic Code: Selection of Efficient Suppressors of Four-base Codons and Identification of "Shifty" Four-base Codons with a Library Approach in Escherichia coli, J. Mol. Biol. 307: 755-769.

Por ejemplo, se han usado codones de cuatro bases para incorporar aminoácidos no naturales en proteínas usando procedimientos biosintéticos *in vitro*. Véanse, por ejemplo, Ma y col., (1993) Biochemistry, 32:7939; y Hohsaka y col., (1999) J. Am. Chem. Soc., 121:34. Se usaron CGGG y AGGU para incorporar simultáneamente 2-naftilalanina y un derivado de NBD de lisina en estreptavidina *in vitro* con dos ARNt supresores del desplazamiento del marco químicamente acilados. Véase, por ejemplo, Hohsaka y col., (1999) J. Am. Chem. Soc., 121:12194. En un estudio *in vivo*, Moore y col. examinaron la capacidad de derivados de ARNtLeu con anti-codones NCUA para suprimir codones UAGN (N puede ser U, A, G o C), y se encontró que el cuadruplete UAGA podía ser descodificado por un ARNtLeu con un anti-codón UCUA con una eficiencia del 13 al 26 % con poca descodificación en el marco 0 o -1. Véase, Moore y col., (2000) J. Mol. Biol., 298:195. En una realización, pueden usarse codones prolongados basados en codones raros o codones de parada en los procedimientos y composiciones descritos en el presente documento, que pueden reducir la ultralectura por cambio de aminoácido y la supresión del desplazamiento del marco en otros sitios no deseados.

Para un sistema dado, un selector de codones también puede incluir uno de los codones de tres bases naturales, en el que el sistema endógeno no usa (o raramente usa) el codón de bases naturales. Por ejemplo, éste incluye un sistema que carece de un ARNt que reconoce el codón de tres bases natural, y/o un sistema en el que el codón de tres bases es un codón raro.

Los selectores de codones opcionalmente incluyen pares de bases no naturales. Estos pares de bases no naturales amplían adicionalmente el alfabeto genético existente. Un par de bases adicional aumenta el número de codones triplete de 64 a 125. Las propiedades de los terceros pares de bases incluyen apareamiento de bases estable y selectivo, incorporación enzimática eficaz en ADN con alta fidelidad por una polimerasa y la eficaz extensión del cebador continuada después de la síntesis del par de bases no natural naciente. Descripciones de pares de bases no naturales que pueden adaptarse para los procedimientos y composiciones incluyen, por ejemplo, Hirao, y col., (2002) An unnatural base pair for incorporating amino acid analogues into protein, Nature Biotechnology, 20:177-182, y véase también Wu, Y. y col. (2002) J. Am. Chem. Soc. 124:14626-14630. Otras publicaciones relevantes se enumeran más adelante.

Para el uso *in vivo*, el nucleósido no natural es permeable a la membrana y se fosforila para formar el trifosfato correspondiente. Además, el aumento de la información genética es estable y no es destruida por enzimas celulares. Los intentos anteriores de Benner y otros aprovecharon los patrones de enlaces de hidrógeno que son diferentes de aquellos en los pares canónicos de Watson-Crick, cuyo ejemplo más notable es el par iso-C:iso-G. Véanse, por ejemplo, Switzer y col., (1989) J. Am. Chem. Soc., 111:8322; y Piccirilli y col., (1990) Nature, 343:33; Kool, (2000) Curr. Opin. Chem. Biol., 4:602. Estas bases en general se aparean incorrectamente en cierta medida con las bases naturales y no pueden replicarse enzimáticamente. Kool y colaboradores demostraron que las interacciones de empaquetamiento hidrófobas entre bases pueden reemplazar los enlaces de hidrógeno para conducir la formación de pares de bases. Véase Kool, (2000) Curr. Opin. Chem. Biol., 4:602, y Guckian y Kool, (1998) Angew. Chem. Int. Ed. Engl., 36, 2825. En un intento de desarrollar un par de bases no naturales que cumplan todos los requisitos anteriores, Schultz, Romesberg y colaboradores han sintetizado y estudiado sistemáticamente varias bases hidrófobas no naturales. Un auto-par PICS:PICS resulta ser más estable que los pares de bases naturales y puede incorporarse eficazmente en el ADN por el fragmento Klenow (KF) de la ADN polimerasa I de *Escherichia coli*. Véanse, por ejemplo, McMinn y col., (1999) J. Am. Chem. Soc., 121:11586; y Ogawa y col., (2000) J. Am. Chem. Soc., 122:3274. Un auto-par 3MN:3MN puede sintetizarse por KF con eficiencia y selectividad suficiente para la función biológica. Véase, por ejemplo, Ogawa y col., (2000) J. Am. Chem. Soc., 122:8803. Sin embargo, ambas bases actúan de terminador de cadena para la replicación adicional. Se ha desarrollado recientemente una ADN polimerasa mutante que puede utilizarse para replicar el auto-par PICS. Además, puede replicarse un auto-par 7AI. Véase, por ejemplo, Tae y col., (2001) J. Am. Chem. Soc., 123:7439. También se ha desarrollado un par de metalobases novedoso, Dipic:Py, que forma un par estable tras su unión a Cu (II). Véase Meggers y col., (2000) J. Am. Chem. Soc., 122:10714. Debido a que los codones prolongados y los codones no naturales son intrínsecamente ortogonales a los codones naturales, los procedimientos de aminoácidos no naturales descritos en el presente documento pueden aprovechar esta propiedad para generar ARNt ortogonales para ellos.

También puede usarse un sistema de derivación de la traducción para incorporar un aminoácido no natural en un polipéptido deseado. En un sistema de derivación de la traducción, se incorpora una gran secuencia en un gen, pero no se traduce en proteína. La secuencia contiene una estructura que sirve de señal para inducir al ribosoma a saltar sobre la secuencia y reanudar la traducción en la dirección 3' de la inserción.

En ciertas realizaciones, la proteína o polipéptido de interés (o porción del mismo) en los procedimientos y/o composiciones descritos en el presente documento está codificada por un ácido nucleico. Normalmente, el ácido nucleico comprende al menos un selector de codones, al menos dos selectores de codones, al menos tres

selectores de codones, al menos cuatro selectores de codones, al menos cinco selectores de codones, al menos seis selectores de codones, al menos siete selectores de codones, al menos ocho selectores de codones, al menos nueve selectores de codones, diez o más selectores de codones.

Los genes que codifican proteínas o polipéptidos de interés pueden mutagenizarse usando procedimientos muy conocidos para un experto en la materia y descritos en el presente documento bajo "Mutagénesis y otras técnicas de biología molecular" que incluyen, por ejemplo, uno o más selectores de codones para la incorporación de un aminoácido no natural. Por ejemplo, un ácido nucleico para una proteína de interés se mutageniza para incluir uno o más selectores de codones, proporcionando la incorporación del uno o más aminoácidos no naturales. Los procedimientos y composiciones descritos en el presente documento incluyen cualquier variante tal, que incluye, pero no se limita a, mutante, versiones de cualquier proteína, por ejemplo, que incluye al menos un aminoácido no natural. Similarmente, los procedimientos y composiciones descritos en el presente documento incluyen ácidos nucleicos correspondientes, es decir, cualquier ácido nucleico con uno o más selectores de codones que codifique uno o más aminoácidos no naturales.

Las moléculas de ácidos nucleicos que codifican una proteína de interés, que incluye a modo de ejemplo solo, polipéptido de GH, pueden mutarse fácilmente para introducir una cisteína en cualquier posición deseada del polipéptido. La cisteína se usa ampliamente para introducir moléculas reactivas, polímeros solubles en agua, proteínas, o una amplia variedad de otras moléculas, sobre una proteína de interés. Procedimientos adecuados para la incorporación de cisteína en una posición deseada de un polipéptido son muy conocidos en la técnica, tales como aquellos descritos en la patente de EE.UU. nº 6.608.183 y técnicas de mutagénesis estándar.

VIII: Generación in vivo de polipéptidos que comprenden aminoácidos no naturales

Los polipéptidos descritos en el presente documento pueden generarse *in vivo* usando ARNt y ARNt sintetasas modificados para añadir o sustituir aminoácidos que no son codificados en sistemas que se producen naturalmente.

Los procedimientos para generar ARNt y ARNt sintetasas que usan aminoácidos que no son codificados en sistemas que se producen naturalmente se describen en, por ejemplo, la patente de EE.UU. nº 7.045.337 titulada "In vivo incorporation of unnatural amino acids" y la patente de EE.UU. nº 7.083.970 titulada "Methods and compositions for the production of orthogonal tRNA-aminoacyl tRNA synthetase pairs". Estos procedimientos implican generar una maquinaria de traducción que funciona independientemente de las sintetasas y ARNt endógenos al sistema de traducción (y, por tanto, algunas veces denominados "ortogonales"). En una realización, el sistema de traducción comprende un polinucleótido que codifica el polipéptido; el polinucleótido puede ser ARNm que se transcribió del ADN correspondiente, o el ARNm puede surgir de un vector viral de ARN; adicionalmente, el polinucleótido comprende un selector de codones correspondiente al sitio previamente designado de incorporación para el aminoácido no natural. La traducción comprende además un ARNt que comprende el aminoácido no natural, en el que el ARNt es específico para el selector de codones anteriormente mencionado; en otras realizaciones, el aminoácido no natural está aminoacilado. En otras realizaciones o realizaciones adicionales, el sistema de traducción comprende una aminoacil sintetasa específica para el ARNt, y en otras realizaciones o realizaciones adicionales el sistema de traducción comprende un ARNt ortogonal y una aminoacil ARNt sintetasa ortogonal. En otras realizaciones o realizaciones adicionales, el sistema de traducción comprende al menos uno de los siguientes: un plásmido que comprende el polinucleótido anteriormente mencionado (normalmente en forma de ADN), ADN genómico que comprende el polinucleótido anteriormente mencionado (normalmente en forma de ADN), o ADN genómico en el que se ha integrado el polinucleótido anteriormente mencionado (en otras realizaciones, la integración es integración estable). En otras realizaciones o realizaciones adicionales del sistema de traducción, el selector de codones está seleccionado del grupo que consiste en un codón ámbar, codón ocre, codón ópalo, un codón único, un codón raro, un codón no natural, un codón de cinco bases y un codón de cuatro bases. En otras realizaciones o realizaciones adicionales del sistema de traducción, el ARNt es un ARNt supresor. En otras realizaciones o realizaciones adicionales, el polipéptido de aminoácidos no naturales se sintetiza por un ribosoma.

En otras realizaciones o realizaciones adicionales, el sistema de traducción comprende un ARNt ortogonal (O-ARNt) y una aminoacil ARNt sintetasa ortogonal (O-RS). Normalmente, la O-RS aminoacila preferencialmente el O-ARNt con al menos un aminoácido no natural en el sistema de traducción y el O-ARNt reconoce al menos un selector de codones que no es reconocido por otros ARNt en el sistema. Así, el sistema de traducción inserta el aminoácido no natural en una proteína producida en el sistema, en respuesta a un selector de codones codificado, "sustituyendo" así un aminoácido en una posición en el polipéptido codificado.

Se ha descrito una amplia variedad de ARNt y aminoacil ARNt sintetasas ortogonales en la materia para insertar aminoácidos sintéticos particulares en polipéptidos, y son generalmente adecuados en los procedimientos descritos en el presente documento para producir los polipéptidos de aminoácidos no naturales descritos en el presente documento. Por ejemplo, O-ARNt/aminoacil-ARNt sintetasas específicas para ceto se describen en Wang, L. y col., Proc. Natl. Acad. Sci. USA 100:56-61 (2003) y Zhang, Z. y col., Biochem. 42(22):6735-6746 (2003). O-RS a modo de ejemplo, o porciones de las mismas, están codificados por secuencias de polinucleótidos e incluyen secuencias de aminoácidos desveladas en las publicaciones de solicitud de patente de EE.UU. 2003/0082575 y 2003/0108885. Moléculas de O-ARNt correspondientes para su uso con los O-RS también se describen en la patente de EE.UU. nº

7.045.337 titulada “In vivo incorporation of unnatural amino acids” y la patente de EE.UU. nº 7.083.970 titulada “Methods and compositions for the production of orthogonal tRNA-aminoacyl tRNA synthetase pairs”. Además, Mehl y col. En J. Am. Chem. Soc. 2003; 125:935-939 y Santoro y col. Nature Biotechnology 2002 Oct; 20:1044-1048, tratan procedimientos de cribado y moléculas de aminoacil ARNt sintetasa y de ARNt para la incorporación de p-aminofenilalanina en polipéptidos.

Secuencias de O-ARNt a modo de ejemplo adecuadas para su uso en los procedimientos descritos en el presente documento incluyen, pero no se limitan a, secuencias de nucleótidos SEC ID Nº: 1-3 como se desvelan en la publicación de solicitud de patente de EE.UU. 2003/0108885 (nº de serie 10/126.931). Otros ejemplos de pares de O-ARNt/aminoacil-ARNt sintetasa específicos para aminoácidos no naturales particulares se describen en la publicación de solicitud de patente de EE.UU. 2003/0082575 (nº de serie 10/126.927). O-RS y O-ARNt que incorporan tanto aminoácidos que contienen ceto como azida en *S. cerevisiae* se describen en Chin, J. W., y col., Science 301:964-967 (2003).

El uso de O-ARNt/aminoacil-ARNt sintetasas implica la selección de un codón específico que codifica el aminoácido no natural. Aunque puede usarse cualquier codón, es generalmente deseable seleccionar un codón que sea raramente o nunca usado en la célula en la que se expresa el O-ARNt/aminoacil-ARNt sintetasa. Por ejemplo, codones a modo de ejemplo incluyen codón de parada tales como codones de terminación (ámbar, ocre y ópalo), codones de cuatro o más bases y otros codones de tres bases naturales que son raramente usados o no se usan.

Pueden introducirse selector(es) de codones específicos en posiciones apropiadas en la secuencia codificante de polinucleótidos usando procedimientos de mutagénesis conocidos en la técnica (que incluyen mutagénesis específica del sitio, mutagénesis en casete, mutagénesis de restricción-selección, etc.).

Procedimientos para generar componentes de la maquinaria biosintética de proteínas, tales como O-RS, O-ARNt y pares de O-ARNt/O-RS ortogonales que pueden usarse para incorporar un aminoácido no natural, se describen en Wang, L., y col., Science 292: 498-500 (2001); Chin, J. W., y col., J. Am. Chem. Soc. 124:9026-9027 (2002); Zhang, Z. y col., Biochemistry 42: 6735-6746 (2003). Procedimientos y composiciones para la incorporación *in vivo* de aminoácidos no naturales se describen en la publicación de solicitud de patente de EE.UU. 2003/0082575 (nº de serie 10/126.927). Procedimientos para seleccionar un par de ARNt-ARNt sintetasa ortogonal para su uso en el sistema de traducción *in vivo* de un organismo también se describen en la patente de EE.UU. nº 7.045.337 titulada “In vivo incorporation of unnatural amino acids” y la patente de EE.UU. nº 7.083.970 titulada “Methods and compositions for the production of orthogonal tRNA-aminoacyl tRNA synthetase pairs”. Además, la publicación PCT WO 04/035743 titulada “Site Specific Incorporation of Keto Amino Acids into proteins” describe pares ortogonales de RS y ARNt para la incorporación de ceto-aminoácidos. La publicación PCT nº WO 04/094593 titulada “Expanding the Eukaryotic Genetic Code” describe pares ortogonales de RS y ARNt para la incorporación de aminoácidos no naturalmente codificados en células huésped eucariotas.

Procedimientos de producción de al menos una aminoacil-ARNt sintetasa ortogonal (O-RS) recombinante comprenden: (a) generar una biblioteca de RS (opcionalmente mutantes) derivadas de al menos una aminoacil-ARNt sintetasa (RS) de un primer organismo, que incluye un organismo procariota, tal como *Methanococcus jannaschii, Methanobacterium thermoautotrophicum, Halobacterium, Escherichia coli, A. fulgidus, P. furiosus, P. horikoshil, A. pernix, T. thermophilus,* o similares, o un organismo eucariota; (b) seleccionar (y/o cribar) la biblioteca de RS (opcionalmente RS mutantes) para miembros que aminoacilan un ARNt ortogonal (O-ARNt) en presencia de un aminoácido no natural y un aminoácido natural, proporcionando así un conjunto de RS activas (opcionalmente mutantes); y/o, (c) seleccionar (opcionalmente mediante selección negativa) el conjunto de RS activas (que incluyen, pero no se limitan a, RS mutantes) que aminoacilan preferencialmente el O-ARNt en ausencia del aminoácido no natural, proporcionando así la al menos una O-RS recombinante; en el que la al menos una O-RS recombinante aminoacila preferencialmente el O-ARNt con el aminoácido no natural.

En una realización, la RS es una RS inactiva. La RS inactiva pueden generarse mutando una RS activa. Por ejemplo, la RS inactiva pueden generarse mutando al menos aproximadamente 1, al menos aproximadamente 2, al menos aproximadamente 3, al menos aproximadamente 4, al menos aproximadamente 5, al menos aproximadamente 6, o al menos aproximadamente 10 o más aminoácidos a diferentes aminoácidos, que incluyen, pero no se limitan a, alanina.

Pueden generarse bibliotecas de RS mutantes usando diversas técnicas conocidas en la técnica, que incluyen, pero no se limitan a, diseño racional basado en estructura de RS tridimensional de la proteína, o mutagénesis de nucleótidos de RS en una técnica de diseño aleatoria o racional. Por ejemplo, las RS mutantes pueden generarse por mutaciones específicas del sitio, mutaciones al azar, mutaciones de recombinación que generan diversidad, construcciones quiméricas, diseño racional y por otros procedimientos descritos en el presente documento o conocidos en la técnica.

En una realización, la selección (y/o cribado) de la biblioteca de RS (opcionalmente RS mutadas) para miembros que son activos, que incluyen aminoacilar un ARNt ortogonal (O-ARNt) en presencia de un aminoácido no natural y un aminoácido natural, incluye: introducir un marcador de selección o de cribado positivo, que incluye, pero no se limita

a, un gen de resistencia a antibióticos, o similares, y la biblioteca de RS (opcionalmente mutantes) en una pluralidad de células, en la que el marcador de selección y/o cribado positivo comprende al menos un selector de codones, que incluye, pero no se limita a, un codón ámbar, codón ocre, codón ópalo, un codón único, un codón raro, un codón no natural, un codón de cinco bases y un codón de cuatro bases; cultivar la pluralidad de células en presencia de un agente de selección; identificar células que sobreviven (o muestran una respuesta específica) en presencia del agente de selección y/o cribado para suprimir el al menos un selector de codones en el marcado de selección o de cribado positivo, proporcionando así un subconjunto de células positivamente seleccionadas que contiene el conjunto de RS activas (opcionalmente mutantes). Opcionalmente, puede variarse la concentración de agente de selección y/o cribado.

En un aspecto, el marcador de selección positivo es un gen cloranfenicol acetiltransferasa (CAT) y el selector de codones es un codón de terminación ámbar en el gen CAT. Opcionalmente, el marcador de selección positivo es un gen β-lactamasa y el selector de codones es un codón de terminación ámbar en el gen β-lactamasa. En otro aspecto, el marcador de cribado positivo comprende un marcador de cribado fluorescente o luminiscente o un marcador de cribado basado en afinidad (que incluye, pero no se limita a, un marcador de la superficie celular).

En una realización, la selección o cribado negativo del conjunto para RS activas (opcionalmente mutantes) que aminoacilan preferencialmente el O-ARNt en ausencia del aminoácido no natural incluye: introducir un marcador de selección o de cribado negativo con el conjunto de RS activas (opcionalmente mutantes) del cribado o selección positivo en una pluralidad de células de un segundo organismo, en el que el marcador de selección o de cribado negativo comprende al menos un selector de codones (que incluye un gen de resistencia a antibióticos, que incluye, pero no se limita a, un gen cloranfenicol acetiltransferasa (CAT)0); e identificar células que sobreviven o muestran una respuesta a la selección específica en un primer medio complementado con el aminoácido no natural y un agente de cribado o de selección, pero dejan de sobrevivir o de mostrar la respuesta específica en un segundo medio no complementado con el aminoácido no natural y el agente de selección o de cribado, proporcionando así células supervivientes o células cribadas con la al menos una O-RS recombinante. Por ejemplo, un protocolo de identificación de CAT actúa opcionalmente de selección positiva y/o cribado negativo en la determinación de O-RS recombinantes apropiadas. Por ejemplo, un conjunto de clones se replica opcionalmente sobre placas de crecimiento que contienen CAT (que comprende al menos un codón selector) tanto con como sin uno o más aminoácidos no naturales. Así, se consideran que las colonias que crecen exclusivamente sobre las placas que contienen aminoácidos no naturales contienen O-RS recombinante. En un aspecto, se varía la concentración del agente de selección (y/o cribado). En algunos aspectos, el primer y segundo organismos son diferentes. Así, el primer y/o segundo organismo comprenden opcionalmente: un procariota, un eucariota, un mamífero, una *Escherichia coli,* un hongo, una levadura, una arqueobacteria, una eubacteria, una planta, un insecto, un protista, etc. En otras realizaciones, el marcador de cribado comprende un marcador de cribado fluorescente o luminiscente o un marcador de cribado basado en afinidad.

En otra realización, el cribar o seleccionar (que incluye seleccionar negativamente) el conjunto para RS activas (opcionalmente mutantes) incluye: aislar el conjunto de RS mutadas activas de la etapa de selección positiva (b); introducir un marcador de selección o de cribado negativo, en el que el marcador de selección o de cribado negativo comprende al menos un selector de codones (que incluye un gen de marcador tóxico, que incluye un gen ribonucleasa barnasa, que comprende al menos un selector de codones) y el conjunto de RS activas (opcionalmente mutantes) en una pluralidad de células de un segundo organismo; e identificar células que sobreviven o muestran una respuesta al cribado específica en un primer medio no complementado con el aminoácido no natural, pero dejan de sobrevivir o mostrar una respuesta al cribado específica en un segundo medio complementado con el aminoácido no natural, proporcionando así células supervivientes o cribadas con la al menos una O-RS recombinante, en el que la al menos una O-RS recombinante es específica para el aminoácido no natural. En un aspecto, el al menos un selector de codones comprende aproximadamente dos o más selectores de codones. Tales realizaciones pueden incluir opcionalmente en las que el al menos un selector de codones comprende dos o más selectores de codones, y en las que el primer y segundo organismo son diferentes (incluyendo cada organismo opcionalmente un procariota, un eucariota, un mamífero, una *Escherichia coli,* un hongo, una levadura, una arqueobacteria, una eubacteria, una planta, un insecto, un protista, etc.). Por tanto, algunos aspectos incluyen en los que el marcador de selección negativo comprende un gen ribonucleasa barnasa (que comprende al menos un selector de codones). Otros aspectos incluyen en los que el marcador de cribado comprende opcionalmente un marcador de cribado fluorescente o luminiscente o un marcador de cribado basado en afinidad. En las realizaciones en el presente documento, los cribados y/o selecciones incluyen opcionalmente variación de la rigurosidad del cribado y/o selección.

En una realización, los procedimientos para producir al menos una aminoacil-ARNt sintetasa ortogonal (O-RS) recombinante pueden comprender además: (d) aislar la al menos una O-RS recombinante; (e) generar un segundo conjunto de -O-RS (opcionalmente mutada) derivada de la al menos una O-RS recombinante; y, (f) repetir las etapas (b) y (c) hasta que se obtiene una O-RS mutada que comprende una capacidad de aminoacilar preferencialmente el O-ARNt. Opcionalmente, las etapas (d)-(f) se repiten, que incluye al menos aproximadamente dos veces. En un aspecto, el segundo conjunto de O-RS mutada derivada de al menos una O-RS recombinante puede generarse por mutagénesis, que incluye mutagénesis al azar, mutagénesis específica del sitio, recombinación o una combinación de los mismos.

La rigurosidad de las etapas de selección/cribado, que incluyen la etapa de selección/cribado positiva (b), la etapa de selección/cribado negativa (c), o tanto las etapas de selección/cribado positivas como negativas (b) y (c), en los procedimientos anteriormente descritos, incluye opcionalmente variar la rigurosidad de la selección/cribado. En otra realización, la etapa de selección/cribado positiva (b), la etapa de selección/cribado negativa (c), o tanto las etapas de selección/cribado positivas como negativas (b) y (c), comprenden usar un indicador, en el que el indicador se detecta por citometría de flujo activada por fluorescencia (FACS) o en el que el indicador se detecta por luminiscencia. Opcionalmente, el indicador se expresa sobre una superficie celular, sobre una expresión en fago o similares y se selecciona basándose en afinidad o actividad catalítica que implica el aminoácido no natural o un análogo. En una realización, la sintetasa mutada se expresa sobre una superficie celular, sobre una expresión en fago o similares.

Los procedimientos de producción de un ARNt ortogonal (O-ARNt) recombinante incluyen: (a) generar una biblioteca de ARNt mutantes derivados de al menos un ARNt, que incluye un ARNt supresor, de un primer organismo; (b) seleccionar (que incluye seleccionar negativamente) o cribar la biblioteca para ARNt (opcionalmente mutantes) que son aminoacilados por una aminoacil-ARNt sintetasa (RS) de un segundo organismo en ausencia de una RS del primer organismo, proporcionando así un conjunto de ARNt (opcionalmente mutantes); y, (c) seleccionar o cribar el conjunto de ARNt (opcionalmente mutantes) para miembros que son aminoacilados por una RS ortogonal (O-RS) introducida, proporcionando así al menos un O-ARNt recombinante; en el que el al menos un O-ARNt recombinante reconoce un selector de codones y no es eficazmente reconocido por la RS del segundo organismo y es preferencialmente aminoacilado por la O-RS. En algunas realizaciones, el al menos un ARNt es un ARNt supresor y/o comprende un único codón de tres bases de bases naturales y/o no naturales, o es un codón de parada, un codón raro, un codón no natural, un codón que comprende al menos 4 bases, un codón ámbar, un codón ocre, o un codón de terminación ópalo. En una realización, el O-ARNt recombinante posee una mejora de la ortogonalidad. Se apreciará que en algunas realizaciones el O-ARNt se importa opcionalmente en un primer organismo desde un segundo organismo sin la necesidad de modificación. En diversas realizaciones, el primer y segundo organismos son tanto iguales como diferentes y se eligen opcionalmente de, que incluyen, pero no se limitan a, procariotas (que incluyen *Methanococcus jannaschii, Methanobacterium thermoautotrophicum, Escherichia coli, Halobacterium,* etc.), eucariotas, mamíferos, hongos, levaduras, arqueobacterias, eubacterias, plantas, insectos, protistas, etc. Adicionalmente, el ARNt recombinante se aminoacila opcionalmente por un aminoácido no natural, en el que el aminoácido no natural se biosintetiza *in vivo* tanto naturalmente como mediante manipulación genética. El aminoácido no natural se añade opcionalmente a un medio de crecimiento para al menos el primer o segundo organismo.

En un aspecto, el seleccionar (que incluye seleccionar negativamente) o cribar la biblioteca para ARNt (opcionalmente mutantes) que son aminoacilados por una aminoacil-ARNt sintetasa (etapa (b)) incluye: introducir un gen de marcador tóxico, en el que el gen de marcador tóxico comprende al menos uno de los selectores de codones (o un gen que conduce a la producción de un agente tóxico o estático o un gen esencial para el organismo en el que tal gen de marcador comprende al menos un selector de codones) y la biblioteca de ARNt (opcionalmente mutantes) en una pluralidad de células del segundo organismo; y, seleccionar células supervivientes, en el que las células supervivientes contienen el conjunto de ARNt (opcionalmente mutantes) que comprende al menos un ARNt ortogonal o ARNt no funcional. Por ejemplo, pueden seleccionarse células supervivientes usando un ensayo de densidad celular por relación de comparación.

En otro aspecto, el gen de marcador tóxico puede incluir dos o más selectores de codones. En otra realización de los procedimientos, el gen de marcador tóxico es un gen ribonucleasa barnasa, en el que el gen ribonucleasa barnasa comprende al menos un codón ámbar. Opcionalmente, el gen ribonucleasa barnasa puede incluir dos o más codones ámbar.

En una realización, el seleccionar o cribar el conjunto de ARNt (opcionalmente mutantes) para miembros que son aminoacilados por una RS ortogonal (O-RS) introducida puede incluir: introducir un gen de marcador de cribado o de selección positivo, en el que el gen de marcador positivo comprende un gen de resistencia a fármacos (que incluye el gen β-lactamasa, que comprende al menos uno de los selectores de codones, tales como al menos un codón de terminación ámbar) o un gen esencial para el organismo, o un gen que conduce a la desintoxicación de un agente tóxico, junto con la O-RS, y el conjunto de ARNt (opcionalmente mutantes) en una pluralidad de células del segundo organismo; e identificar células supervivientes o cribadas cultivadas en presencia de un agente de selección o de cribado, que incluye un antibiótico, proporcionando así un conjunto de células que poseen el al menos un ARNt recombinante, en el que el al menos un ARNt recombinante está aminoacilado por la O-RS e inserta un aminoácido en un producto de traducción codificado por el gen de marcador positivo, en respuesta a el al menos un selector de codones. En otra realización, se varía la concentración del agente de selección y/o de cribado.

Se proporcionan procedimientos para generar pares de O-ARNt/O-RS específicos. Los procedimientos incluyen: (a) generar una biblioteca de ARNt mutantes derivados de al menos un ARNt de un primer organismo; (b) seleccionar o cribar negativamente la biblioteca para ARNt (opcionalmente mutantes) que son aminoacilados por una aminoacil-ARNt sintetasa (RS) de un segundo organismo en ausencia de una RS del primer organismo, proporcionando así un conjunto de ARNt (opcionalmente mutantes); (c) seleccionar o cribar el conjunto de ARNt (opcionalmente mutantes) para miembros que son aminoacilados por una RS ortogonal (O-RS) introducida, proporcionando así al menos un O-

ARNt recombinante. El al menos un O-ARNt recombinante reconoce un selector de codones y no es eficazmente reconocido por la RS del segundo organismo y es preferencialmente aminoacilado por la O-RS. El procedimiento también incluye (d) generar una biblioteca de RS (opcionalmente mutantes) derivadas de al menos un aminoacil-ARNt sintetasa (RS) de un tercer organismo; (e) seleccionar o cribar la biblioteca de RS mutadas para miembros que aminoacilan preferencialmente el al menos un O-ARNt recombinante en presencia de un aminoácido no natural y un aminoácido natural, proporcionando así un conjunto de RS activas (opcionalmente mutantes); y (f) seleccionar o cribar negativamente el conjunto para RS activas (opcionalmente mutantes) que preferencialmente aminoacilan el al menos un O-ARNt recombinante en ausencia del aminoácido no natural, proporcionando así el al menos un par de O-ARNt/O-RS específico, en el que el al menos un par de O-ARNt/O-RS específico comprende al menos una O-RS recombinante que es específica para el aminoácido no natural y el al menos un O-ARNt recombinante. Se incluyen pares de O-ARNt/O-RS específicos producidos mediante los procedimientos. Por ejemplo, el par de O-ARNt/O-RS específico puede incluir, que incluye pero no se limita a, un par de muARNtTyr-mutTyrRS, tal como un par de muARNtTyr-SS12TyrRS, un par de muARNtLeu-mutLeuRS, un par de muARNtThr-mutThrRS, un par de muARNtGlu-mutGluRS, o similares. Adicionalmente, tales procedimientos incluyen en los que el primer y tercer organismo son iguales (incluyendo *Methanococcus jannaschii).*

También están incluidos procedimientos de selección de una par ortogonal de ARNt-ARNt sintetasa para su uso en un sistema de traducción *in vivo* de un segundo organismo en los procedimientos descritos en el presente documento. Los procedimientos incluyen: introducir un gen de marcador, un ARNt y una aminoacil-ARNt sintetasa (RS) aislada o derivada de un primer organismo en un primer conjunto de células del segundo organismo; introducir el gen de marcador y el ARNt en un conjunto de células duplicado de un segundo organismo; y, seleccionar células supervivientes en el primer conjunto que dejan de sobrevivir en el conjunto de células duplicado o cribar células que muestran una respuesta al cribado específica que deja de dar tal respuesta en el conjunto de células duplicado, en el que el primer conjunto y el conjunto de células duplicado se cultivan en presencia de un agente de selección o de cribado, en el que las células supervivientes o cribadas comprenden el par ortogonal de ARNt-ARNt sintetasa para su uso en el sistema de traducción *in vivo* del segundo organismo. En una realización, el comparar y seleccionar o cribar incluye un ensayo de complementación *in vivo*. La concentración del agente de selección o de cribado puede variarse.

Los organismos descritos en el presente documento comprenden una variedad de organismos y una variedad de combinaciones. En una realización, los organismos son opcionalmente un organismo procariota, que incluye, *Methanococcus jannaschii, Methanobacterium thermoautotrophicum, Halobacterium, Escherichia coli, A. fulgidus, P. furiosus, P. horikoshii, A. pernix, T. thermophilus,* o similares. Alternativamente, los organismos comprenden opcionalmente un organismo eucariota, que incluye, pero no se limita a, plantas (que incluyen, pero no se limitan a, plantas complejas tales como monocotiledóneas, o dicotiledóneas), algas, protistas, hongos (que incluyen, pero no se limitan a, levadura, etc.), animales (que incluyen, pero no se limitan a, mamíferos, insectos, artrópodos, etc.), o similares.

A. Expresión en no eucariotas y eucariotas

Las técnicas desveladas en esta sección pueden aplicarse a la expresión en no eucariotas y eucariotas de los polipéptidos de aminoácidos no naturales descritos en el presente documento.

Para obtener expresión de alto nivel de un polinucleótido clonado, normalmente se subclonan polinucleótidos que codifican un polipéptido deseado en un vector de expresión que contiene un promotor fuerte para dirigir la transcripción, un terminador de la transcripción/traducción, y si es para un ácido nucleico que codifica una proteína, un sitio de unión al ribosoma para el inicio de la traducción. Promotores bacterianos adecuados se describen en, por ejemplo, en Sambrook y col. y Ausubel y col.

Los sistemas de expresión bacterianos para expresar polipéptidos están disponibles en, que incluyen, *E. coli, Bacillus sp., Pseudomonas fluorescens, Pseudomonas aeruginosa, Pseudomonas putida* y *Salmonella* (Palva y col., Gene 22:229-235 (1983); Mosbach y col., Nature 302:543-545 (1983). Kits para tales sistemas de expresión están comercialmente disponibles. También están comercialmente disponibles sistemas de expresión eucariotas para células de mamífero, levadura y células de insecto. En casos en los que se usan los ARNt y aminoacil ARNt sintetasas ortogonales (descritos en cualquier parte en el presente documento) para expresar los polipéptidos, las células huésped para la expresión se seleccionan basándose en su capacidad para usar los componentes ortogonales. Células huésped a modo de ejemplo incluyen bacterias Gram-positivas (que incluyen *B. brevis* o *B. subtilis,* o *Streptomyces)* y bacterias Gram-negativas (E. *coli* o *Pseudomonas fluorescens, Pseudomonas aeruginosa, Pseudomonas putida),* además de levadura y otras células eucariotas. Las células que comprenden pares de O-ARNt/O-RS pueden usarse como se describe en el presente documento.

Una célula huésped eucariota o célula huésped no eucariota, como se describe en el presente documento, proporciona la capacidad para sintetizar polipéptidos que comprenden aminoácidos no naturales en grandes cantidades útiles. En un aspecto, la composición incluye opcionalmente, pero no se limita a, al menos 10 microgramos, al menos 50 microgramos, al menos 75 microgramos, al menos 100 microgramos, al menos 200 microgramos, al menos 250 microgramos, al menos 500 microgramos, al menos 1 miligramo, al menos 10

miligramos, al menos 100 miligramos, al menos un gramo, o más de los polipéptidos que comprenden un aminoácido no natural, o una cantidad que puede lograrse con procedimientos de producción de polipéptidos *in vivo* (los detalles sobre la producción y purificación de proteínas recombinantes se proporcionan en el presente documento). En otro aspecto, la proteína está opcionalmente presente en la composición a una concentración de, que incluye al menos 10 microgramos de polipéptido por litro, al menos 50 microgramos de polipéptido por litro, al menos 75 microgramos de polipéptido por litro, al menos 100 microgramos de polipéptido por litro, al menos 200 microgramos de polipéptido por litro, al menos 250 microgramos de polipéptido por litro, al menos 500 microgramos de polipéptido por litro, al menos 1 miligramo de polipéptido por litro, o al menos 10 miligramos de polipéptido por litro o más, en, que incluye, pero no se limita a, un lisado celular, un tampón, un tampón farmacéutico u otra suspensión líquida (que incluye en un volumen de, en cualquier parte de aproximadamente 1 nl a aproximadamente 100 l o más). La producción de grandes cantidades (que incluyen más de las normalmente posibles con otros procedimientos, que incluyen traducción *in vitro*) de una proteína en una célula eucariota que incluye al menos un aminoácido no natural es una característica de los procedimientos, técnicas y composiciones descritos en el presente documento.

Una célula huésped eucariota o célula huésped no eucariota como se describe en el presente documento proporciona la capacidad de biosintetizar proteínas que comprenden aminoácidos no naturales en grandes cantidades útiles. Por ejemplo, pueden producirse proteínas que comprenden un aminoácido no natural a una concentración de, que incluye al menos 10 µg/litro, al menos 50 µg/litro, al menos 75 µg/litro, al menos 100 µg/litro, al menos 200 µg/litro, al menos 250 µg/litro, o al menos 500 µg/litro, al menos 1 mg/litro, al menos 2 mg/litro, al menos 3 mg/litro, al menos 4 mg/litro, al menos 5 mg/litro, al menos 6 mg/litro, al menos 7 mg/litro, al menos 8 mg/litro, al menos 9 mg/litro, al menos 10 mg/litro, al menos 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900 mg/litro, 1 g/litro, 5 g/litro, 10 g/litro o más de proteína en un extracto celular, lisado celular, medio de cultivo, un tampón, y/o similares.

1. Sistemas de expresión, cultivo y aislamiento

Las técnicas desveladas en esta sección pueden aplicarse a los sistemas de expresión, cultivo y aislamiento de los polipéptidos de aminoácidos no naturales descritos en el presente documento. Los polipéptidos de aminoácidos no naturales pueden expresarse en cualquier número de sistemas de expresión adecuados, que incluyen, pero no se limitan a, levadura, células de insecto, células de mamífero y bacterias. Una descripción de sistemas de expresión a modo de ejemplo se proporciona en el presente documento.

Levadura Como se usa en el presente documento, el término “levadura” incluye cualquiera de las diversas levaduras que pueden expresar un gen que codifica el polipéptido de aminoácidos no naturales. Tales levaduras incluyen levaduras ascosporógenas (*Endomycetales*), levaduras basidiosporógenas y levaduras que pertenecen al grupo de los hongos imperfectos (*Blastomyceles*). Las levaduras ascosporógenas se dividen en dos familias, *Spermophthoraceae* y *Saccharomycetaceae*. La última comprende cuatro subfamilias, *Schizosaccharomycoideae* (por ejemplo, género *Schizosaccharomyces), Nadsonioideae, Lipomycoideae* y *Saccharomycoideae* (por ejemplo, géneros *Pichia, Kluyveromyces* y *Saccharomyces).* Las levaduras basidiosporógenas incluyen los géneros *Leucosporidium, Rhodosporidium, Sporidiobolus, Filobasidium* y *Filobasidiella.* Las levaduras que pertenecen al grupos de los hongos imperfectos *(Blastomyceles)* se dividen en dos familias, *Sporobolomycetaceae* (por ejemplo, géneros *Sporobolomyces* y *Bullera)* y *Cryptococcaceae* (por ejemplo, género *Candida).*

En ciertas realizaciones, las especies dentro de los géneros *Pichia, Kluyveromyces, Saccharomyces*, *Schizosaccharomyces, Hansenula, Torulopsis* y *Candida*, que incluyen *P. pastoris, P. guilliermondii, S. cerevisiae, S. carlsbergensis, S. diastaticus, S. douglasii, S. kluyveri, S, norbensis, S. oviformis, K. lactis, K. fragilis, C. albicans, C. maltosa* y *H. polymorpha* se usan en los procedimientos, técnicas y composiciones descritos en el presente documento.

La selección de levadura adecuada para la expresión del polipéptido de aminoácidos no naturales está dentro de la experiencia de un experto habitual en la materia. En la selección de huéspedes de levadura para la expresión, huéspedes adecuados pueden incluir aquellos que se muestra que tienen, a modo de ejemplo, buena capacidad de secreción, baja actividad proteolítica y robustez global. Las levaduras están generalmente disponibles de una variedad de fuentes que incluyen, pero no se limitan a, the Yeast Genetic Stock Center, Departamento de Biofísica y Física Médica, Universidad de California (Berkeley, CA), y la Colección Americana de Cultivos Tipo (“ATCC”) (Manassas, VA).

El término “huésped de levadura” o “célula huésped de levadura” incluye levadura que puede usarse, o se ha usado, como receptor para vectores recombinantes u otro ADN de transferencia. El término incluye la progenie de la célula huésped de levadura original que ha recibido los vectores recombinantes u otro ADN de transferencia. Se entiende que la progenie de una única célula parental puede no ser necesariamente completamente idéntica en morfología o en complemento de ADN genómico o total a la parental original, debido a mutación accidental o deliberada. La progenie de la célula parental que es suficientemente similar a la parental que va a caracterizarse por la propiedad relevante, tal como la presencia de una secuencia de nucleótidos que codifica un polipéptido de aminoácidos no naturales, está incluida en la progenie prevista por esta definición.

Se han desarrollado vectores de expresión y transformación, que incluyen replicones extracromosómicos o vectores integrantes, para la transformación en muchos huéspedes de levadura. Por ejemplo, se han desarrollado vectores de expresión para *S. cerevisiae* (Sikorski y col., Genetics (1998) 122:19; Ito y col., J. Bacteriol. (1983) 153:163; Hinnen y col., Proc. Natl. Acad. Sci. USA (1978) 75:1929); *C. albicans* (Kurtz y col., Mol. Cell. Biol. (1986) 6:142); *C. maltosa* (Kunze y col., J. Basic Microbiol. (1985) 25:141); *H. polymorpha* (Gleeson y col., J. Gen. Microbiol. (1986) 132:3459; Roggenkamp y col., Mol. Gen. Genet. (1986) 202:302); *K. fragilis* (Das y col., J. Bacteriol. (1984) 158:1165); *K. lactis* (De Louvencourt y col., J. Bacteriol. (1983) 154:737; Van den Berg y col., Bio/Technology (1990) 8:135); *P. guilliermondii* (Kunze y col., J. Basic Microbiol. (1985) 25:141); *P. pastoris* (patentes de EE.UU. nº 5.324.639; 4.929.555; y 4.837.148; Cregg y col., Mol. Cell. Biol. (1985) 5:3376); *Schizosaccharomyces pombe* (Beach y col., Nature (1982) 300:706); *A. nidulans* (Ballance y col., Biochem. Biophys. Res. Commun. (1983) 112:284-89; Tilbum y col., Gene (1983) 26:205-221; y Yelton y col., Proc. Natl. Acad. Sci. USA (1984) 81:1470-74); *A. niger* (Kelly y Hynes, EMBO J. (1985) 4:475-479); *T. reesia* (EP 0 244 234); y hongos filamentosos tales como, por ejemplo, *Neurospora*, *Penicillium*, *Tolypocladium* (documento WO 91/00357).

Secuencias de control para vectores de levadura incluyen regiones promotoras de genes tales como alcohol deshidrogenasa (ADH) (documento EP 0 284 044); enolasa; glucocinasa; glucosa-6-fosfato isomerasa; gliceraldehído-3-fosfato-deshidrogenasa (GAP o GAPDH); hexocinasa; fosfofructocinasa; 3-fosfoglicerato mutasa; y piruvato cinasa (PyK) (documento EP 0 329 203). El gen PHO5 de levadura, que codifica fosfatasa ácida, también puede proporcionar secuencias promotoras útiles (Miyanohara y col., PROC. NATL. ACAD. SCI. USA (1983) 80:1). Otras secuencias promotoras adecuadas para su uso con huéspedes de levadura pueden incluir los promotores para 3-fosfoglicerato cinasa (Hitzeman y col., J. BIOL. CHEM. (1980) 255:12073); y otras enzimas glucolíticas, tales como piruvato descarboxilasa, triosafosfato isomerasa y fosfoglucosa isomerasa (Holland y col., BIOCHEMISTRY (1978) 17:49,00; Hess y col., J. ADV. ENZYME REG. (1969) 7:149). Promotores de levadura inducibles que tienen la ventaja adicional de la transcripción controlada por las condiciones de crecimiento pueden incluir regiones promotoras para la alcohol deshidrogenasa 2; isocitocromo C; fosfatasa ácida; metalotioneína; gliceraldehído-3-fosfato deshidrogenasa; enzimas degradativas asociadas al metabolismo del nitrógeno; y enzimas responsables de la utilización de maltosa y galactosa. Vectores y promotores adecuados para su uso en la expresión de levadura se describen adicionalmente en el documento EP 0 073 657.

También pueden usarse potenciadores de levadura con promotores de levadura. Además, promotores sintéticos también pueden servir de promotores de levadura. A modo de ejemplo, las secuencias activantes en la dirección 5' (UAS) de un promotor de levadura pueden unirse con la región de activación de la transcripción de otro promotor de levadura, creando un promotor híbrido sintético. Ejemplos de tales promotores híbridos incluyen la secuencia reguladora de ADH ligada a la región de activación de la transcripción de GAP. Véanse las patentes de EE.UU. nº 4.880.734 y 4.876.197. Otros ejemplos de promotores híbridos incluyen promotores que consisten en las secuencias reguladoras de los genes ADH2, GAL4, GAL10 o PHO5, combinadas con la región de activación de la transcripción de un gen de enzima glucolítica tal como GAP o PyK. Véase el documento EP 0 164 556. Además, un promotor de levadura puede incluir promotores que se producen naturalmente de origen no levadura que tienen la capacidad de unirse a ARN polimerasa de levadura e iniciar la transcripción.

Otros elementos de control que pueden comprender parte de los vectores de expresión en levadura incluyen terminadores, por ejemplo, de GAPDH o los genes enolasa (Holland y col., J. BIOL. CHEM. (1981) 256:1385). Además, el origen de replicación del origen de plásmido 2µ es adecuado para levadura. Un gen de selección adecuado para su uso en levadura es el gen *trpl* presente en el plásmido de levadura. Véanse Tschumper y col., GENE (1980) 10:157; Kingsman y col., GENE (1979) 7:141. El gen *trpl* proporciona un marcador de selección para una cepa mutante de levadura que carece de la capacidad para crecer en triptófano. Similarmente, cepas de levadura deficientes en Leu2 (ATCC 20.622 ó 38.626) se complementan por plásmidos conocidos que llevan el gen *Leu2*.

Procedimientos de introducción de ADN exógeno en huéspedes de levadura incluyen, pero no se limitan a, tanto la transformación de esferoplastos como de células huésped de levadura intactas tratadas con cationes alcalinos. A modo de ejemplo, la transformación de levadura puede llevarse a cabo según el procedimiento descrito en Hsiao y col., PROC. NATL. ACAD. SCI. USA (1979) 76:3829 y Van Solingen y col., J. BACT. (1977) 130:946. Sin embargo, también pueden usarse otros procedimientos para introducir ADN en células, tales como por inyección nuclear, electroporación o fusión de protoplastos, como se describe generalmente en SAMBROOK Y COL., MOLECULAR CLONING: A LAB. MANUAL (2001). A continuación, las células huésped de levadura pueden cultivarse usando técnicas convencionales conocidas para aquellos expertos habituales en la materia.

Otros procedimientos para expresar proteínas heterólogas en células huésped de levadura son muy conocidos para aquellos expertos habituales en la materia. Véanse generalmente la publicación de patente de EE.UU. nº 20020055169, las patentes de EE.UU. nº 6.361.969; 6.312.923; 6.183.985; 6.083.723; 6.017.731; 5.674.706; 5.629.203; 5.602.034; y 5.089.398; patentes reexaminadas de EE.UU. nº RE37.343 y RE35.749; solicitudes de patente publicadas PCT WO 99/07862; WO 98/37208; y WO 98/26080; solicitudes de patente europea EP 0 946 736; EP 0 732 403; EP 0 480 480; WO 90/10277; EP 0 340 986; EP 0 329 203; EP 0 324 274; y EP 0 164 556. Véanse también Gellissen y col., ANTONIE VAN LEEUWENHOEK (1992) 62(1-2):79-93; Romanos y col., YEAST

(1992) 8(6):423-488; Goeddel, METHODS IN ENZYMOLOGY (1990) 185:3-7.

La cepas huésped de levadura pueden cultivarse en fermentadores durante la etapa de amplificación usando procedimientos de fermentación de alimentación discontinua convencionales. Los procedimientos de fermentación pueden adaptarse para explicar diferencias en una ruta de utilización de carbono del huésped de levadura particular o modo de control de la expresión. A modo de ejemplo, la fermentación de un huésped de levadura de *Saccharomyces* puede requerir una única alimentación de glucosa, fuente de nitrógeno compleja (por ejemplo, hidrolizados de caseína) y suplementación de múltiples vitaminas, mientras que la levadura metilotrófica *P. pastoris* puede requerir glicerol, metanol y alimentaciones de oligominerales, pero solo sales de amonio simple (nitrógeno) para el óptimo crecimiento y expresión. Véanse, por ejemplo, la patente de EE.UU. nº 5.324.639; Elliott y col., J. PROTEIN CHEM. (1990) 9:95; y Fieschko y col., BIOTECH. BIOENG. (1987) 29:1113.

Sin embargo, tales procedimientos de fermentación pueden tener ciertas características comunes independientes de la cepa de huésped de levadura empleada. A modo de ejemplo, un nutriente limitante del crecimiento, normalmente carbono, puede añadirse al fermentador durante la fase de amplificación para permitir el máximo crecimiento. Además, los procedimientos de fermentación generalmente emplean un medio de fermentación diseñado para contener cantidades adecuadas de carbono, nitrógeno, sales basales, fósforo y otros nutrientes minoritarios (vitaminas, oligominerales y sales, etc.). Ejemplos de medios de fermentación adecuados para su uso con *Pichia* se describen en las patentes de EE.UU. nº 5.324.639 y 5.231.178.

Células de insecto infectadas con baculovirus El término “huésped de insecto” o “célula huésped de insecto” se refiere a un insecto que puede usarse, o se ha usado, como receptor para vectores recombinantes u otro ADN de transferencia. El término incluye la progenie de la célula huésped de insecto original que se ha transfectado. Se entiende que la progenie de una única célula parental puede no ser necesariamente completamente idéntica en morfología o en complemento de ADN genómico o total a la parental original, debido a mutación accidental o deliberada. La progenie de la célula parental que es suficientemente similar a la parental que va a caracterizarse por la propiedad relevante, tal como la presencia de una secuencia de nucleótidos que codifica un polipéptido de aminoácidos no naturales, está incluida en la progenie prevista por esta definición.

La selección de células de insecto adecuadas para la expresión de un polipéptido es muy conocida para aquellos expertos habituales en la materia. Varias especies de insectos están bien descritas en la materia y están comercialmente disponibles, que incluyen, pero no se limitan a, *Aedes aegypti, Bombyx mori, Drosophila melanogaster, Spodoptera frugiperda* y *Trichoplusia ni.* En la selección de huéspedes de insecto para la expresión, huéspedes adecuados pueden incluir aquellos que se muestra que tienen, entre otras cosas, buena capacidad de secreción, baja actividad proteolítica y robustez global. Los insectos están generalmente disponible de una variedad de fuentes que incluyen, the Insect Genetic Stock Center, Departamento de Biofísica y Física Médica, Universidad de California (Berkeley, CA); y la Colección Americana de Cultivos Tipo (“ATCC”) (Manassas, VA).

Generalmente, los componentes de un sistema de expresión de insecto infectado con baculovirus incluyen un vector de transferencia, normalmente un plásmido bacteriano, que contiene tanto un fragmento del genoma de baculovirus como un sitio de restricción conveniente para la inserción del gen heterólogo que va a expresarse; un baculovirus natural con una secuencia homóloga al fragmento específico para baculovirus en el vector de transferencia (esto permite la recombinación homóloga del gen heterólogo en el genoma de baculovirus); y células huésped de insecto apropiadas y medios de crecimiento. Los materiales, procedimientos y técnicas usados en construir vectores, transfectar células, recoger placas, cultivar células en cultivo y similares se conocen en la técnica y están disponibles manuales para describir estas técnicas.

Después de insertar el gen heterólogo en el vector de transferencia, el vector y el genoma viral natural se transfectan en una célula huésped de insecto en la que se recombinan el vector y el genoma viral. Se expresa el virus recombinante encapsidado y se identifican y purifican placas recombinantes. Materiales y procedimientos para los sistemas de expresión de baculovirus/células de insecto están comercialmente disponibles en forma de kit de, por ejemplo, Invitrogen Corp. (Carlsbad, CA). Estas técnicas se describen en SUMMERS Y SMITH, TEXAS AGRICULTURAL EXPERIMENT STATION BULLETIN No. 1555 (1987), incorporado en el presente documento por referencia. Véanse también RICHARDSON, 39 METHODS IN MOLECULAR BIOLOGY: BACULOVIRUS EXPRESSION PROTOCOLS (1995); AUSUBEL Y COL., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY 16.9-16.11 (1994); KING Y POSSEE, THE BACULOVIRUS SYSTEM: A LABORATORY GUIDE (1992); y O'REILLY Y COL., BACULOVIRUS EXPRESSION VECTORS: A LABORATORY MANUAL (1992).

La producción de diversas proteínas heterólogas usando sistemas de expresión de baculovirus/células de insecto se describe en las siguientes referencias y tales técnicas pueden adaptarse para producir los polipéptidos de aminoácidos no naturales descritos en el presente documento. Véanse, por ejemplo, las patentes de EE.UU. nº 6.368.825; 6.342.216; 6.338.846; 6.261.805; 6.245.528; 6.225.060; 6.183.987; 6.168.932; 6.126.944; 6.096.304; 6.013.433; 5.965.393; 5.939.285; 5.891.676; 5.871.986; 5.861.279; 5.858.368; 5.843.733; 5.762.939; 5.753.220; 5.605.827; 5.583.023; 5.571.709; 5.516.657; 5.290.686; WO 02/06305; WO 01/90390; WO 01/27301; WO 01/05956; WO 00/55345; WO 00/20032 WO 99/51721; WO 99/45130; WO 99/31257; WO 99/10515; documento WO 99/09193; WO 97/26332; WO 96/29400; WO 96/25496; WO 96/06161; WO 95/20672; WO 93/03173; WO 92/16619; WO

92/02628; WO 92/01801; WO 90/14428; WO 90/10078; WO 90/02566; WO 90/02186; WO 90/01556; WO 89/01038; WO 89/01037; WO 88/07082.

Se conocen vectores que son útiles en sistemas de expresión de baculovirus/células de insecto e incluyen vectores de expresión y de transferencia de insecto derivados del virus de la poliedrosis nuclear de *Autographa californica* (AcNPV) de baculovirus, que es un vector de expresión viral independiente de colaborador. Vectores de expresión virales derivados de este sistema normalmente usan el promotor viral fuerte del gen de la poliedrina para dirigir la expresión de genes heterólogos. Véase generalmente O'Reilly Y COL., BACULOVIRUS EXPRESSION VECTORS: A LABORATORY MANUAL (1992).

Antes de insertar el gen extraño en el genoma del baculovirus, los componentes anteriormente descritos, que comprenden un promotor, conductor (si se desea), secuencia codificante de interés y secuencia de terminación de la transcripción, se ensamblan normalmente en una construcción de sustitución intermedia (vector de transferencia). Las construcciones de sustitución intermedia se mantienen frecuentemente en un replicón, tales como un elemento extracromosómico (por ejemplo, plásmidos) que puede mantenerse establemente en un huésped, tal como bacterias. El replicón tendrá un sistema de replicación, permitiendo así que se mantenga en un huésped adecuado para la clonación y amplificación. Más específicamente, el plásmido puede contener la señal de poliadenilación poliédrica (Miller, ANN. REV. MICROBIOL. (1988) 42:177) y un gen procariota de resistencia a ampicilina (*amp*) y origen de replicación para la selección y propagación en *E. coli.*

Un vector de transferencia comúnmente usado para introducir genes extraños en AcNPV es pAc373. También se han diseñado muchos otros vectores, conocidos para aquellos expertos en la materia, que incluyen, por ejemplo, pVL985, que altera el codón de iniciación de poliedrina de ATG a ATT, y que introduce un sitio de clonación BamHI de 32 pares de bases en la dirección 3' de ATT. Véase Luckow y Summers, VIROLOGY 170:31 (1989). Otros vectores comercialmente disponibles incluyen, por ejemplo, PBlueBac4.5/V5-His; pBlueBacHis2; pMelBac; pBlueBac4.5 (Invitrogen Corp., Carlsbad, CA).

Después de la inserción del gen heterólogo, el vector de transferencia y genoma baculoviral natural se co-transfectan en un huésped de célula de insecto. Procedimientos ilustrativos para introducir ADN heterólogo en el sitio deseado en el virus de baculovirus se describen en SUMMERS Y SMITH, TEXAS AGRICULTURAL EXPERIMENT STATION BULLETIN NO. 1555 (1987); Smith y col., MOL. CELL. BIOL. (1983) 3:2156; Luckow y Summers, VIROLOGY (1989) 170:31. A modo de ejemplo, la inserción puede ser en un gen tal como el gen de la poliedrina, por recombinación cruzada doble homóloga; la inserción también puede ser en un sitio de enzima de restricción manipulado en el gen de baculovirus deseado. Véase Miller y col., BIOESSAYS (1989) 11(4):91.

La transfección puede llevarse a cabo por electroporación usando procedimientos descritos en TROTTER Y WOOD, 39 METHODS IN MOLECULAR BIOLOGY (1995); Mann and King, J.-GEN.-VIROL. (1989) 70:3501. Alternativamente, pueden usarse liposomas para transfectar las células de insecto con el vector de expresión recombinante y el baculovirus. Véanse, por ejemplo, Liebman y col., BIOTECHNIQUES (1999) 26(1):36; Graves y col., BIOCHEMISTRY (1998) 37:6050; Nomura y col., J. BIOL. CHEM. (1998) 273(22):13570; Schmidt y col., PROTEIN EXPRESSION AND PURIFICATION (1998) 12:323; Siffert y col., NATURE GENETICS (1998) 18:45; TILKINS Y COL., CELL BIOLOGY: A LABORATORY HANDBOOK 145-154 (1998); Cai y col., PROTEIN EXPRESSION AND PURIFICATION (1997) 10:263; Dolphin y col., NATURE GENETICS (1997) 17:491; Kost y col., GENE (1997) 190:139; Jakobsson y col., J. BIOL. CHEM. (1996) 271:22203; Rowles y col., J. BIOL. CHEM. (1996) 271(37):22376; Reverey y col., J. BIOL. CHEM. (1996) 271(39):23607-10; Stanley y col., J. BIOL. CHEM. (1995) 270:4121; Sisk y col., J. VIROL. (1994) 68(2):766; y Peng y col., BIOTECHNIQUES (1993) 14(2):274. Liposomas comercialmente disponibles incluyen, por ejemplo, Cellfectin® y Lipofectin® (Invitrogen, Corp., Carlsbad, CA). Además, puede usarse transfección con fosfato de calcio. Véanse TROTTER Y WOOD, 39 METHODS IN MOLECULAR BIOLOGY (1995); Kitts, NAR (1990) 18(19):5667; y Mann and King, J. GEN. VIROL. (1989) 70:3501.

Los vectores de expresión de baculovirus normalmente contienen un promotor de baculovirus. Un promotor de baculovirus es cualquier secuencia de ADN que pueda unirse a una ARN polimerasa de baculovirus e iniciar la transcripción aguas abajo (3') de una secuencia codificante (por ejemplo, gen estructural) en ARNm. Un promotor tendrá una región de iniciación de la transcripción que está normalmente ubicada proximal al extremo 5' de la secuencia codificante. Esta región de iniciación de la transcripción normalmente incluye un sitio de unión de ARN polimerasa y un sitio de iniciación de la transcripción. Un promotor de baculovirus también puede tener un segundo dominio llamado un potenciador, que, si está presente, está normalmente distal al gen estructural. Además, la expresión puede ser tanto regulada como constitutiva.

Genes estructurales abundantemente transcritos en momentos tardíos en el ciclo de infección proporcionan secuencias promotoras particularmente útiles. Ejemplos incluyen secuencias derivadas del gen que codifica la proteína poliédrica viral (FRIESEN Y COL., The Regulation of Baculovirus Gene Expression in THE MOLECULAR BIOLOGY OF BACULOVIRUSES (1986); documentos EP 0 127 839 y 0 155 476) y el gen que codifica la proteína p10 (Vlak y col., J. GEN. VIROL. (1988) 69:765).

El vector de expresión de baculovirus recientemente formado se encapsida en un baculovirus recombinante

infeccioso y posteriormente pueden purificarse placas cultivadas por técnicas tales como aquellas descritas en Miller y col., BIOESSAYS (1989) 11(4):91; SUMMERS Y SMITH, TEXAS AGRICULTURAL EXPERIMENT STATION BULLETIN No. 1555 (1987).

Se han desarrollado vectores de expresión de baculovirus recombinantes para la infección en varias células de insecto. Por ejemplo, se han desarrollado baculovirus recombinantes para, entre otros, *Aedes aegypti* (ATCC Nº CCL-125), *Bombyx mori* (ATCC Nº CRL-8910), *Drosophila melanogaster* (ATCC Nº 1963), *Spodoptera frugiperda* y *Trichoplusia ni.* Véanse Wright, NATURE (1986) 321:718; Carbonell y col., J. VIROL. (1985) 56:153; Smith y col., MOL. CELL. BIOL. (1983) 3:2156. Véase generalmente Fraser y col., IN VITRO CELL. DEV. BIOL. (1989) 25:225. Más específicamente, las líneas celulares usadas para sistemas de vectores de expresión de baculovirus comúnmente incluyen, pero no se limitan a, Sf9 (*Spodoptera frugiperda*) (ATCC Nº CRL-1711), Sf21 (*Spodoptera frugiperda*) (Invitrogen Corp., nº de cat. 11497-013 (Carlsbad, CA)), Tri-368 (*Trichoplusia ni*) y High-Five™ BTI-TN-5B1-4 (*Trichoplusia ni*).

Están comercialmente disponibles células y medios de cultivo para tanto la expresión directa como de fusión de polipéptidos heterólogos en un baculovirus/expresión.

E. coli, Pseudomonas y otros procariotas: Las técnicas de expresión bacteriana son muy conocidas en la técnica. Están disponibles una amplia variedad de vectores para su uso en huéspedes bacterianos. Los vectores pueden ser vectores de una única copia o de multicopia alta o baja. Los vectores pueden servir para clonación y/o expresión. En vista de la amplia bibliografía referente a vectores, disponibilidad comercial de muchos vectores, e incluso manuales que describen vectores y sus mapas y características de restricción, no se requiere aquí una amplia discusión. Como es muy conocido, los vectores normalmente implican marcadores que permiten la selección, marcadores que pueden proporcionar resistencia a agentes citotóxicos, prototrofia o inmunidad. Frecuentemente, están presentes una pluralidad de marcadores, que proporcionan diferentes características.

Un promotor bacteriano es cualquier secuencia de ADN que puede unirse a ARN polimerasa bacteriana e iniciar la transcripción aguas abajo (3') de una secuencia codificante (por ejemplo, gen estructural) en ARNm. Un promotor tendrá una región de iniciación de la transcripción que está normalmente ubicada proximal al extremo 5' de la secuencia codificante. Este región de iniciación de la transcripción normalmente incluye un sitio de unión de ARN polimerasa y un sitio de iniciación de la transcripción. Un promotor bacteriano también puede tener un segundo dominio llamado un operador, que puede solapar un sitio de unión de ARN polimerasa adyacente en el que empieza la síntesis del ARN. El operador permite la transcripción regulada negativa (inducible), ya que una proteína represora de genes puede unirse al operador y así inhibir la transcripción de un gen específico. Puede producirse expresión constitutiva en ausencia de elementos reguladores negativos, tales como el operador. Además, puede lograrse regulación positiva por una secuencia de unión de proteínas activadoras de genes, que, si está presente, es normalmente proximal (5') a la secuencia de unión de ARN polimerasa. Un ejemplo de una proteína activadora de genes es la proteína activadora de catabolitos (CAP), que ayuda a iniciar la transcripción del operón lac en *Escherichia coli* (*E. coli*) [Raibaud y col., ANNU. REV. GENET. (1984) 18:173]. Por tanto, la expresión regulada puede ser tanto positiva como negativa, tanto potenciando así como reduciendo la transcripción.

Secuencias que codifican enzimas de rutas metabólicas proporcionan secuencias promotoras particularmente útiles. Ejemplos incluyen secuencias promotoras derivadas de enzimas metabolizantes de azúcares, tales como galactosa, lactosa (lac) [Chang y col., NATURE (1977) 198:1056] y maltosa. Ejemplos adicionales incluyen secuencias promotoras derivadas de enzimas biosintéticas tales como triptófano (trp) [Goeddel y col., NUC. ACIDS RES. (1980) 8:4057; Yelverton y col., NUCL. ACIDS RES. (1981) 9:731; patente de EE.UU. nº 4.738.921; IFNPub. Nos. 036 776 y 121 775]. Los sistemas promotores el sistema de promotor de β-galactosidasa (bla) [Weissmann (1981) “The cloning of interferon and other mistakes”. En Interferon 3 (Ed. I. Gresser)], bacteriófago lambda PL [Shimatake y col., NATURE (1981) 292:128] y T5 [patente de EE.UU. nº 4.689.406]] también proporcionan secuencias promotoras útiles. Procedimientos preferidos englobados en el presente documento utilizan promotores fuertes, tales como el promotor T7 para inducir la producción de polipéptidos a altos niveles. Ejemplos de tales vectores incluyen, pero no se limitan a, las series pET29 de Novagen y los vectores pPOP descritos en el documento WO99/05297. Tales sistemas de expresión producen altos niveles de polipéptido en el huésped sin comprometer la viabilidad celular del huésped o los parámetros de crecimiento.

Además, promotores sintéticos que no se producen en la naturaleza también funcionan como promotores bacterianos. Por ejemplo, secuencias de activación de la transcripción de un promotor bacteriano o de bacteriófago pueden unirse con las secuencias de operón de otro promotor bacteriano o de bacteriófago, creando un promotor híbrido sintético [patente de EE.UU. nº 4.551.433]. Por ejemplo, el promotor tac es un promotor híbrido de trp-lac que comprende tanto las secuencias del promotor trp como del operón lac que está regulado por el represor lac [Amann y col., GENE (1983) 25:167; de Boer y col., PROC. NATL. ACAD. SCI. (1983) 80:21]. Además, un promotor bacteriano puede incluir promotores que se producen naturalmente de origen no bacteriano que tienen la capacidad de unirse a ARN polimerasa bacteriana e iniciar la transcripción. Un promotor que se produce naturalmente de origen no bacteriano también puede acoplarse con una ARN polimerasa compatible para producir altos niveles de expresión de algunos genes en procariotas. El sistema de ARN polimerasa del bacteriófago T7/promotor es un ejemplo de un sistema de promotores acoplados [Studier y col., J. MOL. BIOL. (1986) 189:113; Tabor y col., Proc

Natl. Acad. Sci. (1985) 82:1074]. Además, un promotor híbrido también puede estar comprendido de un promotor de bacteriófago y una región de operador de *E. coli* (IFNPub. No. 267 851).

Además de una secuencia promotora de funcionamiento, también es útil un sitio de unión al ribosoma eficaz para la expresión de genes extraños en procariotas. En *E. coli*, el sitio de unión al ribosoma se llama la secuencia de Shine-Dalgarno (SD) e incluye un codón de iniciación (ATG) y una secuencia de 3-9 nucleótidos en la longitud localizada 3-11 nucleótidos en la dirección 5' del codón de iniciación [Shine y col., NATURE (1975) 254:34]. Se cree que la secuencia de SD promueve la unión del ARNm al ribosoma por el apareamiento de bases entre la secuencia de SD y el 3' y de 16S ARNr de *E. coli* [Steitz y col. "Genetic signals and nucleotide sequences in messenger RNA", en Biological Regulation and Development: Gene Expression (Ed. R. F. Goldberger, 1979)]. Para expresar genes eucariotas y genes procariotas con sitio de unión al ribosoma débil [Sambrook y col. "Expression of cloned genes in Escherichia coli", Molecular Cloning: A Laboratory Manual, 1989].

El término "huésped bacteriano" o "célula huésped bacteriana" se refiere a una bacteria que puede usarse, o se ha usado, como receptor para vectores recombinantes u otro ADN de transferencia. El término incluye la progenie de la célula huésped bacteriana original que se ha transfectado. Se entiende que la progenie de una única célula parental puede no ser necesariamente completamente idéntica en morfología o en complemento de ADN genómico o total a la parental original, debido a mutación accidental o deliberada. La progenie de la célula parental que es suficientemente similar a la parental que va a caracterizarse por la propiedad relevante, tal como la presencia de una secuencia de nucleótidos que codifica un polipéptido, está incluida en la progenie prevista por esta definición.

La selección de bacterias huésped adecuadas para la expresión de un polipéptido es muy conocida para aquellos expertos habituales en la materia. En la selección de huéspedes bacterianos para la expresión, huéspedes adecuados pueden incluir aquellos que se muestra que tienen, entre otras cosas, buena capacidad de formación de cuerpos de inclusión, baja actividad proteolítica, buena capacidad de secreción, buena capacidad de producción de proteínas solubles y robustez global. Están generalmente disponibles huéspedes bacterianos de una variedad de fuentes que incluyen, pero no se limitan a, the Bacterial Genetic Stock Center, Departamento de Biofísica y Física Médica, Universidad de California (Berkeley, CA); y la Colección Americana de Cultivos Tipo ("ATCC") (Manassas, VA). La fermentación industrial/farmacéutica generalmente usa bacterias derivadas de cepas K (por ejemplo, W3110) o de bacterias derivadas de cepas B (por ejemplo, BL21). Estas cepas son particularmente útiles debido a que sus parámetros de crecimiento son extremadamente muy conocidos y robustos. Además, estas cepas no son patógenas, que es comercialmente importante por motivos de seguridad y medioambientales. En una realización de los procedimientos descritos y englobados en el presente documento, el huésped de *E. coli* incluye, pero no se limita a, cepas de BL21, DH10B, o derivados de las mismas. En otra realización de los procedimientos descritos y englobados en el presente documento, el huésped de *E. coli* es una cepa de proteasa menos que incluye OMP- y LON-. En otra realización, el huésped bacteriano es una especie de *Pseudomonas,* tal como *P. fluorescens, P. aeruginosa* y *P. putida.* Un ejemplo de una cepa de *Pseudomonas* es *P. fluorescens* biovar I, cepa MB101 (Dow Chemical).

Una vez se ha establecido una cepa de células huésped recombinantes (es decir, la construcción de expresión se ha introducido en la célula huésped y las células huésped se aíslan con la construcción de expresión apropiada), la cepa de células huésped recombinantes se cultiva en condiciones apropiadas para la producción de polipéptidos. El procedimiento de cultivo de la cepa de células huésped recombinantes dependerá de la naturaleza de la construcción de expresión utilizada y la identidad de la célula huésped. Las cepas huésped recombinantes normalmente se cultivan usando procedimientos que son muy conocidos para la materia. Las células huésped recombinantes normalmente se cultivan en medio líquido que contiene fuentes de carbono asimilables, nitrógeno y sales inorgánicas y, opcionalmente, que contiene vitaminas, aminoácidos, factores de crecimiento y otros suplementos de cultivo proteináceos muy conocidos para la materia. Los medios líquidos para el cultivo de células huésped pueden contener opcionalmente antibióticos o antifúngicos para prevenir el crecimiento de microorganismos no deseables y/o compuestos que incluyen, pero no se limitan a, antibióticos para seleccionar células huésped que contiene el vector de expresión.

Las células huésped recombinantes pueden cultivarse en formatos discontinuo o continuo, con tanto recogida de células (en el caso en el que el polipéptido se acumule intracelularmente) como recogida de sobrenadante de cultivo en formatos tanto discontinuos como continuos. Para la producción en células huésped procariotas, se prefieren cultivo discontinuo y recogida de células.

En una realización, los polipéptidos de aminoácidos no naturales descritos en el presente documento se purifican después de la expresión en sistemas recombinantes. Los polipéptidos pueden purificarse a partir de células huésped o medio de cultivo mediante una variedad de procedimientos conocidos en la técnica. Normalmente, muchos polipéptidos producidos en células huésped bacterianas pueden ser poco solubles o insolubles (en forma de cuerpos de inclusión). En una realización, pueden hacerse fácilmente sustituciones de aminoácidos en los polipéptidos que están seleccionadas con el fin de aumentar la solubilidad del polipéptido recombinantemente producido utilizando los procedimientos desvelados en el presente documento, además de aquellos conocidos en la técnica. En el caso de polipéptido insoluble, el polipéptido puede recogerse de lisados de células huésped por centrifugación o filtrado y puede ir adicionalmente seguido de homogenización de las células. En el caso de polipéptido poco soluble,

compuestos que incluyen, pero no se limitan a, polietilenimina (PEI) pueden añadirse para inducir la precipitación de polipéptido parcialmente soluble. La proteína precipitada puede a continuación recogerse convenientemente por centrifugación o filtrado. Las células huésped recombinantes pueden romperse u homogeneizarse para liberar los cuerpos de inclusión de dentro de las células usando una variedad de procedimientos muy conocidos para aquellos expertos habituales en la materia. La rotura u homogenización de células huésped puede realizarse usando técnicas muy conocidas que incluyen, pero no se limitan a, rotura enzimática de células, sonicación, homogenización por Dounce, o rotura por liberación a alta presión. En una realización de los procedimientos descritos y englobados en el presente documento, la técnica de liberación a alta presión se usa para romper las células huésped de *E. coli* para liberar los cuerpos de inclusión de los polipéptidos. Cuando se manipulan cuerpos de inclusión de polipéptidos, es ventajoso minimizar el tiempo de homogenización en las repeticiones con el fin de maximizar el rendimiento de los cuerpos de inclusión sin pérdida debido a factores tales como solubilización, cizallamiento mecánico o proteólisis.

A continuación, polipéptidos insolubles o precipitados pueden solubilizarse usando cualquiera de varios agentes de solubilización adecuados conocidos en la técnica. A modo de ejemplo, los polipéptidos se solubilizan con urea o clorhidrato de guanidina. El volumen de los polipéptidos solubilizados debe minimizarse de manera que puedan producirse grandes lotes usando tamaños de lote convenientemente manejables. Este factor puede ser significativo en un marco comercial a gran escala en el que el huésped recombinante puede cultivarse en lotes que son miles de litros de volumen. Además, cuando se fabrican polipéptidos en un marco comercial a gran escala, en particular para usos farmacéuticos humanos, el evitar productos químicos agresivos que pueden dañar la maquinaria y el recipiente, o el propio producto de proteína, debe evitarse, si es posible. Se ha mostrado en los procedimientos descritos y englobados en el presente documento que el agente desnaturalizante más suave urea puede usarse para solubilizar los cuerpos de inclusión de polipéptido en lugar del agente desnaturalizante más agresivo clorhidrato de guanidina. El uso de urea reduce significativamente el riesgo de daño al equipo de acero inoxidable utilizado en el procedimiento de fabricación y purificación de un polipéptido mientras que solubiliza eficazmente los cuerpos de inclusión de polipéptidos.

En el caso de polipéptidos solubles, los péptidos pueden secretarse en el espacio periplásmico o en el medio de cultivo. Además, pueden estar presentes péptidos solubles en el citoplasma de las células huésped. El péptido soluble puede concentrarse antes para realizar las etapas de purificación. Pueden usarse técnicas convencionales, que incluyen aquellas descritas en el presente documento, para concentrar péptido soluble de, a modo de ejemplo, lisados celulares o medio de cultivo. Además, pueden usarse técnicas convencionales, que incluyen aquellas descritas en el presente documento, para romper células huésped y liberar péptido soluble del citoplasma o espacio periplásmico de las células huésped.

Cuando el polipéptido se produce como una proteína de fusión, la secuencia de fusión se elimina preferentemente. La eliminación de una secuencia de fusión puede llevarse a cabo mediante procedimientos que incluyen escisión enzimática o química. La eliminación enzimática de secuencias de fusión puede llevarse a cabo usando procedimientos muy conocidos para aquellos en la materia. La elección de enzima para la eliminación de la secuencia de fusión se determinará por la identidad de la fusión y las condiciones de reacción se especificarán por la elección de enzima. La escisión química puede llevarse a cabo usando reactivos, que incluyen, pero no se limitan a, bromuro de cianógeno, TEV proteasa y otros reactivos. El polipéptido escindido se purifica opcionalmente a partir de la secuencia de fusión escindida por procedimientos muy conocidos. Tales procedimientos se determinarán por la identidad y propiedades de la secuencia de fusión y el polipéptido. Procedimientos para la purificación pueden incluir cromatografía de exclusión por tamaño, cromatografía de interacción hidrófoba, cromatografía de intercambio iónico o diálisis, o cualquier combinación de las mismas.

El polipéptido también se purifica opcionalmente para eliminar ADN de la disolución de proteína. El ADN puede eliminarse por cualquier procedimiento adecuado conocido en la técnica, que incluye precipitación o cromatografía de intercambio iónico. En una realización, el ADN se elimina mediante precipitación con un agente precipitante de ácido nucleico, tal como, pero no se limita a, sulfato de protamina. El polipéptido puede separarse del ADN precipitado usando procedimientos estándar muy conocidos que incluyen centrifugación o filtración. La eliminación de moléculas de ácidos nucleicos huésped es un factor importante en un entorno en el que el polipéptido va a usarse para tratar seres humanos y los procedimientos descritos en el presente documento reducen el ADN de células huésped a niveles farmacéuticamente aceptables.

También pueden usarse procedimientos para fermentación a pequeña escala o gran escala en la expresión de proteínas que incluyen, pero no se limitan a, fermentadores, matraces oscilantes, biorreactores de lecho fluidizado, biorreactores de fibra hueca, sistemas de cultivo de botellas rotatorias y sistemas de biorreactor de tanques con agitación. Cada uno de estos procedimientos puede realizarse en un procedimiento en modo discontinuo, de alimentaciones discontinuas o continuo.

Las formas humanas de los polipéptidos de aminoácidos no naturales descritos en el presente documento pueden generalmente recuperarse usando procedimientos convencionales en la materia. Por ejemplo, el medio de cultivo o lisado celular puede centrifugarse o filtrarse para eliminar residuo celular. El sobrenadante puede concentrarse o diluirse a un volumen deseado o diafiltrarse en un tampón adecuado para acondicionar la preparación para la purificación adicional. La purificación adicional de los polipéptidos de aminoácidos no naturales descritos en el

presente documento incluye separar formas desamidadas y cortadas de una variante de polipéptido de la forma intacta correspondiente.

Cualquiera de los siguientes procedimientos a modo de ejemplo puede emplearse para la purificación de un polipéptido de aminoácidos no naturales descrito en el presente documento: cromatografía de afinidad; cromatografía de intercambio aniónico o catiónico (usando, que incluye DEAE-SEPHAROSE); cromatografía sobre sílice; HPLC de fase inversa; filtración en gel (usando, que incluye, pero no se limita a, SEPHADEX G-75); cromatografía de interacción hidrófoba; cromatografía de exclusión por tamaño, cromatografía con quelato de metal; ultrafiltración/diafiltración; precipitación con etanol; precipitación con sulfato de amonio; cromatoenfoque; cromatografía de desplazamiento; procedimientos electroforéticos (que incluyen enfoque isoeléctrico preparativo), solubilidad diferencial (que incluye, pero no se limita a, precipitación con sulfato de amonio), SDS-PAGE o extracción.

Los polipéptidos englobados dentro de los procedimientos y composiciones descritos en el presente documento, que incluye polipéptidos que comprenden aminoácidos no naturales, anticuerpos para polipéptidos que comprenden aminoácidos no naturales, componentes de unión para polipéptidos que comprenden aminoácidos no naturales, etc., pueden purificarse, tanto parcialmente como sustancialmente a homogeneidad, según procedimientos convencionales conocidos para y usados por aquellos expertos en la materia. Por consiguiente, los polipéptidos descritos en el presente documento pueden recuperarse y purificarse por cualquiera de varios procedimientos muy conocidos en la técnica, que incluyen precipitación con sulfato de amonio o etanol, extracción con ácido o base, cromatografía en columna, cromatografía en columna de afinidad, cromatografía de intercambio aniónico o catiónico, cromatografía en fosfocelulosa, cromatografía de interacción hidrófoba, cromatografía en hidroxilapatita, cromatografía en lectina, electroforesis en gel y cualquier combinación de los mismos. Pueden usarse etapas de replegamiento de proteínas, según se desee, en la preparación de proteínas maduras correctamente plegadas. Pueden emplearse cromatografía líquida de alta resolución (HPLC), cromatografía de afinidad u otros procedimientos adecuados en etapas de purificación finales en los que se desea alta pureza. En una realización, anticuerpos preparados contra aminoácidos no naturales (o polipéptidos que comprenden aminoácidos no naturales) se usan como reactivos de purificación, que incluye purificación de polipéptidos basada en afinidad que comprende uno o más aminoácidos no naturales). Una vez purificados, parcialmente o a homogeneidad, según se desee, los polipéptidos se usan opcionalmente para una amplia variedad de utilidades, que incluyen como componentes de ensayo, terapéuticos, profilaxis, diagnóstico, reactivos de investigación, y/o como inmunógenos para la producción de anticuerpos.

Además de otras referencias indicadas en el presente documento, una variedad de procedimientos de purificación/plegamiento de proteínas son muy conocidos en la técnica, que incluyen aquellos expuestos R. Scopes, Protein Purification, Springer-Verlag, N.Y. (1982); Deutscher, Methods in Enzymology Vol. 182: Guide to Protein Purification, Academic Press, Inc. N.Y. (1990); Sandana (1997) Bioseparation of Proteins, Academic Press, Inc.; Bollag y col. (1996) Protein Methods, 2ª edición Wiley-Liss, NY; Walker (1996) The Protein Protocols Handbook Humana Press, NJ, Harris and Angal (1990) Protein Purification Applications: A Practical Approach IRL Press at Oxford, Oxford, England; Harris and Angal Protein Purification Methods: A Practical Approach IRL Press at Oxford, Oxford, England; Scopes (1993) Protein Purification: Principles and Practice 3ª edición Springer Verlag, NY; Janson and Ryden (1998) Protein Purification: Principles. High Resolution Methods and Applications, Second Edition Wiley-VCH, NY; y Walker (1998) Protein Protocols on CD-ROM Humana Press, NJ; y las referencias citadas en su interior.

Una ventaja de producir polipéptidos que comprenden al menos un aminoácido no natural en una célula huésped eucariota o célula huésped no eucariota es que normalmente los polipéptidos se plegarán en sus conformaciones nativas. Sin embargo, en ciertas realizaciones de los procedimientos y composiciones descritos en el presente documento, después de la síntesis, expresión y/o purificación, los polipéptidos pueden poseer una conformación diferente de las conformaciones deseadas de los polipéptidos relevantes. En un aspecto de los procedimientos y composiciones descritos en el presente documento, la proteína expresada se desnaturaliza opcionalmente y a continuación se renaturaliza. Esta desnaturalización y renaturalización opcional se lleva a cabo utilizando procedimientos conocidos en la técnica, que incluyen añadir una chaperonina al polipéptido de interés, y solubilizar las proteínas en un agente caotrópico que incluye HCl de guanidina, y utilizar la proteína disulfuro isomerasa.

En general, es ocasionalmente deseable desnaturalizar y reducir los polipéptidos expresados y a continuación producir los polipéptidos para replegarlos en la conformación preferida. A modo de ejemplo, tal replegamiento puede llevarse a cabo con la adición guanidina, urea, DTT, DTE y/o una chaperonina a un producto de traducción de interés. Los procedimientos de reducción, desnaturalización y renaturalización de proteínas son muy conocidos para aquellos expertos en la materia (véanse las referencias anteriores, y Debinski, y col. (1993) J. Biol, Chem., 268: 14065-14070; Kreitman y Pastan (1993) Bioconjug. Chem.,4: 581-585; y Buchner, y col., (1992) Anal. Biochem., 205: 263-270). Debinski, y col., por ejemplo, describen la desnaturalización y reducción de proteínas de cuerpos de inclusión en guanidina-DTE. Las proteínas pueden replegarse en un tampón rédox que contiene, que incluye glutatión oxidado y L-arginina. Los reactivos de replegamiento pueden hacerse circular o moverse de otro modo en contacto con el uno o más polipéptidos u otro producto de expresión, o viceversa.

En el caso de la producción de procariotas de un polipéptido de aminoácidos no naturales, el polipéptido así

producido puede plegarse erróneamente y así carece de o tiene actividad biológica reducida. La bioactividad de la proteína puede restaurarse por "replegamiento". En una realización, un polipéptido erróneamente plegado se repliega solubilizando (en la que el polipéptido también es insoluble), desplegando y reduciendo la cadena de polipéptidos usando, a modo de ejemplo, uno o más agentes caotrópicos (que incluyen urea y/o guanidina) y un agente reductor que puede reducir enlaces disulfuro (que incluye ditiotreitol, DTT o 2-mercaptoetanol, 2-ME). A una concentración moderada de caótropo, a continuación se añade un agente de oxidación (por ejemplo, oxígeno, cistina o cistamina), que permite que se vuelvan a formar enlaces disulfuro. Un polipéptido desplegado o erróneamente plegado puede volver a plegarse usando procedimientos convencionales conocidos en la técnica, tales como aquellos descritos en las patentes de EE.UU. nº 4.511.502, 4.511.503 y 4.512.922. El polipéptido también puede co-plegarse con otras proteínas para formar heterodímeros o heteromultímeros. Después del replegamiento o coplegamiento, el polipéptido puede purificarse adicionalmente.

La purificación de polipéptidos de aminoácidos no naturales puede llevarse a cabo usando una variedad de técnicas, que incluyen aquellas descritas en el presente documento, a modo de ejemplo cromatografía de interacción hidrófoba, cromatografía de exclusión por tamaño, cromatografía de intercambio iónico, cromatografía líquida de alta resolución de fase inversa, cromatografía de afinidad, y similares, o cualquier combinación de las mismas. La purificación adicional también puede incluir una etapa de secado o precipitación de la proteína purificada.

Después de la purificación, los polipéptidos de aminoácidos no naturales pueden intercambiarse en diferentes tampones y/o concentrarse por cualquiera de una variedad de procedimientos conocidos en la técnica, que incluyen diafiltración y diálisis. La hGH que se proporciona como una proteína purificada individual puede someterse a agregación y precipitación. En ciertas realizaciones, los polipéptidos de aminoácidos no naturales purificados pueden ser al menos el 90 % puros (como se mide por cromatografía líquida de alta resolución de fase inversa, RP-HPLC o electroforesis en gel de dodecilsulfato de sodio-poliacrilamida, SDS-PAGE). En ciertas otras realizaciones, los polipéptidos de aminoácidos no naturales purificados pueden ser al menos el 95 % puros, o al menos el 98 % puros, o al menos el 99 % o de mayor pureza. Independientemente del valor numérico exacto de la pureza de los polipéptidos de aminoácidos no naturales, los polipéptidos de aminoácidos no naturales son suficientemente puros para su uso como un producto farmacéutico o para procesamiento adicional, que incluye conjugación con un polímero soluble en agua tal como PEG.

En ciertas realizaciones, las moléculas de polipéptidos de aminoácidos no naturales pueden usarse como agentes terapéuticos en ausencia de otros principios activos o proteínas (distintos de excipientes, vehículos y estabilizadores, albúmina de suero y similares), y en ciertas realizaciones, las moléculas de polipéptidos de aminoácidos no naturales pueden complejarse con otro polipéptido o un polímero.

2. Purificación de polipéptidos de aminoácidos no naturales

Procedimientos de purificación general Las técnicas desveladas en esta sección pueden aplicarse a la purificación general de los polipéptidos de aminoácidos no naturales descritos en el presente documento. Una cualquiera de una variedad de etapas de aislamiento puede realizarse sobre el extracto de lisado celular, medio de cultivo, cuerpos de inclusión, espacio periplásmico de las células huésped, citoplasma de las células huésped, u otro material que comprenda el polipéptido deseado o sobre cualquier mezcla de polipéptidos resultante de cualquier etapa de aislamiento que incluye, pero no se limita a, cromatografía de afinidad, cromatografía de intercambio iónico, cromatografía de interacción hidrófoba, cromatografía por filtración en gel, cromatografía líquida de alta resolución ("HPLC"), HPLC de fase inversa ("RP-HPLC"), adsorción en lecho expandido, o cualquier combinación y/o repetición de las mismas y en cualquier orden apropiado.

Equipo y otros materiales necesarios usados en realizar las técnicas descritas en el presente documento están comercialmente disponibles. Bombas, colectores de fracciones, monitores, grabadoras y sistemas enteros están disponibles de, por ejemplo, Applied Biosystems (Foster City, CA), Bio-Rad Laboratories, Inc. (Hercules, CA) y Amersham Biosciences, Inc. (Piscataway, NJ). Materiales cromatográficos que incluyen materiales de matriz de intercambio, medios y tampones también están disponibles de tales empresas.

La equilibración, y otras etapas en los procedimientos de cromatografía en columna descritos en el presente documento tales como lavado y elución, puede realizarse más rápidamente usando equipo especializado tal como una bomba. Bombas comercialmente disponibles incluyen bomba P-50 HILOAD®, bomba peristáltica P-1, bomba P-901 y bomba P-903 (Amersham Biosciences, Piscataway, NJ).

Ejemplos de colectores de fracciones incluyen el colector de fracciones RediFrac, los colectores de fracciones FRAC-100 y FRAC-200, y el colector de fracciones SUPERFRAC® (Amersham Biosciences, Piscataway, NJ). También están disponibles mezcladoras para formar gradientes de pH y de concentración lineal. Mezcladoras comercialmente disponibles incluyen mezcladora en gradiente GM-1 y mezcladoras en línea (Amersham Biosciences, Piscataway, NJ).

El procedimiento cromatográfico puede monitorizarse usando cualquier monitor comercialmente disponible. Tales monitores pueden usarse para conseguir información como UV, fluorescencia, pH y conductividad. Ejemplos de

detectores incluyen Monitor UV-1, UVICORD® S II, Monitor UV-M II, Monitor UV-900, Monitor UPC-900, Monitor pH/C-900 y monitor de conductividad (Amersham Biosciences, Piscataway, NJ). De hecho, están comercialmente disponibles sistemas enteros que incluyen los diversos sistemas AKTA® de Amersham Biosciences (Piscataway, NJ).

En una realización de los procedimientos y composiciones descritos en el presente documento, por ejemplo, el polipéptido puede reducirse y desnaturalizarse desnaturalizando primero el polipéptido purificado resultante en urea, seguido de dilución en tampón Tris que contiene un agente reductor (tal como DTT) a un pH adecuado. En otra realización, el polipéptido se desnaturaliza en urea en un intervalo de concentración de entre aproximadamente 2 M y aproximadamente 9 M, seguido de dilución en tampón Tris a un pH en el intervalo de aproximadamente 5,0 a aproximadamente 8,0. La mezcla de replegamiento de esta realización puede entonces incubarse. En una realización, la mezcla de replegamiento se incuba a temperatura ambiente durante cuatro a veinticuatro horas. La mezcla de polipéptido reducido y desnaturalizado puede entonces aislarse o purificarse adicionalmente.

Como se ha establecido en el presente documento, el pH de la primera mezcla de polipéptidos puede ajustarse antes de realizar cualquier etapa de aislamiento posterior. Además, la primera mezcla de polipéptidos o cualquier mezcla posterior de la misma puede concentrarse usando técnicas conocidas en la técnica. Además, el tampón de elución que comprende la primera mezcla de polipéptidos o cualquier mezcla posterior de los mismos, puede intercambiarse por un tampón adecuado para la siguiente etapa de aislamiento usando técnicas muy conocidas para aquellos expertos habituales en la materia.

Cromatografía de intercambio iónico Las técnicas desveladas en esta sección pueden aplicarse a la cromatografía iónica de los polipéptidos de aminoácidos no naturales descritos en el presente documento. En una realización, y como una etapa adicional opcional, la cromatografía de intercambio iónico puede realizarse sobre la primera mezcla de polipéptidos. Véase generalmente ION EXCHANGE CHROMATOGRAPHY: PRINCIPLES AND METHODS (Cat. nº 18-1114-21, Amersham Biosciences (Piscataway, NJ)). Columnas de intercambio iónico comercialmente disponibles incluyen columnas HITRAP®, HIPREP® y HILOAD® (Amersham Biosciences, Piscataway, NJ). Tales columnas utilizan fuertes intercambiadores aniónicos tales como Q SEPHAROSE® Fast Flow, Q SEPHAROSE® High Performance y Q SEPHAROSE® XL; fuertes intercambiadores catiónicos tales como SP SEPHAROSE® High Performance, SP SEPHAROSE® Fast Flow y SP SEPHAROSE® XL; intercambiadores aniónicos débiles tales como DEAE SEPHAROSE® Fast Flow; e intercambiadores catiónicos débiles tales como CM SEPHAROSE® Fast Flow (Amersham Biosciences, Piscataway, NJ). La cromatografía en columna de intercambio aniónico o catiónico puede realizarse sobre el polipéptido en cualquier etapa del procedimiento de purificación para aislar el polipéptido sustancialmente purificado. La etapa de cromatografía de intercambio catiónico puede realizarse usando cualquier matriz de intercambio catiónico adecuada. Las matrices de intercambio catiónico incluyen, pero no se limitan a, materiales de matriz de intercambio catiónico fibrosos, porosos, no porosos, microgranulares, perlados o reticulados. Tales materiales de matriz de intercambio catiónico incluyen celulosa, agarosa, dextrano, poliacrilato, polivinilo, poliestireno, sílice, poliéter o materiales compuestos de cualquiera de los anteriores. Tras la adsorción del polipéptido a la matriz del intercambiador catiónico, el polipéptido sustancialmente purificado puede eluirse poniendo en contacto la matriz con un tampón que tiene un pH o fuerza iónica suficientemente alto para desplazar el polipéptido de la matriz. Tampones adecuados para su uso en elución a pH alto de polipéptido sustancialmente purificado incluyen tampones citrato, fosfato, formiato, acetato, HEPES y MES que oscilan en concentración de al menos aproximadamente 5 mM a al menos aproximadamente 100 mM.

Cromatografía de fase inversa Las técnicas desveladas en esta sección pueden aplicarse a la cromatografía de fase inversa de los polipéptidos de aminoácidos no naturales descritos en el presente documento. Puede realizarse RP-HPLC para purificar proteínas siguiendo protocolos adecuados que son conocidos para aquellos expertos habituales en la materia. Véase, por ejemplo, Pearson y col., ANAL BIOCHEM. (1982) 124:217-230 (1982); Rivier y col., J.CHROM. (1983) 268:112-119; Kunitani y col., J. CHROM. (1986) 359:391-402. La RP-HPLC puede realizarse sobre el polipéptido para aislar polipéptido sustancialmente purificado. A este respecto, pueden usarse resinas derivatizadas de sílice con funcionalidades alquilo con una amplia variedad de longitudes, que incluyen, pero no se limitan a, resinas de al menos aproximadamente C_3 a al menos aproximadamente C_{30}, al menos aproximadamente C_3 a al menos aproximadamente C_{20}, o al menos aproximadamente C_3 a al menos aproximadamente C_{18}. Alternativamente, puede usarse una resina polimérica. Por ejemplo, puede usarse la resina TosoHaas Amberchrome CG1000sd, que es una resina de polímero de estireno. También pueden usarse resinas de ciano o poliméricas con una amplia variedad de longitudes de cadena de alquilo. Además, la columna de RP-HPLC puede lavarse con un disolvente tal como etanol. Un tampón de elución adecuado que contiene un agente de apareamiento de iones y un modificador orgánico tal como metanol, isopropanol, tetrahidrofurano, acetonitrilo o etanol, puede usarse para eluir el polipéptido de la columna de RP-HPLC. Los agentes de apareamiento de iones más comúnmente usados incluyen ácido acético, ácido fórmico, ácido perclórico, ácido fosfórico, ácido trifluoroacético, ácido heptafluorobutírico, trietilamina, tetrametilamonio, tetrabutilamonio, acetato de trietilamonio. La elución puede realizarse usando uno o más gradientes o condiciones isocráticas, siendo las condiciones en gradiente preferidas para reducir el tiempo de separación y para reducir la anchura de picos. Otro procedimiento implica el uso de dos gradientes con diferentes intervalos de concentración de disolventes. Ejemplos de tampones de elución adecuados para su uso en el presente documento pueden incluir disoluciones de acetato de amonio y acetonitrilo.

Técnicas de purificación por cromatografía de interacción hidrófoba Las técnicas desveladas en esta sección pueden aplicarse a la purificación por cromatografía de interacción hidrófoba de los polipéptidos de aminoácidos no naturales descritos en el presente documento. Puede realizarse cromatografía de interacción hidrófoba (HIC) en el polipéptido. Véase generalmente HYDROPHOBIC INTERACTION CHROMATOGRAPHY HANDBOOK: PRINCIPLES AND METHODS (Cat. nº 18-1020-90, Amersham Biosciences (Piscataway, NJ), que se incorpora por referencia en el presente documento. Matrices de HIC adecuadas pueden incluir matrices sustituidas con alquilo o arilo, tales como matrices sustituidas con butilo, hexilo, octilo o fenilo que incluyen matrices de agarosa, agarosa reticulada, Sepharose, celulosa, sílice, dextrano, poliestireno, poli(metacrilato), y resinas en modo mixto, que incluyen una resina de polietilenamina o una matriz de poli(metacrilato) sustituida con butilo o fenilo. Fuentes comercialmente disponibles para cromatografía en columna de interacción hidrófoba incluyen columnas HITRAP®, HIPREP® y HILOAD® (Amersham Biosciences, Piscataway, NJ). Brevemente, antes de la carga, la columna de HIC puede equilibrarse usando tampones estándar conocidos para aquellos expertos habituales en la materia, tales como una disolución de ácido acético/cloruro sódico o sulfato de amonio que contiene HEPES. El sulfato de amonio puede usarse como tampón para la carga de la columna de HIC. Después de la carga del polipéptido, la columna puede a continuación lavarse usando tampones y condiciones estándar para eliminar materiales no deseados, pero reteniendo el polipéptido sobre la columna de HIC. El polipéptido puede eluirse con aproximadamente 3 a aproximadamente 10 volúmenes de columna de un tampón estándar, tal como un tampón HEPES que contiene EDTA y menor concentración de sulfato de amonio que el tampón de equilibrio, o un tampón ácido acético/cloruro sódico, entre otros. También puede usarse un gradiente de sal lineal decreciente que usa, por ejemplo, un gradiente de fosfato de potasio, para eluir las moléculas de polipéptido. El eluyente puede entonces concentrarse, por ejemplo, por filtración tal como diafiltración o ultrafiltración. La diafiltración puede utilizarse para eliminar la sal usada para eluir el polipéptido.

Otras técnicas de purificación Las técnicas desveladas en esta sección pueden aplicarse a otras técnicas de purificación de los polipéptidos de aminoácidos no naturales descritos en el presente documento. Todavía puede realizarse otra etapa de aislamiento usando, por ejemplo, filtración en gel (GEL FILTRATION: PRINCIPLES AND METHODS (Cat. nº 18-1022-18, Amersham Biosciences, Piscataway; NJ, que se incorpora en el presente documento por referencia en su totalidad), cromatografía en hidroxiapatita (matrices adecuadas incluyen HA-Ultrogel, High Resolution (Calbiochem), CHT Ceramic Hydroxyapatite (BioRad), Bio - Gel HTP - Hydroxyapatite (BioRad)), HPLC, adsorción en lecho expandido, ultrafiltración, diafiltración, liofilización, y similares, en la primera mezcla de polipéptidos o cualquier mezcla posterior de los mismos, para eliminar cualquier exceso de sales y para sustituir el tampón con un tampón adecuado para la siguiente etapa de aislamiento o incluso la formulación del medicamento final. El rendimiento de polipéptido, que incluye polipéptido sustancialmente purificado, puede monitorizarse en cada etapa descrita en el presente documento usando diversas técnicas, que incluyen, pero no se limitan a, aquellas descritas en el presente documento. Tales técnicas pueden también usarse para evaluar el rendimiento del péptido sustancialmente purificado tras la última etapa de aislamiento. A modo de ejemplo, el rendimiento de polipéptido puede monitorizarse usando cualquiera de varias columnas de cromatografía líquida de alta presión de fase inversa, que tienen una variedad de longitudes de cadena de alquilo tales como ciano RP-HPLC, C_{18}RP-HPLC; además de HPLC de intercambio catiónico y HPLC de filtración en gel.

En ciertas realizaciones, el rendimiento de polipéptido después de cada etapa de purificación puede ser al menos aproximadamente el 30 %, al menos aproximadamente el 35 %, al menos aproximadamente el 40 %, al menos aproximadamente el 45 %, al menos aproximadamente el 50 %, al menos aproximadamente el 55 %, al menos aproximadamente el 60 %, al menos aproximadamente el 65 %, al menos aproximadamente el 70 %, al menos aproximadamente el 75 %, al menos aproximadamente el 80 %, al menos aproximadamente el 85 %, al menos aproximadamente el 90 %, al menos aproximadamente el 91 %, al menos aproximadamente el 92 %, al menos aproximadamente el 93 %, al menos aproximadamente el 94 %, al menos aproximadamente el 95 %, al menos aproximadamente el 96 %, al menos aproximadamente el 97 %, al menos aproximadamente el 98 %, al menos aproximadamente el 99 %, al menos aproximadamente el 99,9 %, o al menos aproximadamente el 99,99 %, del polipéptido en el material de partida para cada etapa de purificación.

La pureza puede determinarse usando técnicas convencionales, tales como SDS-PAGE, o midiendo polipéptido usando ensayos de transferencia Western y de ELISA. Por ejemplo, pueden generarse anticuerpos policlonales contra proteínas aisladas de la fermentación de levadura de control negativo y recuperación por intercambio catiónico. Los anticuerpos también pueden usarse para investigar la presencia de proteínas de células huésped contaminantes.

El material de RP-HPLC Vydac C4 (Vydac) consiste en partículas de gel de sílice, cuyas superficies llevan cadenas de alquilo C4. La separación de polipéptido de las impurezas proteináceas se basa en diferencias en la concentración de interacciones hidrófobas. La elución se realiza con un gradiente de acetonitrilo en ácido trifluoroacético diluido. Se realiza HPLC preparativa usando una columna de acero inoxidable (llenada con 2,8 a 3,2 litros de gel de sílice Vydac C4). El eluato de Hydroxyapatite Ultrogel se acidifica añadiendo ácido trifluoroacético y se carga sobre la columna Vydac C4. Para el lavado y elución se usa un gradiente de acetonitrilo en ácido trifluoroacético diluido. Se recogen fracciones y se neutralizan inmediatamente con tampón fosfato. Se reúnen las fracciones de polipéptido que están dentro de los límites de IPC.

El material DEAE-Sepharose (Pharmacia) en grupos de dietilaminoetilo (DEAE) que están covalentemente unidos a la superficie de perlas de Sepharose. La unión del polipéptido a los grupos DEAE está mediada por interacciones iónicas. El acetonitrilo y el ácido trifluoroacético pasan a través de la columna sin ser retenidos. Después de haberse separado por lavado estas sustancias, las impurezas traza se eliminan lavando la columna con tampón acetato a un pH bajo. Entonces, la columna se lava con tampón fosfato neutro y el polipéptido se eluye con un tampón con elevada fuerza iónica. La columna se rellena con DEAE-Sepharose fast flow. Se ajusta el volumen de columna para asegurar una carga de polipéptido en el intervalo de 3-10 mg de polipéptido/ml de gel. La columna se lava con agua y tampón de equilibrio (fosfato de sodio/potasio). Se cargan las fracciones reunidas del eluato de HPLC y la columna se lava con tampón de equilibrio. A continuación, la columna se lava con tampón de lavado (tampón acetato sódico), seguido de lavado con tampón de equilibrio. Posteriormente, el polipéptido se eluye de la columna con tampón de elución (cloruro sódico, fosfato de sodio/potasio) y se recoge en una única fracción según el perfil de elución maestro. El eluato de la columna de DEAE-Sepharose se ajusta a la conductividad especificada. El principio activo resultante se esteriliza por filtración en botellas de teflón y se guarda a -70 ºC.

Procedimientos adicionales incluyen etapas para eliminar endotoxinas. Las endotoxinas son lipopolisacáridos (LPS) que se localizan sobre la membrana externa de células huésped Gram-negativas tales como, por ejemplo, *Escherichia coli*. Procedimientos para reducir los niveles de endotoxina incluyen técnicas de purificación usando soportes de sílice, polvo de vidrio o hidroxiapatita, cromatografía de fase inversa, de afinidad, de exclusión por tamaño, de intercambio aniónico, cromatografía de interacción hidrófoba, una combinación de estos procedimientos, y similares. Pueden requerirse modificaciones o procedimientos adicionales para eliminar contaminantes tales como proteínas co-migrantes del polipéptido de interés. Los procedimientos para medir los niveles de endotoxina son conocidos para un experto habitual en la materia e incluyen ensayos de lisado de amebocitos de Limulus (LAL).

Métodos y procedimientos adicionales incluyen SDS-PAGE acoplada a procedimientos de tinción de proteína, inmunotransferencia, desorción/ionización láser asistida por matriz-espectrometría de masas (MALDI-EM), cromatografía de líquidos/espectrometría de masas, isoelectroenfoque, intercambio aniónico analítico, cromatoenfoque y dicroísmo circular.

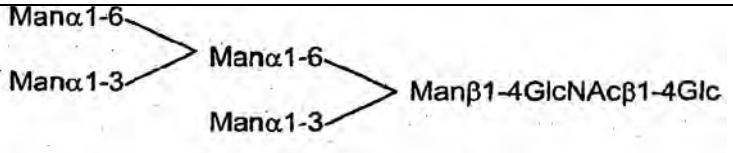
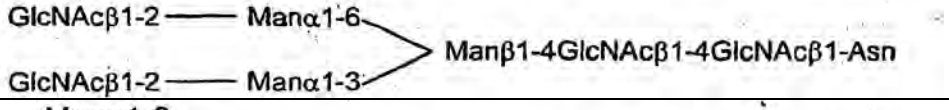
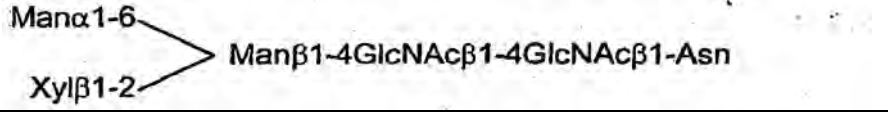
En ciertas realizaciones, los aminoácidos no naturales descritos en el presente documento pueden incorporarse biosintéticamente en polipéptidos, preparando así polipéptidos de aminoácidos no naturales. En otras realizaciones, tales aminoácidos se incorporan en un sitio específico dentro del polipéptido. En otras realizaciones, tales aminoácidos se incorporan en el polipéptido usando un sistema de traducción. En otras realizaciones, tales sistemas de traducción comprenden: (i) un polinucleótido que codifica el polipéptido, en el que el polinucleótido comprende un selector de codones correspondiente al sitio de incorporación previamente designado de los aminoácidos anteriores, y (ii) un ARNt que comprende el aminoácido, en el que el ARNt es específico para el selector de codones. En otras realizaciones de tales sistemas de traducción, el polinucleótido es ARNm producido en el sistema de traducción. En otras realizaciones de tales sistemas de traducción, el sistema de traducción comprende un plásmido o un fago que comprende el polinucleótido. En otras realizaciones de tales sistemas de traducción, el sistema de traducción comprende ADN genómico que comprende el polinucleótido. En otras realizaciones de tales sistemas de traducción, el polinucleótido está establemente integrado en el ADN genómico. En otras realizaciones de tales sistemas de traducción, el sistema de traducción comprende ARNt específico para un selector de codones seleccionado del grupo que consiste en un codón ámbar, codón ocre, codón ópalo, un codón único, un codón raro, un codón no natural, un codón de cinco bases y un codón de cuatro bases. En otras realizaciones de tales sistemas de traducción, el ARNt es un ARNt supresor. En otras realizaciones de tales sistemas de traducción, el sistema de traducción comprende un ARNt que está aminoacilado por los aminoácidos anteriores. En otras realizaciones de tales sistemas de traducción, el sistema de traducción comprende una aminoacil sintetasa específica para el ARNt. En otras realizaciones de tales sistemas de traducción, el sistema de traducción comprende un ARNt ortogonal y una aminoacil ARNt sintetasa ortogonal. En otras realizaciones de tales sistemas de traducción, el polipéptido se sintetiza por un ribosoma, y en otras realizaciones el sistema de traducción es un sistema de traducción *in vivo* que comprende una célula seleccionada del grupo que consiste en una célula bacteriana, célula arqueobacteriana y célula eucariota. En otras realizaciones, la célula es una célula de *Escherichia coli*, célula de levadura, una célula de una especie de *Pseudomonas*, célula de mamífero, célula de planta, o una célula de insecto. En otras realizaciones de tales sistemas de traducción, el sistema de traducción es un sistema de traducción *in vitro* que comprende extracto celular de una célula bacteriana, célula arqueobacteriana o célula eucariota. En otras realizaciones, el extracto celular es de una célula de *Escherichia coli*, una célula de una especie de *Pseudomonas*, célula de levadura, célula de mamífero, célula de planta o una célula de insecto. En otras realizaciones, al menos una porción del polipéptido se sintetiza por síntesis de péptidos en fase sólida o en fase de disolución, o una combinación de las mismas, mientras que en otras realizaciones comprende además ligar el polipéptido a otro polipéptido. En otras realizaciones, los aminoácidos no naturales descritos en el presente documento pueden incorporarse biosintéticamente en polipéptidos, en los que el polipéptido es una proteína homóloga a una proteína terapéutica seleccionada del grupo que consiste en: alfa-1 antitripsina, angiostatina, factor antihemolítico, anticuerpo, apolipoproteína, apoproteína, factor natriurético auricular, polipéptido natriurético auricular, péptido auricular, quimiocina C-X-C, T39765, NAP-2, ENA-78, gro-a, gro-b, gro-c, IP-10, GCP-2, NAP-4, SDF-1, PF4, MIG, calcitonina, ligando c-kit, citocina, quimiocina CC, proteína 1 quimioatrayente de monocitos, proteína 2 quimioatrayente de monocitos, proteína 3 quimioatrayente de monocitos, proteína 1 alfa inflamatoria de monocitos, proteína 1 beta inflamatoria de monocitos, RANTES, 1309, R83915, R91733, HCC1, T58847, D31065, T64262, CD40, ligando

CD40, ligando c-kit, colágeno, factor estimulante de colonias (CSF), factor 5a del complemento, inhibidor del complemento, receptor 1 del complemento, citocina, péptido 78 activante de neutrófilos epiteliales, MIP-16, MCP-1, factor de crecimiento epidérmico (EGF), péptido activante de neutrófilos epiteliales, eritropoyetina (EPO), toxina exfoliante, factor IX, factor VII, factor VIII, factor X, factor de crecimiento de fibroblastos (FGF), fibrinógeno, fibronectina, proteína de cuatro haces helicoidales, G-CSF, glp-1, GM-CSF, glucocerebrosidasa, gonadotropina, factor de crecimiento, receptor de factores de crecimiento, grf, proteína hedgehog, hemoglobina, factor de crecimiento de hepatocitos (hGH), hirudina, hormona de crecimiento humana (hGH), albúmina de suero humano, ICAM-1, receptor de ICAM-1, LFA-1, receptor de LFA-1, insulina, factor de crecimiento similar a la insulina (IGF), IGF-I, IGF-II, interferón (IFN), IFN-alfa, IFN-beta, IFN-gamma, cualquier molécula similar a interferón o miembro de la familia de IFN, interleucina (IL), IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, factor de crecimiento de queratinocitos (KGF), lactoferrina, factor inhibidor de la leucemia, luciferasa, neurturina, factor inhibidor de neutrófilos (NIF), oncostatina M, proteína osteogénica, producto oncogénico, paracitonina, hormona paratiroidea, PD-ECSF, PDGF, hormona peptídica, pleiotropina, proteína A, proteína G, pth, endotoxina A pirogénica, exotoxina B pirogénica, exotoxina C pirogénica, pyy, relaxina, renina, SCF, proteína biosintética pequeña, receptor I del complemento soluble, I-CAM 1 soluble, receptor de interleucina soluble, receptor de TNF soluble, somatomedina, somatostatina, somatotropina, estreptocinasa, superantígenos, enterotoxina estafilocócica, SEA, SEB, SEC1, SEC2, SEC3, SED, SEE, receptor de hormonas esteroideas, superóxido dismutasa, toxina del síndrome del choque tóxico, timosina alfa 1, activador tisular del plasminógeno, factor de crecimiento tumoral (TGF), factor de necrosis tumoral, factor de necrosis tumoral alfa, factor de necrosis tumoral beta, receptor del factor de necrosis tumoral (TNFR), proteína VLA-4, proteína VCAM-1, factor de crecimiento endotelial vascular (VEGF); urocinasa, 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 corticosterona.

B. Modificaciones postraduccionales in vivo

Produciendo proteínas o polipéptidos de interés con al menos un aminoácido no natural en células eucariotas, tales polipéptidos pueden incluir modificaciones postraduccionales eucariotas. En ciertas realizaciones, una proteína incluye al menos un aminoácido no natural y al menos una modificación postraduccional que se hace *in vivo* por una célula eucariota, en la que la modificación postraduccional no se hace por una célula procariota. A modo de ejemplo, la modificación postraduccional incluye acetilación, acilación, modificación por lípidos, palmitoilación, adición de palmitato, fosforilación, modificación del enlace con glucolípido, glucosilación y similares. En un aspecto, la modificación postraduccional incluye la unión de un oligosacárido (que incluye, pero no se limita a, $(GlcNAc\text{-}Man)_2$-Man-GlcNAc-GlcNAc)) a una asparagina por un enlace de GlcNAc-asparagina. Véase la Tabla 1 que enumera algunos ejemplos de oligosacáridos ligados en N de proteínas eucariotas (también puede estar presentes residuos adicionales, que no se muestran). En otro aspecto, la modificación postraduccional incluye la unión de un oligosacárido (que incluye Gal-GalNAc, Gal-GlcNAc, etc.) a una serina o treonina por un enlace GalNAc-serina o GalNAc-treonina, o un enlace GlcNAc-serina o GlcNAc-treonina.

Tabla 1: Ejemplos de oligosacáridos mediante enlace de GlcNAc

Tipo	Estructura base
ALTO CONTENIDO DE MANOSA	 [01]
HIBRIDO	
COMPLEJO	
XILOSA	

En otro aspecto más, la modificación postraduccional incluye procesamiento proteolítico de precursores (que incluyen precursor de calcitonina, precursor de péptido relacionado con el gen calcitonina, hormona preproparatiroidea, preproinsulina, proinsulina, prepro-opiomelanocortina, pro-opiomelanocortina y similares), ensamblaje en una proteína de multisubunidad o ensamblaje macromolecular, traducción a otro sitio en la célula (que incluye a orgánulos, tales como el retículo endoplásmico, el aparato de Golgi, el núcleo, lisosomas, peroxisomas, mitocondrias, cloroplastos, vacuolas, etc., o mediante la ruta secretora). En ciertas realizaciones, la

proteína comprende una secuencia de secreción o de localización, una marca de epítope, una marca FLAG, una marca de polihistidina, una fusión de GST, o similares.

Una ventaja de un aminoácido no natural es que presenta restos químicos adicionales que pueden usarse para añadir moléculas adicionales. Estas modificaciones pueden hacerse *in vivo* en una célula eucariota o no eucariota, o *in vitro*. Así, en ciertas realizaciones, la modificación postraduccional es mediante el aminoácido no natural. Por ejemplo, la modificación postraduccional puede ser mediante una reacción nucleófila-electrófila. La mayoría de las reacciones actualmente usadas para la modificación selectiva de proteínas implica la formación de enlaces covalentes entre componentes de reacción nucleófilos y electrófilos, que incluyen, pero no se limitan a, la reacción de cc-halocetonas con histidina o cadenas laterales de cisteína. La selectividad en estos casos se determina por el número y accesibilidad de los residuos nucleófilos en la proteína. En los polipéptidos descritos en el presente documento o producidos usando los procedimientos descritos en el presente documento, pueden usarse otras reacciones más selectivas, que incluyen, pero no se limitan a, la reacción de un ceto-aminoácido no natural con hidrazidas o compuestos de aminooxi, *in vitro* e *in vivo*. Véanse, por ejemplo, Cornish y col., (1996) J. Am. Chem. Soc., 118:8150-8151; Mahal y col., (1997) Science, 276:1125-1128; Wang y col., (2001) Science 292:498-500; Chin y col., (2002) J. Am. Chem. Soc. 124:9026-9027; Chin y col., (2002) Proc. Natl. Acad. Sci., 99:11020-11024; Wang y col., (2003) Proc. Natl. Acad. Sci., 100:56-61; Zhang y col., (2003) Biochemistry, 42:6735-6746; y, Chin y col., (2003) Science, 301:964-7. Esto permite el marcado selectivo de prácticamente cualquier proteína con un huésped de reactivos que incluyen fluoróforos, agentes de reticulación, derivados de sacáridos y moléculas citotóxicas. Véase, por tanto, la patente de EE.UU. nº 6.927.042 titulada “Glycoprotein synthesis” presentada el 16 de enero de 2003. También pueden hacerse modificaciones postraduccionales, que incluyen mediante un aminoácido azido, mediante la ligación de Staudinger (que incluye, pero no se limita a, con reactivos de triarilfosfina). Véase, por ejemplo, Kiick y col., (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation, PNAS 99:19-24.

IX. Sistemas alternativos para producir polipéptidos de aminoácidos no naturales

Se han empleado varias estrategias para introducir aminoácidos no naturales en proteínas en células huésped no recombinantes, células huésped mutagenizadas, o en sistemas sin células. Los sistemas alternativos desvelados en esta sección pueden aplicarse a la producción de aminoácido no natural. A modo de ejemplo, la derivatización de aminoácidos con cadenas laterales reactivas tales como Lys, Cys y Tyr produjeron la conversión de lisina en N^2-acetil-lisina. La síntesis química también proporciona un procedimiento directo para incorporar aminoácidos no naturales. Con el reciente desarrollo de la ligación enzimática y ligación química nativa de fragmentos de péptidos es posible preparar proteínas más grandes. Véanse, por ejemplo, P. E. Dawson y S. B. H. Kent, Annu. Rev. Biochem., 69:923 (2000). La ligación química de péptidos y la ligación química nativa se describen en la patente de EE.UU. nº 6.184.344, la publicación de patente de EE.UU. nº 2004/0138412, la publicación de patente de EE.UU. nº 2003/0208046, documentos WO 02/098902 y WO 03/042235. Se ha usado un procedimiento biosintético *in vitro* general en el que un ARNt supresor acilado químicamente con el aminoácido no natural deseado se añade a un extracto *in vitro* que puede soportar la biosíntesis de proteínas para incorporar específicamente en el sitio más de 100 aminoácidos no naturales en una variedad de proteínas de prácticamente cualquier tamaño. Véanse, por ejemplo, V. W. Cornish, D. Mendel y P. G. Schultz, Angew. Chem. Int. Ed. Engl., 1995,34:621 (1995); C.J. Noren, S.J. Anthony-Cahill, M.C. Griffth, P.G. Schultz, A general method for site-specific incorporation of unnatural amino acids into proteins, Science 244 182-188 (1989); y J.D. Bain, C.G. Glabe, T.A. Dix, A.R. Chamberlin, E.S. Diala, Biosynthetic site-specific incorporation of a non-natural amino acid into a polypeptide, J. Am. Chem. Soc. 111 8013-8014 (1989). Se ha introducido una amplia variedad de grupos funcionales en proteínas para estudios de estabilidad de proteínas, plegamiento de proteínas, mecanismo enzimático y transducción de señales.

Se desarrolló un procedimiento *in vivo*, llamado incorporación por presión selectiva, para explotar la promiscuidad de las sintetasas naturales. Véase, por ejemplo, N. Budisa, C. Minks, S. Alefelder, W. Wenger, F. M. Dong, L. Moroder y R. Huber, FASEB J., 13:41 (1999). Una cepa auxotrófica, en la que la ruta metabólica relevante que provee la célula con un aminoácido natural particular está desconectada, se cultiva en medio mínimo que contiene concentraciones limitadas del aminoácido natural, mientras que se reprime la transcripción del gen diana. En la aparición de una fase de crecimiento estacionario, el aminoácido natural se empobrece y se sustituye con el análogo de aminoácido no natural. La inducción de la expresión de la proteína recombinante produce la acumulación de una proteína que contiene el análogo no natural. Por ejemplo, usando esta estrategia, o, m y p-fluorofenilalaninas se han incorporado en proteínas, y presentan dos hombros característicos en el espectro UV que pueden ser fácilmente identificados, véase, por ejemplo, C. Minks, R. Huber, L. Moroder y N. Budisa, Anal. Biochem., 284:29 (2000); se ha usado trifluorometionina para sustituir metionina en lisozima del bacteriófago T4 para estudiar su interacción con ligandos de quitooligosacárido por RMN ^{19}F, véase, por ejemplo, H. Duewel, E. Daub, V. Robinson y J. F. Honek, Biochemistry, 36:3404 (1997); y se ha incorporado trifluoroleucina en lugar de leucina, produciendo elevada estabilidad térmica y química de una proteína de cremallera de leucina. Véanse, por ejemplo, Y. Tang, G. Ghirlanda, W. A. Petka, T. Nakajima, W. F. DeGrado y D. A. Tirrell, Angew. Chem. Int. Ed, Engl., 40:1494 (2001). Además, se incorporan selenometionina y telurometionina en diversas proteínas recombinantes para facilitar la disolución de fases en cristalografía de rayos X. Véanse, por ejemplo, W. A. Hendrickson, J. R. Horton y D. M. Lemaster, EMBO J., 9:1665 (1990); J. O. Boles, K. Lewinski, M. Kunkle, J. D. Odom, B. Dunlap, L. Lebioda y M. Hatada, Nat. Struct. Biol., 1:283 (1994); N. Budisa, B. Steipe, P. Demange, C. Eckerskorn, J. Kellermann y R. Huber, Eur. J. Biochem.,

230:788 (1995); y N. Budisa, W. Karnbrock, S. Steinbacher, A. Humm, L. Prade, T. Neuefeind, L. Moroder y R. Huber, J. Mol. Biol., 270:616 (1997). También se han incorporado eficazmente análogos de metionina con funcionalidades alqueno o alquino, permitiendo la modificación adicional de proteínas por medios químicos. Véanse, por ejemplo, J. C. M. van Hest y D. A. Tirrell, FEBS Lett., 428:68 (1998); J. C. M. van Hest, K. L. Kiick y D. A. Tirrell, J. Am. Chem. Soc., 122:1282 (2000); y, K. L. Kiick y D. A. Tirrell, Tetrahedron, 56:9487 (2000); patente de EE.UU. nº 6.586.207; publicación de patente de EE.UU. 2002/0042097.

El éxito de este procedimiento depende del reconocimiento de los análogos de aminoácidos no naturales por aminoacil-ARNt sintetasas, que, en general, requieren alta selectividad para asegurar la fidelidad de la traducción de proteínas. Una forma de ampliar el alcance de este procedimiento es relajar la especificidad del sustrato de aminoacil-ARNt sintetasas, que se ha conseguido en un número limitado de casos. A modo de ejemplo solo, la sustitución de Ala^{294} por Gly en fenilalanil-ARNt sintetasa (PheRS) de *Escherichia coli* aumenta el tamaño del sitio de unión a sustrato, y produce la acilación de ARNtPhe por p-Cl-fenilalanina (p-Cl-Phe).Véanse M. Ibba, P. Kast y H. Hennecke, Biochemistry, 33:7107 (1994). Una cepa de *Escherichia coli* que aloja esta PheRS mutante permite la incorporación de p-Cl-fenilalanina o p-Br-fenilalanina en lugar de fenilalanina. Véanse, por ejemplo, M. Ibba y H. Hennecke, FEBS Lett., 364:272 (1995); y N. Sharma, R. Furter, P. Kast y D. A. Tirrell, FEBS Lett., 467:37 (2000). Similarmente, se mostró que una mutación puntual Phe130Ser próxima al sitio de unión del aminoácido de tirosil-ARNt sintetasa de *Escherichia coli* permitía que se incorporara azatirosina más eficazmente que la tirosina. Véase F. Hamano-Takaku, T. Iwama, S. Saito-Yano, K. Takaku, Y. Monden, M. Kitabatake, D. Soll y S. Nishimura, J. Biol. Chem., 275:40324 (2000).

Otra estrategia para incorporar aminoácidos no naturales en proteínas *in vivo* es modificar sintetasas que tienen mecanismos de corrección. Estas sintetasas no pueden discriminar y, por tanto, activan aminoácidos que son estructuralmente similares a los aminoácidos naturales relacionados. Este error se corrige en un sitio separado, que desacila el aminoácido erróneamente cargado del ARNt para mantener la fidelidad de la traducción de proteínas. Si la actividad de corrección de la sintetasa se inhabilita, análogos estructurales que son erróneamente activados pueden escapar de la función de edición e incorporarse. Este enfoque se ha demostrado recientemente con la valil-ARNt sintetasa (ValRS). Véanse V. Doring, H. D. Mootz, L. A. Nangle, T. L. Hendrickson, V. de Crecy-Lagard, P. Schimmel y P. Marliere, Science. 292:501 (2001). ValRS puede acilar erróneamente ARNtVal con Cys, Thr o aminobutirato (Abu); estos aminoácidos no relacionados se hidrolizan posteriormente por el dominio de edición. Después de la mutagénesis al azar del cromosoma de *Escherichia coli*, se seleccionó una cepa de *Escherichia coli* mutante que tenía una mutación en el sitio de edición de ValRS. Esta ValRS defectuosa en la edición carga incorrectamente ARNtVal con Cys. Debido a que Abu se parece estéricamente a Cys (el grupo -SH de Cys se sustituye por -CH3 en Abu), la ValRS mutante también incorpora Abu en proteínas cuando esta cepa de *Escherichia coli* mutante se cultiva en presencia de Abu. El análisis de espectrometría de masas muestra que aproximadamente el 24 % de las valinas están sustituidas con Abu en cada posición de valina en la proteína nativa.

Procedimientos de síntesis en fase sólida y semisintéticos también han permitido la síntesis de varias proteínas que contienen aminoácidos novedosos. Por ejemplo, véanse las siguientes publicaciones y referencias citadas dentro de, que son las siguientes: Crick, F.H.C., Barrett, L. Brenner, S. Watts-Tobin, R. General nature of the genetic code for proteins. Nature, 192:1227-1232 (1961); Kaiser, E.T. Synthetic approaches to biologically active peptides and proteins including enzymes, Acc Chem Res, 22:47-54 (1989); Nakatsuka, T., Sasaki, T., Kaiser, E.T. Peptide segment coupling catalyzed by the semisynthetic enzyme thiosubtilisin, J Am Chem Soc, 109:3808-3810 (1987); Schnolzer, M., Kent, S B H. Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease, Science, 256(5054):221-225 (1992); Chaiken, I.M. Semisynthetic peptides and proteins, CRC Crit Rev Biochem, 11(3):255-301 (1981); Offord, R.E. Protein engineering by chemical means? Protein Eng., 1(3):151-157 (1987); y Jackson, D.Y., Burnier, J., Quan, C., Stanley, M., Tom, J., Wells, J.A. A Designed Peptide Ligase for Total Synthesis of Ribonuclease A with Unnatural Catalytic Residues, Science, 266(5183):243 (1994).

Se ha usado modificación química para introducir una variedad de cadenas laterales no naturales, que incluyen cofactores, marcas de espín y oligonucleótidos, en proteínas *in vitro*. Véase, por ejemplo, Corey, D.R., Schultz, P.G. Generation of a hybrid sequence-specific single-stranded deoxyribonuclease, Science, 238(4832):1401-1403 (1987); Kaiser, E.T., Lawrence D.S., Rokita, S.E. The chemical modification of enzymatic specificity, Annu. Rev Biochem, 54:565-595 (1985); Kaiser, E.T., Lawrence, D.S. Chemical mutation of enzyme active sites, Science, 226(4674):505-511 (1984); Neet, K.E., Nanci A, Koshland, D.E. Properties of thiol-subtilisin, J Biol. Chem, 243(24):6392-6401 (1968); Polgar, L. (ed.)., M.L. Bender, A new enzyme containing a synthetically formed active site. Thiol-subtilisin. J. Am Chem Soc, 88:3153-3154 (1966); y Pollack, S.J., Nakayama, G. Schultz, P.G. Introduction of nucleophiles and spectroscopic probes into antibody combining sites, Science, 224(4881):1038-1040 (1988).

Alternativamente, se han usado procedimientos biosintéticos que emplean aminoacil-ARNt químicamente modificados para incorporar varias sondas biofísicas en proteínas sintetizadas *in vitro*. Véanse las siguientes publicaciones y referencias citadas dentro de: Brunner, J. New Photolabeling and crosslinking methods, Annu. Rev Biochem, 483-514 (1993); y Krieg, U.C., Walter, P., Hohnson, A.E. Photocrosslinking of the signal sequence of nascent preprolactin of the 54-kilodalton polypeptide of the signal recognition particle, Proc. Natl. Acad. Sci. 8604-8608 (1986).

Previamente, se ha mostrado que aminoácidos no naturales pueden incorporarse específicamente en el sitio en proteínas *in vitro* mediante la adición de ARNt supresores químicamente aminoacilados para reacciones de síntesis de proteínas programadas con un gen que contiene una mutación terminadora ámbar deseada. Usando estos enfoques, pueden sustituirse varios de los veinte aminoácidos comunes con homólogos estructurales parecidos, por ejemplo, fluorofenilalanina por fenilalanina, usando cepas auxotróficas para un aminoácido particular. Véanse, por ejemplo, Noren, C.J., Anthony-Cahill, Griffith, M.C., Schultz, P.G. A general method for site-specific incorporation of unnatural amino acids into proteins, Science, 244: 182-188 (1989); M.W. Nowak y col., Science 268:439-42 (1995); Bain, J.D., Glabe, C.G., Dix, T.A., Chamberlin, A.R., Diala, E.S. Biosynthetic site-specific Incorporation of a non-natural amino acid into a polypeptide, J. Am Chem Soc, 111:8013-8014 (1989); N. Budisa y col., FASEB J. 13:41-51 (1999); Ellman, J.A., Mendel, D., Anthony-Cahill, S., Noren, C.J., Schultz, P.G. Biosynthetic method for introducing unnatural amino acids site-specifically into proteins, Methods in Enz., vol. 202, 301-06 (199); y Mendel, D., Cornish, V.W. & Schultz, P.G. Site-Directed Mutagenesis with an Expanded Genetic Code, Annu Rev Biophys. Biomol Struct. 24, 435-62 (1995).

Por ejemplo, se preparó un ARNt supresor que reconoció el codón de terminación UAG y se aminoaciló químicamente con un aminoácido no natural. Se usó mutagénesis dirigida al sitio convencional para introducir el codón de terminación TAG, en el sitio de interés en el gen de proteína. Véase, por ejemplo, Sayers, J.R., Schmidt, W. Eckstein, F. 5', 3' Exonuclease in phosphorothioate-based oligonucleotide-directed mutagenesis, Nucleic Acids Res, 791-802 (1988). Cuando el ARNt supresor acilado y el gen mutante se combinaron en un sistema de transcripción/traducción *in vitro*, el aminoácido no natural se incorporó en respuesta al codón UAG que dio una proteína que contiene ese aminoácido en la posición especificada. Experimentos usando [^{3}H]-Phe y experimentos con α-hidroxiácidos demostraron que solo el aminoácido deseado se incorpora en la posición especificada por el codón UAG y que este aminoácido no se incorpora en cualquier otro sitio en la proteína. Véanse, por ejemplo, Noren, y col., arriba*;* Kobayashi y col., (2003) Nature Structural Biology 10(6):425-432; y, Ellman, J.A., Mendel, D., Schultz, P.G. Site-specific incorporation of novel backbone structures into proteins, Science, 197-200 (1992):

También se han usado técnicas de microinyección para incorporar aminoácidos no naturales en proteínas. Véase, por ejemplo, M. W. Nowak, P. C. Kearney, J. R. Sampson, M. E. Saks, C. G. Labarca, S. K. Silverman, W. G. Zhong, J. Thorson, J. N. Abelson, N. Davidson, P. G. Schultz, D. A. Dougherty y H. A. Lester, Science, 268:439 (1995); y D. A. Dougherty, Curr. Opin. Chem. Biol., 4:645 (2000). Se coinyectó un ovocito de Xenopus con dos especies de ARN preparadas *in vitro*: un ARNm que codifica la proteína diana con un codón de terminación UAG en la posición del aminoácido de interés y un ARNt supresor ámbar aminoacilado con el aminoácido no natural deseado. La maquinaria traduccional del ovocito inserta entonces el aminoácido no natural en la posición especificada por UAG. Este procedimiento ha permitido estudios de estructura-función *in vivo* de proteínas de membrana integral, que generalmente no son aceptados para sistemas de expresión *in vitro*. Ejemplos incluyen la incorporación de un aminoácido fluorescente en el receptor de taquinina neuroquinina-2 para medir distancias por transferencia de energía por resonancia de fluorescencia, véanse, por ejemplo, G. Turcatti, K. Nemeth, M. D. Edgerton, U. Meseth, F. Talabot, M. Peitsch, J. Knowles, H. Vogel y A. Chollet, J. Biol. Chem., 271:19991 (1996); la incorporación de aminoácidos biotinilados para identificar residuos expuestos en la superficie en canales de iones, véase, por ejemplo, J. P. Gallivan, H. A. Lester y D. A. Dougherty, Chem. Biol., 4:739 (1997); el uso de análogos de tirosina enjaulada para monitorizar cambios conformacionales en un canal de iones en tiempo real, véase, por ejemplo, J. C. Miller, S. K. Silverman, P. M. England, D. A. Dougherty y H. A. Lester, Neuron, 20:619 (1998); y, el uso de alfa-hidroxiaminoácidos para cambiar esqueletos de canales de iones para averiguar sus mecanismos de apertura. Véanse, por ejemplo, P. M. England, Y. Zhang, D. A. Dougherty y H. A. Lester, Cell, 96:89 (1999); y, T. Lu, A. Y. Ting, J. Mainland, L. Y. Jan, P. G. Schultz y J. Yang, Nat. Neurosci., 4:239 (2001).

La capacidad para incorporar aminoácidos no naturales directamente en proteínas *in vivo* ofrece una amplia variedad de ventajas que incluyen altos rendimientos de proteínas mutantes, facilidad técnica, el potencial de estudiar las proteínas mutantes en células o posiblemente en organismos vivos y el uso de estas proteínas mutantes en tratamientos terapéuticos. La capacidad para incluir aminoácidos no naturales con diversos tamaños, acideces, nucleofilias, hidrofobias y otras propiedades en proteínas puede expandir enormemente nuestra capacidad para manipular racionalmente y sistemáticamente las estructuras de proteínas, tanto para averiguar la función de proteínas como para crear nuevas proteínas u organismos con propiedades novedosas.

En un intento por incorporar específicamente en el sitio para-F-Phe, se usó un par de ARNtPheCUA supresor ámbar de levadura/fenilalanil-ARNt sintetasa en una cepa de *Escherichia coli* auxotrófica para Phe resistente a p-F-Phe. Véase, por ejemplo, R. Furter, Protein Sci., 7:419 (1998).

También puede ser posible obtener la expresión de un polinucleótido deseado fusionando un sistema de traducción sin células (*in vitro*). Los sistemas de traducción pueden ser celulares o sin células, y pueden ser procariotas o eucariotas. Sistemas de traducción celulares incluyen preparaciones de células completas tales como células permeabilizadas o cultivos celulares en los que una secuencia de ácidos nucleicos deseada puede transcribirse en ARNm y traducirse el ARNm. Los sistemas de traducción sin células están comercialmente disponibles y son muy conocidos muchos tipos y sistemas diferentes. Ejemplos de sistemas sin células incluyen lisados procariotas tales como lisados de *Escherichia coli* y lisados eucariotas tales como extractos de germen de trigo, lisados de células de insecto, lisados de reticulocitos de conejo, lisados de ovocitos de conejo y lisados de células humanas. Pueden

preferirse extractos o lisados eucariotas cuando la proteína resultante está glucosilada, fosforilada o modificada de otro modo debido a que muchas de tales modificaciones son solo posibles en sistemas eucariotas. Algunos de estos extractos y lisados están disponibles comercialmente (Promega; Madison, Wis.; Stratagene; La Jolla, Calif.; Amersham; Arlington Heights, III.; GIBCO/BRL; Grand Island, N.Y.). También están disponibles extractos membranosos, tales como los extractos pancreáticos caninos que contienen membranas de microsomas, que son útiles para traducir proteínas secretoras. En estos sistemas, que pueden incluir tanto ARNm como molde (traducción *in vitro*) como ADN como molde (transcripción y traducción combinada *in vitro*), la síntesis *in vitro* está dirigida por los ribosomas. Se ha aplicado un esfuerzo considerable al desarrollo de sistemas de expresión de proteínas sin células. Véanse, por ejemplo, Kim, D.-M. y J.R. Biotechnology and Bioengineering, 74 :309-316 (2001); Kim, D.-M. y J.R. Swartz, Biotechnology Letters, 22, 1537-1542, (2000); Kim, D.-M. y J.R. Swartz, Biotechnology Progress, 16, 385-390, (2000); Kim, D.-M. y J.R. Swartz, Biotechnology and Bioengineering, 66, 180-188, (1999); y Patnaik, R. y J.R. Swartz, Biotechniques 24, 862-868, (1998); patente de EE.UU. nº 6.337.191; publicación de patente de EE.UU. nº 2002/0081660; documentos WO 00/55353; WO 90/05785. Otro enfoque que puede aplicarse a la expresión de polipéptidos que comprenden un aminoácido no natural incluye la técnica de fusión de ARNm-péptido. Véanse, por ejemplo, R. Roberts y J. Szostak, Proc. Natl Acad. Sci. (USA) 94 12297-12302 (1997); A. Frankel, y col., Chemistry & Biology 10, 1043-1050 (2003). En este enfoque, un molde de ARNm ligado a puromicina se traduce en péptido sobre el ribosoma. Si se han modificado una o más moléculas de ARNt, también pueden incorporarse aminoácidos no naturales en el péptido. Después de haberse leído el último codón de ARNm, la puromicina captura el extremo C del péptido. Si se encuentra que el conjugado de ARNm-péptido resultante tiene propiedades interesantes en un ensayo *in vitro*, su identidad puede revelarse fácilmente a partir de la secuencia de ARNm. De esta forma, pueden cribarse bibliotecas de polipéptidos que comprenden uno o más aminoácidos no naturales para identificar polipéptidos que tienen propiedades deseadas. Más recientemente, se ha informado de traducciones de ribosomas *in vitro* con componentes purificados que permiten la síntesis de péptidos sustituidos con aminoácidos no naturales. Véase, por ejemplo, A. Forster y col., Proc. Natl Acad. Sci. (USA) 100 6353 (2003).

También pueden usarse sistemas de traducción reconstituidos. También se han usado satisfactoriamente mezclas de factores de traducción purificados para traducir ARNm en proteína, además de combinaciones de lisados o lisados complementados con factores de traducción purificados tales como factor 1 de iniciación (SI-1), SI-2, SI-3, factor T de elongación (EF-Tu) o factores de terminación. Los sistemas sin células también pueden acoplarse a sistemas de transcripción/traducción en los que el ADN se introduce en el sistema, se transcribe en ARNm y el ARNm se traduce como se ha descrito en Current Protocols in Molecular Biology (F. M. Ausubel y col. editors, Wiley Interscience, 1993), que se incorpora por este documento específicamente por referencia. El ARN transcrito en el sistema de transcripción eucariota puede estar en forma de ARN heteronuclear (hnRNA) o tapa del extremo 5' (7-metilguanosina) y ARNm maduro con colas de poli A en el extremo 3', que pueden ser una ventaja en ciertos sistemas de traducción. Por ejemplo, los ARNm tapados se traducen con alta eficiencia en el sistema de lisado de reticulocitos.

UN ARNt puede ser aminoacilado con un aminoácido deseado por cualquier procedimiento o técnica, que incluye aminoacilación química o enzimática.

La aminoacilación puede llevarse a cabo por aminoacil ARNt sintetasas o por otras moléculas enzimáticas, que incluyen ribozimas. El término “ribozima” es intercambiable con “ARN catalítico”. Cech y colaboradores (Cech, 1987, Science, 236:1532-1539; McCorkle y col., 1987, Concepts Biochem. 64:221-226) demostraron la presencia de ARN que se producen naturalmente que pueden actuar de catalizadores (ribozimas). Sin embargo, aunque solo se ha mostrado que estos catalizadores de ARN naturales actúan sobre sustratos de ácido ribonucleico para escisión y corte y empalme, el reciente desarrollo de evolución artificial de ribozimas ha ampliado el repertorio de la catálisis a diversas reacciones químicas. Estudios han identificado moléculas de ARN que pueden catalizar enlaces de aminoacil-ARN en sus propios extremos (2')3' (Illangakekare y col., 1995 Science 267:643-647), y una molécula de ARN que puede transferir un aminoácido de una molécula de ARN a otra (Lohse y col., 1996, Nature 381:442-444).

La publicación de solicitud de patente de EE.UU. 2003/0228593 describe procedimientos de construcción de ribozimas y su uso en la aminoacilación de ARNt con aminoácidos naturalmente codificados y no naturalmente codificados. Formas inmovilizadas en sustrato de moléculas enzimáticas que pueden aminoacilar ARNt, que incluyen, pero no se limitan a, ribozimas, pueden permitir la eficaz purificación por afinidad de los productos aminoacilados. Ejemplos de sustratos adecuados incluyen agarosa, Sepharose y perlas magnéticas. La producción y uso de una forma inmovilizada en sustrato de ribozima para la aminoacilación se describe en Chemistry and Biology 2003, 10:1077-1084 y la publicación de solicitud de patente de EE.UU. 2003/0228593.

Procedimientos de aminoacilación química incluyen aquellos introducidos por Hecht y colaboradores (Hecht, S. M. Acc. Chem. Res. 1992, 25, 545; Heckler, T. G.; Roesser, J. R.; Xu, C.; Chang, P.; Hecht, S. M. Biochemistry 1988, 27, 7254; Hecht, S. M.; Alford, B. L.; Kuroda, Y.; Kitano, S. J. Biol. Chem. 1978, 253, 4517) y por Schultz, Chamberlin, Dougherty y otros (Cornish, V. W.; Mendel, D.; Schultz, P. G. Angew. Chem. Int. Ed. Engl. 1995, 34, 621; Robertson, S. A.; Ellman, J. A.; Schultz, P. G. J. Am. Chem. Soc. 1991, 113, 2722; Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. Science 1989,244, 182; Bain, J. D.; Glabe, C. G.; Dix, T. A.; Chamberlin, A. R. J. Am. Chem. Soc. 1989, 111, 8013; Bain, J. D. y col. Nature 1992, 356, 537; Gallivan, J. P.; Lester, H. A.; Dougherty, D. A: Chem. Biol. 1997, 4, 740; Turcatti, y col. J. Biol. Chem. 1996,271, 19991; Nowak, M. W. y col. Science, 1995,

268, 439; Saks, M. E. y col. J. Biol. Chem. 1996, 271, 23169; Hohsaka, T. y col. J. Am. Chem. Soc. 1999, 121, 34) para evitar el uso de sintetasas en la aminoacilación. Tales procedimientos u otros procedimientos de aminoacilación química pueden usarse para aminoacilar moléculas de ARNt.

Procedimientos para generar ARN catalítico pueden implicar generar conjuntos separados de secuencias de ribozimas aleatorizadas, realizar la evolución dirigida sobre los conjuntos, cribar los conjuntos para actividad de aminoacilación deseable y seleccionar secuencias de aquellas ribozimas que presentan actividad de aminoacilación deseada.

Las ribozimas pueden comprender motivos y/o regiones que facilitan la actividad de acilación, tales como un motivo GGU y una región rica en U. Por ejemplo, se ha informado que regiones ricas en U pueden facilitar el reconocimiento de un sustrato de aminoácido, y un motivo de GGU puede formar pares de bases con los extremos 3' de un ARNt. En combinación, el motivo GGU y la región rica en U facilitan el reconocimiento simultáneo de tanto el aminoácido como el ARNt simultáneamente, y así facilitan la aminoacilación del extremo 3' del ARNt.

Las ribozimas pueden generarse por selección *in vitro* usando un r24 mini parcialmente aleatorizado conjugado con ARNtAsnCCCG, seguido de manipulación sistemática de una secuencia consenso encontrada en los clones activos. Una ribozima a modo de ejemplo obtenida por este procedimiento se llama “ribozima Fx3” y se describe en la solicitud de publicación de EE.UU. nº 2003/0228593 actúa de catalizador versátil para la síntesis de diversos aminoacil-ARNt cargados con aminoácidos no naturales relacionados.

La inmovilización sobre un sustrato puede usarse para permitir la eficaz purificación por afinidad de los ARNt aminoacilados. Ejemplos de sustratos adecuados incluyen agarosa, Sepharose y perlas magnéticas. Las ribozimas pueden inmovilizarse sobre resinas aprovechando la estructura química del ARN, tal como el 3'-cis-diol sobre la ribosa de ARN puede oxidarse con peryodato para dar el dialdehído correspondiente para facilitar la inmovilización del ARN sobre la resina. Pueden usarse diversos tipos de resinas que incluyen resinas de hidrazida económicas en las que la aminación reductora hace de la interacción entre la resina y la ribozima un enlace irreversible. La síntesis de aminoacil-ARNt puede facilitarse significativamente por esta técnica de aminoacilación en columna. Kourouklis y col. Methods 2005; 36:239-4 describen un sistema de aminoacilación basado en columna.

El aislamiento de los ARNt aminoacilados puede llevarse a cabo en una variedad de formas. Un procedimiento adecuado es eluir los ARNt aminoacilados de una columna con un tampón tal como una disolución de acetato sódico con EDTA 10 mM, un tampón que contiene ácido N-(2-hidroxietil)piperazin-N'-(3-propanosulfónico) 50 mM, KCl 12,5 mM, pH 7,0, EDTA 10 mM, o simplemente agua tamponada con EDTA (pH 7,0).

Los ARNt aminoacilados pueden añadirse a reacciones de traducción con el fin de incorporar el aminoácido con el que el ARNt se aminoacilo en una posición de elección en un polipéptido preparado por la reacción de traducción. Ejemplos de sistemas de traducción en los que los ARNt aminoacilados de la presente invención pueden usarse incluyen, pero no se limitan a, lisados celulares. Los lisados celulares proporcionan componentes de reacción necesarios para la traducción *in vitro* de un polipéptido de un ARNm de entrada. Ejemplos de tales componentes de reacción incluyen proteínas ribosómicas, ARNr, aminoácidos, ARNt, GTP, ATP, factores de iniciación y elongación de la traducción y factores adicionales asociados a la traducción. Adicionalmente, los sistemas de traducción pueden ser traducciones discontinuas o traducción compartimentalizada. Los sistemas de traducción discontinua combinan componentes de reacción en un único compartimento mientras que los sistemas de traducción compartimentalizada separan los componentes de la reacción de traducción de los productos de reacción que pueden inhibir la eficiencia de traducción. Tales sistemas de traducción están disponibles comercialmente.

Además, puede usarse un sistema de transcripción/traducción acoplado. Los sistemas de transcripción/traducción acoplados permiten tanto la transcripción de un ADN de entrada en un ARNm correspondiente, que a su vez se traduce mediante los componentes de reacción. Un ejemplo de una transcripción/traducción acoplada comercialmente disponible es el sistema de traducción Rapid (RTS, Roche Inc.). El sistema incluye una mezcla que contiene lisado de *E. coli* para proporcionar componentes traduccionales tales como ribosomas y factores de traducción. Adicionalmente, se incluye una ARN polimerasa para la transcripción del ADN de entrada en un molde de ARNm para su uso en la traducción. RTS puede usar compartimentalización de los componentes de reacción a modo de una membrana interpuesta entre los compartimentos de reacción, que incluye un compartimento de suministro/desecho y un compartimento de transcripción/traducción.

La aminoacilación de ARNt puede realizarse por otros agentes, que incluyen transferasas, polimerasas, anticuerpos catalíticos, proteínas multi-funcionales, y similares.

Stephan en Scientist, 10 de octubre de 2005, páginas 30-33, describe procedimientos adicionales para incorporar aminoácidos no naturalmente codificados en proteínas. Lu y col. en Mol Cell. 2001 Oct;8(4):759-69 describen un procedimiento en el que una proteína se liga químicamente a un péptido sintético que contiene aminoácidos no naturales (ligación de proteínas expresadas).

X. Modificaciones postraduccionales de componentes de aminoácidos no naturales de un polipéptido

Se han desarrollado procedimientos, composiciones, técnicas y estrategias para incorporar específicamente en el sitio aminoácidos no naturales durante la traducción *in vivo* de proteínas. Incorporando un aminoácido no natural con una química de cadena lateral que es ortogonal a aquellos de los aminoácidos que se producen naturalmente, esta tecnología hace posible la derivatización específica del sitio de proteínas recombinantes. Como resultado, un mayor ventaja de los procedimientos, composiciones, técnicas y estrategias descritos en el presente documento es que las proteínas derivatizadas pueden ahora prepararse como productos homogéneos definidos. Sin embargo, los procedimientos, composiciones, mezclas de reacción, técnicas y estrategias descritos en el presente documento no se limitaron a polipéptidos de aminoácidos no naturales formados por técnicas de traducción *in vivo* de proteínas, sino que incluyen polipéptidos de aminoácidos no naturales formados por cualquier técnica, que incluye a modo de ejemplo solo, técnicas *in vitro*, ligación de proteínas expresadas, síntesis química, técnicas basadas en ribozimas (véase, por ejemplo, la sección en el presente documento titulada “Expresión en sistemas alternativos”).

La capacidad para incorporar aminoácidos no naturales en proteínas recombinantes extiende ampliamente las químicas que pueden implementarse para la derivatización postraduccional, en las que tal derivatización se produce tanto *in vivo* como *in vitro.* Más específicamente, la derivatización de proteínas utilizando reacciones de alquilación reductora o de aminación reductora entre un compuesto que contiene carbonilo tal como, a modo de ejemplo, aldehídos, y una amina aromática para formar un enlace de amina alquilada, que incluye una amina secundaria o una amina terciaria, sobre una porción de aminoácido no natural de un polipéptido ofrece varias ventajas. Primera, los aminoácidos que se producen naturalmente (a) no contienen grupos amina aromática que puedan reaccionar con grupos carbonilo para formar aminas aromáticas alquiladas, creando así enlaces amina secundaria o amina terciaria, y (b) los grupos amina aromática pueden reaccionar con grupos que contienen carbonilo para formar aminas alquiladas, que incluyen aminas secundarias o aminas terciarias, y así reactivos diseñados para formar tales aminas alquiladas reaccionarán específicamente en el sitio con el componente de aminoácido no natural del polipéptido (suponiendo, por supuesto, que el aminoácido no natural y el reactivo correspondiente se hayan diseñado para formar aminas alquiladas y el enlace de amina correspondiente). Tal derivatización específica del sitio se ilustra en las FIG. 19-34, en las que la cadena lateral que contiene una amina aromática se alquila preferencialmente reductoramente con respecto a cadenas laterales que contienen aminas protonadas o restos imidizol. Segunda, tales aminas aromáticas alquiladas tienen enlaces amina que son estables en condiciones biológicas, sugiriendo que las proteínas derivatizadas por tales enlaces son candidatos válidos para aplicaciones terapéuticas. Tercera, la estabilidad de tal enlace de amina puede manipularse basándose en la identidad (es decir, los grupos funcionales y/o estructura) del aminoácido no natural con el que se ha formado el enlace. En algunas realizaciones, el enlace de la amina alquilada, que incluye una amina secundaria o una amina terciaria, con el polipéptido de aminoácidos no naturales tiene una semivida de descomposición inferior a una hora, en otras realizaciones inferior a 1 día, en otras realizaciones inferior a 2 días, en otras realizaciones inferior a 1 semana y en otras realizaciones superior a 1 semana. En todavía otras realizaciones, la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante al menos cuatro semanas en condiciones biológicas. En todavía otras realizaciones, la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante al menos tres semanas en condiciones biológicas. En todavía otras realizaciones, la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante al menos dos semanas en condiciones biológicas. En todavía otras realizaciones, la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante al menos una semana en condiciones biológicas. En otras realizaciones, el enlace de la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante al menos 6 días en condiciones biológicas. En otras realizaciones, el enlace de la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante al menos 5 días en condiciones biológicas. En otras realizaciones, el enlace de la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante al menos 4 días en condiciones biológicas. En otras realizaciones, el enlace de la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante al menos 3 días en condiciones biológicas. En otras realizaciones, el enlace de la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante al menos 2 días en condiciones biológicas. En otras realizaciones, el enlace de la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante al menos 1 día en condiciones biológicas. En otras realizaciones, el enlace de la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante hasta 24 horas en condiciones biológicas. En otras realizaciones, el enlace de la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante hasta 12 horas en condiciones biológicas. En otras realizaciones, el enlace de la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante hasta 6 horas en condiciones biológicas. En otras realizaciones, el enlace de la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante hasta 3 horas en condiciones biológicas. En otras realizaciones, el enlace de la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante hasta 2 horas en condiciones biológicas. En otras realizaciones, el enlace de la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante hasta 1 hora en condiciones biológicas. En todavía otras realizaciones, la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante al menos dos semanas bajo condiciones moderadamente ácidas. En otras realizaciones, el enlace de la amina alquilada resultante, que incluye una amina secundaria o una amina terciaria, es estable durante al menos 5 días bajo condiciones moderadamente ácidas. En otras realizaciones, el polipéptido de aminoácidos no naturales es estable

durante al menos 1 día a un pH entre aproximadamente 2 y aproximadamente 8; en otras realizaciones, de un pH de aproximadamente 2 a aproximadamente 6; en otras realización, a un pH de aproximadamente 2 a aproximadamente 4. En otras realizaciones, el polipéptido de aminoácidos no naturales es estable durante al menos 1 día a un pH entre aproximadamente 6 y aproximadamente 10; en otras realizaciones, de un pH de aproximadamente 4 a aproximadamente 8; en otras realización, a un pH de aproximadamente 4 a aproximadamente 10. En otras realizaciones, el polipéptido de aminoácidos no naturales es estable durante al menos 1 día a un pH entre aproximadamente 2 y aproximadamente 5; en otras realizaciones, de un pH de aproximadamente 3 a aproximadamente 7; en otras realización, a un pH de aproximadamente 3 a aproximadamente 10. En otras realizaciones, usando las estrategias, procedimientos, composiciones y técnicas descritos en el presente documento, la síntesis de enlaces de amina alquilada, que incluye amina secundaria o amina terciaria, con un polipéptido de aminoácidos no naturales puede variar para alterar la semivida de descomposición a las necesidades requeridas (por ejemplo, para un uso terapéutico tal como liberación sostenida, o un uso de diagnóstico, productos comerciales, o un uso industrial o un uso miliar).

Los polipéptidos de aminoácidos no naturales descritos anteriormente son útiles para, que incluye, terapéuticos novedosos, diagnósticos, enzimas catalíticas, enzimas industriales, proteínas de unión (incluyendo anticuerpos y fragmentos de anticuerpos), y que incluye el estudio de la estructura y función de proteínas. Véase, por ejemplo, Dougherty, (2000) Unnatural Amino Acids as Probes of Protein Structure and Function, Current Opinion in Chemical Biology, 4:645-652. Otros usos para los polipéptidos de aminoácidos no naturales descritos anteriormente incluyen, a modo de ejemplo solo, usos basados en ensayos, cosméticos, de biología vegetal, medioambientales, de producción de energía y/o militares. Sin embargo, los polipéptidos de aminoácidos no naturales descritos anteriormente pueden someterse a modificaciones adicionales de manera que incorporen funcionalidades nuevas o modificadas, que incluyen manipular la eficacia terapéutica del polipéptido, mejorar el perfil de seguridad del polipéptido, ajustar la farmacocinética, farmacológica y/o farmacodinámica del polipéptido (por ejemplo, aumentando la solubilidad en agua, biodisponibilidad, aumentando la semivida en suero, aumentando la semivida terapéutica, modulando la inmunogenicidad, modulando la actividad biológica, o prolongando el tiempo en circulación), proporcionar funcionalidad adicional al polipéptido, incorporar una etiqueta, marca o señal detectable en el polipéptido, facilitar las propiedades de aislamiento del polipéptido, y cualquier combinación de las modificaciones anteriormente mencionadas.

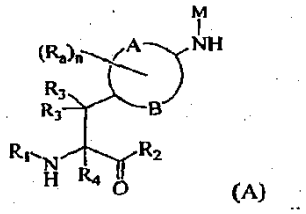
La invención se define por las reivindicaciones. Los procedimientos, composiciones, estrategias y técnicas descritos en el presente documento no se limitan a un tipo particular, clase o familia de polipéptidos o proteínas. De hecho, prácticamente cualquier polipéptido puede incluir al menos un aminoácido no natural descrito en el presente documento. A modo de ejemplo solo, el polipéptido puede ser homólogo a una proteína terapéutica seleccionada del grupo que consiste en: alfa-1 antitripsina, angiostatina, factor antihemolítico, anticuerpo, apolipoproteína, apoproteína, factor natriurético auricular, polipéptido natriurético auricular, péptido auricular, quimiocina C-X-C, T39765, NAP-2, ENA-78, gro-a, gro-b, gro-c, IP-10, GCP-2, NAP-4, SDF-1, PF4, MIG, calcitonina, ligando c-kit, citocina, quimiocina CC, proteína 1 quimioatrayente de monocitos, proteína 2 quimioatrayente de monocitos, proteína 3 quimioatrayente de monocitos, proteína 1 alfa inflamatoria de monocitos, proteína 1 beta inflamatoria de monocitos, RANTES, 1309, R83915, R91733, HCC1, T58847, D31065, T64262, CD40, ligando CD40, ligando c-kit, colágeno, factor estimulante de colonias (CSF), factor 5a del complemento, inhibidor del complemento, receptor 1 del complemento, citocina, péptido 78 activante de neutrófilos epiteliales, MIP-16, MCP-1, factor de crecimiento epidérmico (EGF), péptido activante de neutrófilos epiteliales, eritropoyetina (EPO), toxina exfoliante, factor IX, factor VII, factor VIII, factor X, factor de crecimiento de fibroblastos (FGF), fibrinógeno, fibronectina, proteína de cuatro haces helicoidales, G-CSF, glp-1, GM-CSF, glucocerebrosidasa, gonadotropina, factor de crecimiento, receptor de factores de crecimiento, grf, proteína hedgehog, hemoglobina, factor de crecimiento de hepatocitos (hGF), hirudina, hormona de crecimiento humana (hGH), albúmina de suero humano, ICAM-1, receptor de ICAM-1, LFA-1, receptor de LFA-1, insulina, factor de crecimiento similar a la insulina (IGF), IGF-I, IGF-II, interferón (IFN), IFN-alfa, IFN-beta, IFN-gamma, cualquier molécula similar a interferón o miembro de la familia de IFN, interleucina (IL), IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, factor de crecimiento de queratinocitos (KGF), lactoferrina, factor inhibidor de la leucemia, luciferasa, neurturina, factor inhibidor de neutrófilos (NIF), oncostatina M, proteína osteogénica, producto oncogénico, paracitonina, hormona paratiroidea, PD-ECSF, PDGF, hormona peptídica, pleiotropina, proteína A, proteína G, pth, endotoxina A pirogénica, exotoxina B pirogénica, exotoxina C pirogénica, pyy, relaxina, renina, SCF, proteína biosintética pequeña, receptor I del complemento soluble, I-CAM 1 soluble, receptor de interleucina soluble, receptor de TNF soluble, somatomedina, somatostatina, somatotropina, estreptocinasa, superantígenos, enterotoxina estafilocócica, SEA, SEB, SEC1, SEC2, SEC3, SED, SEE, receptor de hormonas esteroideas, superóxido dismutasa, toxina del síndrome del choque tóxico, timosina alfa 1, activador tisular del plasminógeno, factor de crecimiento tumoral (TGF), factor de necrosis tumoral, factor de necrosis tumoral alfa, factor de necrosis tumoral beta, receptor del factor de necrosis tumoral (TNFR), proteína VLA-4, proteína VCAM-1, factor de crecimiento endotelial vascular (VEGF), urocinasa, 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 corticosterona. El polipéptido de aminoácidos no naturales también puede ser homólogo a cualquier miembro de polipéptido de la familia de supergenes de la hormona de crecimiento.

Tales modificaciones incluyen la incorporación de funcionalidad adicional sobre el componente de aminoácido no natural del polipéptido, que incluye una marca; un colorante; un polímero; un polímero soluble en agua; un derivado

de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos.

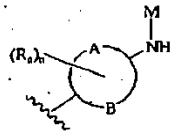
Así, a modo de ejemplo solo, un polipéptido de aminoácidos no naturales que contienen uno cualquiera de los siguientes aminoácidos puede modificarse adicionalmente usando los procedimientos y composiciones descritos en el presente documento:

(a)



(A)

en la que:



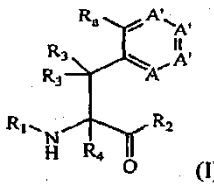
está seleccionado del grupo que consiste en un anillo de arilo monocíclico, un anillo de arilo bicíclico, un anillo de arilo multicíclico, un anillo de heteroarilo monocíclico, un anillo de heteroarilo bicíclico y un anillo de heteroarilo multicíclico;

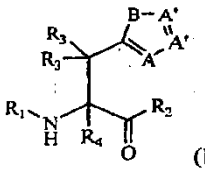
A	es independientemente CR_a o N;
B	es independientemente CR_a, N, O o S;
cada R_a	está seleccionado independientemente del grupo que consiste en H, halógeno, alquilo, $-NO_2$, -CN, alquilo sustituido, $-N(R')_2$, $-C(O)_kR'$, $-C(O)N(R')_2$, -OR' y $-S(O)_kR'$ en la que k es 1, 2 ó 3; y n es 0, 1, 2, 3, 4, 5 ó 6;
R_1	es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y
R_2	es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;
cada uno de R_3	y R_4 es independientemente H, halógeno, alquilo inferior o alquilo inferior sustituido, o R_3 y R_4 o dos grupos R_3 forman opcionalmente un cicloalquilo o un heterocicloalquilo;
M	es H o $-CH_2R_5$; o el resto M-N-C(R_5) puede formar una estructura de anillo de 4 a 7 miembros;
R_5	es alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alquinilo, alquinilo sustituido, alcoxi, alcoxi sustituido, alquilalcoxi, alquilalcoxi sustituido, poli(óxido de alquileno), poli(óxido de alquileno) sustituido, cicloalquilo, cicloalquilo sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido, -C(O)R", -C(O)OR", $-C(O)N(R")_2$, $-C(O)NHCH(R")_2$, -(alquilen o alquilen sustituido)-$N(R")_2$, -(alquenilen o alquenilen sustituido)-$N(R")_2$, -(alquilen o alquilen sustituido)-(arilo o arilo sustituido), -(alquenilen o alquenilen sustituido)-(arilo o arilo sustituido), -(alquilen o alquilen sustituido)-$ON(R")_2$, -(alquilen o alquilen sustituido)-C(O)SR", -(alquilen o alquilen sustituido)-S-S-(arilo o arilo sustituido), en las que cada R" es independientemente hidrógeno, alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alcoxi, alcoxi sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido o -C(O)OR';
o R_5	es L-X, en la que X se selecciona del grupo que consiste en una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o

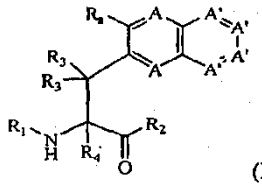
fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos; y L es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno, alquileno sustituido, alquenileno, alquenileno sustituido, -O-, -O-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-O-, -C(O)-, -C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)-, -C(O)N(R')-, -C(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)N(R'), -OC(O)N(R')-, -OC(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-OC(O)N(R')-, -N(R')C(O)-, -NR'C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-NR'C(O)-, -S-, -S-(alquilen o alquilen sustituido)-, $-S(O)_k-$ en la que k es 1, 2 ó 3, $-S(O)_k$(alquilen o alquilen sustituido)-, -C(S)-, -C(S)-(alquilen o alquilen sustituido)-, -CSN(R')-, -CSN(R')-(alquilen o alquilen sustituido)-, -N(R')CO-(alquilen o alquilen sustituido)-, -N(R')C(O)O-, -(alquilen o alquilen sustituido)-O-N=CR'-, -(alquilen o alquilen sustituido)-C(O)NR'-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-$S(O)_k$-(alquilen o alquilen sustituido)-S-, -(alquilen o alquilen sustituido)-S-S-, $-S(O)_kN(R')-$, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, $-N(R')S(O)_kN(R')-$, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, $-C(R')_2-N=N-$ y $-C(R')_2-N(R')-N(R')-$;

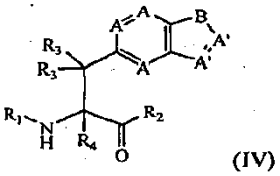
o R_5 y cualquier R_a forman opcionalmente un cicloalquilo o un heterocicloalquilo; y
cada R' es independientemente H, alquilo o alquilo sustituido.

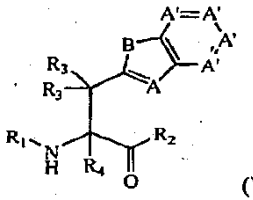
Tales aminoácidos no naturales incluyen los aminoácidos con las siguientes estructuras:


(I)

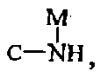

(II)


(III)


(IV)


(V),

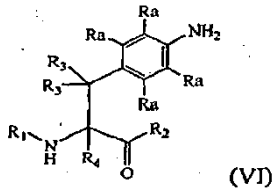
en las que cada A' está seleccionado independientemente de CR_a, N, o



y hasta dos A' pueden ser

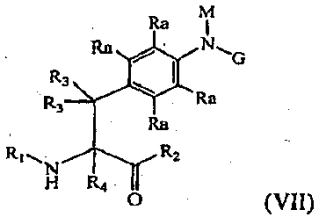
seleccionándose el A' restante de CR_a, o N.

Además, tales aminoácidos no naturales incluyen los aminoácidos con las siguientes estructuras:



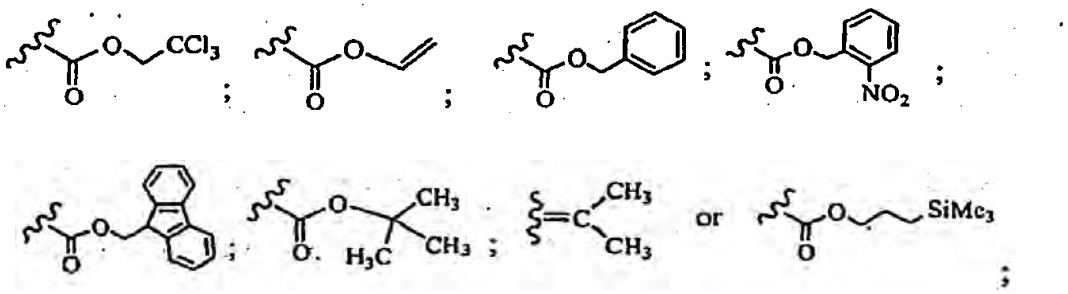
(VI)

y

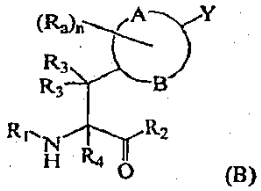


(VII)

en las que G es un grupo protector de amina que incluye

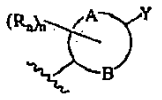


(b)



(B)

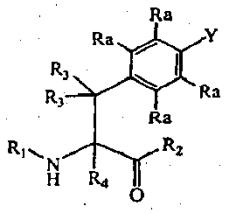
en la que:



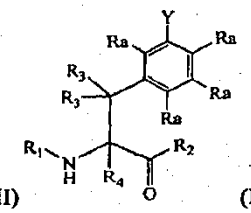
está seleccionado del grupo que consiste en un anillo de arilo monocíclico, un anillo de arilo bicíclico, un anillo de arilo multicíclico, un anillo de heteroarilo monocíclico, un anillo de heteroarilo bicíclico y un anillo de heteroarilo multicíclico;

A	es independientemente CR_a o N;
B	es independientemente CR_a, N, O o S;
cada R_a	está seleccionado independientemente del grupo que consiste en H, halógeno, alquilo, $-NO_2$, -CN, alquilo sustituido, $-N(R')_2$, $-C(O)_kR'$, $-C(O)N(R')_2$, -OR' y $-S(O)_kR'$ en la que k es 1, 2 ó 3; y n es 0, 1, 2, 3, 4, 5 ó 6;
R_1	es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y
R_2	es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;
cada uno de R_3	y R_4 es independientemente H, halógeno, alquilo inferior o alquilo inferior sustituido, o R_3 y R_4 o dos grupos R_3 forman opcionalmente un cicloalquilo o un heterocicloalquilo;
Y	es $-NH-NH_2$, -NH-NHR', -CR'=NR', $-NO_2$ o $-N_3$, y cada R' es independientemente H, alquilo o alquilo sustituido.

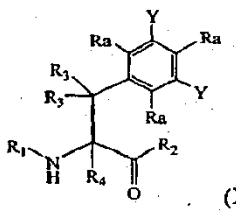
Además, tales aminoácidos no naturales incluyen los aminoácidos con las siguientes estructuras:



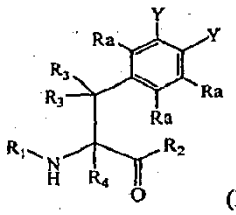
(VIII)



(IX)

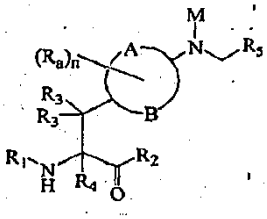


(X)



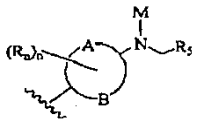
(XI).

(c)



(C)

en la que:

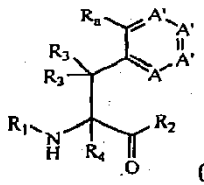


está seleccionado del grupo que consiste en un anillo de arilo monocíclico, un anillo de arilo bicíclico, un anillo de arilo multicíclico, un anillo de heteroarilo monocíclico, un anillo de heteroarilo bicíclico y un anillo de heteroarilo multicíclico;

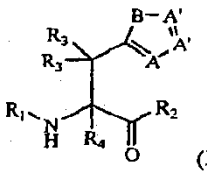
A	es independientemente CR_a o N;
B	es independientemente CR_a, N, O o S;
cada R_a	está seleccionado independientemente del grupo que consiste en H, halógeno, alquilo, $-NO_2$, -CN, alquilo sustituido, $-N(R')_2$, $-C(O)_kR'$, $-C(O)N(R')_2$, -OR' y $-S(O)_kR'$ en la que k es 1, 2 ó 3; y n es 0, 1, 2, 3, 4, 5 ó 6;
R_1	es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y
R_2	es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;
cada uno de R_3	y R_4 es independientemente H, halógeno, alquilo inferior o alquilo inferior sustituido, o R_3 y R_4 o dos grupos R_3 forman opcionalmente un cicloalquilo o un heterocicloalquilo;
M	es H o $-CH_2R_5$; o el resto M-N-C(R_5) puede formar una estructura de anillo de 4 a 7 miembros;
R_5	es alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alquinilo, alquinilo sustituido, alcoxi, alcoxi sustituido, alquilalcoxi, alquilalcoxi sustituido, poli(óxido de alquileno), poli(óxido de alquileno) sustituido, cicloalquilo, cicloalquilo sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido, -C(O)R", -C(O)OR", $-C(O)N(R")_2$, $-C(O)NHCH(R")_2$, -(alquilen o alquilen sustituido)-$N(R")_2$, -(alquenilen o alquenilen sustituido)-$N(R")_2$, -(alquilen o alquilen sustituido)-(arilo o arilo sustituido), -(alquenilen o alquenilen sustituido)-(arilo o arilo sustituido), -(alquilen o alquilen sustituido)-$ON(R")_2$, -(alquilen o alquilen sustituido)-C(O)SR", -(alquilen o alquilen sustituido)-S-S-(arilo o arilo sustituido), en las que cada R" es independientemente hidrógeno, alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alcoxi, alcoxi sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido o -C(O)OR';
o R_5	es L-X, en la que X se selecciona del grupo que consiste en una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo;

una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos; y L es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno, alquileno sustituido, alquenileno, alquenileno sustituido, -O-, -O-(alquilen o alquilen sustituido)-, (alquilen o alquilen sustituido)-O-, -C(O)-, -C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)-, -C(O)N(R')-, -C(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)N(R'), -OC(O)N(R')-, -OC(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-OC(O)N(R')-, -N(R')C(O)-, -NR'C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-NR'C(O)-, -S-, -S-(alquilen o alquilen sustituido)-, $-S(O)_k-$ en la que k es 1, 2 ó 3, $-S(O)_k$(alquilen o alquilen sustituido)-, -C(S)-, -C(S)-(alquilen o alquilen sustituido)-, -CSN(R')-, -CSN(R')-(alquilen o alquilen sustituido)-, -N(R')CO-(alquilen o alquilen sustituido)-, -N(R')C(O)O-, -(alquilen o alquilen sustituido)-O-N=CR'-, -(alquilen o alquilen sustituido)-C(O)NR'-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-$S(O)_k$-(alquilen o alquilen sustituido)-S-, -(alquilen o alquilen sustituido)-S-S-, $-S(O)_kN(R')-$, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, $-N(R')S(O)_kN(R')-$, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, $-C(R')_2-N=N-$ y $-C(R')_2-N(R')-N(R')-$;
o R_5 y cualquier R_a forman opcionalmente un cicloalquilo o un heterocicloalquilo; y cada R' es independientemente H, alquilo o alquilo sustituido.

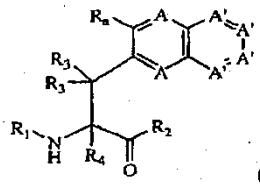
Tales aminoácidos no naturales incluyen los aminoácidos con las siguientes estructuras:



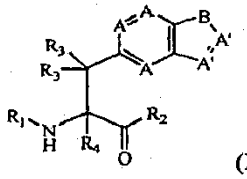
(XII)



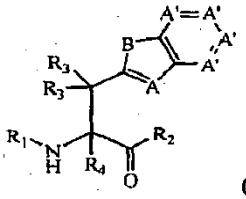
(XIII)



(XIV)

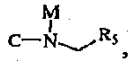


(XV)

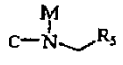


(XVI),

en las que cada A' está seleccionado independientemente de CR_a, N, o

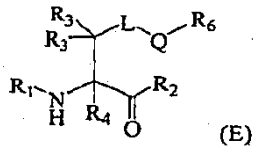


y hasta dos A' puede ser



seleccionándose el A' restante de CR_a o N.

(d)



(E)

en la que:

L es opcional, y cuando está presente es alquileno inferior, alquileno inferior sustituido, cicloalquileno inferior, cicloalquileno inferior sustituido, alquenileno inferior, alquenileno inferior sustituido, alquinileno, heteroalquileno inferior, heteroalquileno sustituido, heterocicloalquileno inferior, heterocicloalquileno inferior sustituido, arileno, arileno sustituido, heteroarileno, heteroarileno sustituido, alcarileno, alcarileno sustituido, aralquileno o aralquileno sustituido;

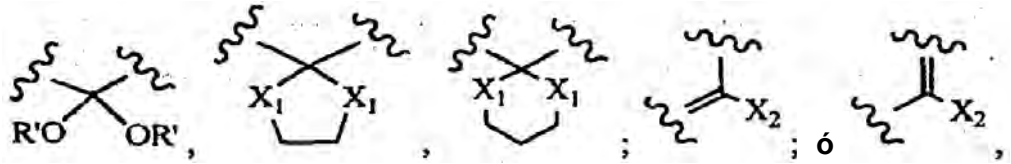
Q es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno inferior, alquileno inferior sustituido, alquenileno inferior, alquenileno inferior sustituido, heteroalquileno inferior, heteroalquileno inferior sustituido, -O-(alquilen o alquilen sustituido)-, -S-(alquilen o alquilen sustituido)-, en la que k es 1, 2 ó 3, $-S(O)_k$(alquilen o alquilen sustituido)-, -C(O)-(alquilen o alquilen sustituido)-, -C(S)-(alquilen o alquilen sustituido)-, -NR'-(alquilen o alquilen sustituido)-, -CON(R")-(alquilen o alquilen sustituido)-, -CSN(R')-(alquilen o alquilen sustituido)-, -N(R')CO-(alquilen o alquilen sustituido)-, y en las que cada R' es independientemente H, alquilo o alquilo sustituido;

R_1 es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y

R_2 es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido; cada uno de R_3 y R_4 es independientemente H, halógeno, alquilo inferior o alquilo inferior sustituido, o R_3 y R_4 o dos grupos R_3 forman opcionalmente un cicloalquilo o un heterocicloalquilo;

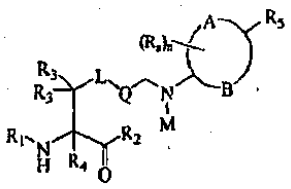
cada R_a está seleccionado independientemente del grupo que consiste en H, halógeno, alquilo, $-NO_2$, -CN, alquilo sustituido, $-N(R')_2$, $-C(O)_kR'$, $-C(O)N(R')_2$, -OR' y $-S(O)_kR'$ en la que k es 1, 2 ó 3;

R_6 es un aldehído protegido o un aldehído enmascarado, en el que el grupo protector incluye, pero no se limita a,



en las que cada X_1 está seleccionado independientemente del grupo que consiste en -O-, -S-, -N(H)-, -N(R)-, -N(Ac)- y -N(OMe)-; X_2 es -OR, -OAc, -SR, $-N(R)_2$, -N(R)(Ac), -N(R)(OMe) o N_3; y en las que cada R' y R es independientemente H, alquilo o alquilo sustituido;

(e)



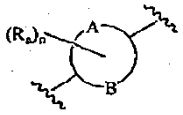
(D)

en la que:

L es opcional, y cuando está presente es alquileno inferior, alquileno inferior sustituido, cicloalquileno inferior, cicloalquileno inferior sustituido, alquenileno inferior, alquenileno inferior sustituido, alquinileno, heteroalquileno inferior, heteroalquileno sustituido, heterocicloalquileno inferior, heterocicloalquileno inferior sustituido, arileno, arileno sustituido, heteroarileno, heteroarileno sustituido, alcarileno, alcarileno sustituido, aralquileno o aralquileno sustituido;

Q es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno inferior, alquileno inferior sustituido, alquenileno inferior, alquenileno inferior sustituido, heteroalquileno inferior, heteroalquileno inferior sustituido, -O-, -O-(alquilen o alquilen sustituido)-, -S-, -S-(alquilen o alquilen sustituido)-, $-S(O)_k$-en la que k es 1, 2 ó 3, $-S(O)_k$(alquilen o alquilen sustituido)-, -C(O)-, -C(O)-(alquilen o alquilen sustituido)-, -C(S)-, -C(S)-(alquilen o alquilen sustituido)-, -N(R')-, -NR'-(alquilen o alquilen sustituido)-, -C(O)N(R')-, -CON(R")-(alquilen o alquilen sustituido)-, -CSN(R")-, -CSN(R')-(alquilen o alquilen

sustituido)-, -N(R')CO-, -N(R')CO-(alquilen o alquilen sustituido)-, -N(R')C(O)O-, -S(O)$_k$N(R')-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R")S(O)$_k$N(R')-, -N(R')-N=, -C(R")=N-, -C(R')=N-N(R')-, -C(R')=N-N=, -C(R')$_2$-N=N- y -C(R')$_2$-N(R')-N(R')-, en las que cada R' es independientemente H, alquilo o alquilo sustituido;
R_1 es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y
R_2 es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido; cada uno de R_3 y R_4 es independientemente H, halógeno, alquilo inferior o alquilo inferior sustituido, o R_3 y R_4 o dos grupos R_3 forman opcionalmente un cicloalquilo o un heterocicloalquilo;
cada R_a está seleccionado independientemente del grupo que consiste en H, halógeno, alquilo, -NO_2, -CN, alquilo sustituido, -N(R')$_2$, -C(O)$_k$R', -C(O)N(R')$_2$, -OR' y -S(O)$_k$R' en la que k es 1, 2 ó 3; n es 0, 1, 2, 3, 4, 5 ó 6;
M es H o -CH_2R_5;
R_5 es alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alquinilo, alquinilo sustituido, alcoxi, alcoxi sustituido, alquilalcoxi, alquilalcoxi sustituido, poli(óxido de alquileno), poli(óxido de alquileno) sustituido, cicloalquilo, cicloalquilo sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido, -C(O)R", -C(O)OR", -C(O)N(R")$_2$, -C(O)NHCH(R")$_2$, -(alquilen o alquilen sustituido)-N(R")$_2$, -(alquenilen o alquenilen sustituido)-N(R")$_2$, -(alquilen o alquilen sustituido)-(arilo o arilo sustituido), -(alquenilen o alquenilen sustituido)-(arilo o arilo sustituido), -(alquilen o alquilen sustituido)-ON(R")$_2$, -(alquilen o alquilen sustituido)-C(O)SR", -(alquilen o alquilen sustituido)-S-S-(arilo o arilo sustituido), en las que cada R" es independientemente hidrógeno, alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alcoxi, alcoxi sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido, o -C(O)OR';
o R_5 es L-X, en la que X se selecciona del grupo que consiste en una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos; y L es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno, alquileno sustituido, alquenileno, alquenileno sustituido, -O-, -O-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-O-, -C(O)-, -C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)-, -C(O)N(R')-, -C(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)N(R'), -OC(O)N(R')-; -OC(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-OC(O)N(R')-, -N(R')C(O)-, -NR'C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-NR'C(O)-, -S-, -S-(alquilen o alquilen sustituido)-, -S(O)$_k$- en la que k es 1, 2 ó 3, -S(O)$_k$(alquilen o alquilen sustituido)-, -C(S)-, -C(S)-(alquilen o alquilen sustituido)-, -CSN(R')-, -CSN(R')-(alquilen o alquilen sustituido)-, -N(R')CO-(alquilen o alquilen sustituido)-, -N(R')C(O)O-, -(alquilen o alquilen sustituido)-O-N=CR'-, -(alquilen o alquilen sustituido)-C(O)NR'-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-S(O)$_k$-(alquilen o alquilen sustituido)-S-, -(alquilen o alquilen sustituido)-S-S-, -S(O)$_k$N(R')-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')S(O)$_k$N(R')-, -N(R')-N= -C(R')=N-, =C(R')=N-N(R')-, -C(R')=N-N=, -C(R')$_2$-N=N- y -C(R')$_2$-N(R')-N(R')-;
o R_5 y cualquier R_a forman opcionalmente un cicloalquilo o un heterocicloalquilo; cada R' es independientemente H, alquilo o alquilo sustituido;



está seleccionado del grupo que consiste en un anillo de arilo monocíclico, un anillo de arilo bicíclico, un anillo de arilo multicíclico, un anillo de heteroarilo monocíclico, un anillo de heteroarilo bicíclico y un anillo de heteroarilo multicíclico;
A es independientemente CR_a o N; y

B es independientemente CR_a, N, O o S.

En un aspecto de los procedimientos y composiciones descritos en el presente documento, las composiciones incluyen al menos un polipéptido con al menos uno, que incluye, pero no se limita a, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, o al menos diez o más aminoácidos no naturales que se han modificado postraduccionalmente. Los aminoácidos no naturales postraduccionalmente modificados pueden ser iguales o diferentes, que incluyen, pero no se limitan a, puede haber 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más sitios diferentes en la proteína que comprenden 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más aminoácidos no naturales postraduccionalmente modificados diferentes. En otro aspecto, una composición incluye un polipéptido con al menos uno, pero menos de todos, de un aminoácido particular presente en el polipéptido está sustituido con el aminoácido no natural postraduccionalmente modificado. Para un polipéptido dado con más de un aminoácido no natural postraduccionalmente modificado, los aminoácidos no naturales postraduccionalmente modificados pueden ser idénticos o diferentes (que incluyen, pero no se limitan a, el polipéptido puede incluir dos o más tipos diferentes de aminoácidos no naturales postraduccionalmente modificados, o puede incluir dos del mismo aminoácido no natural postraduccionalmente modificado). Para un polipéptido dado con más de dos aminoácidos no naturales postraduccionalmente modificados, los aminoácidos no naturales postraduccionalmente modificados pueden ser iguales, diferentes o una combinación de un aminoácido no natural múltiple postraduccionalmente modificado del mismo tipo con al menos un aminoácido no natural postraduccionalmente modificado diferente.

A. Procedimientos para modificar postraduccionalmente polipéptidos de aminoácidos no naturales usando una única etapa de modificación postraduccional: reacciones de alquilación reductora de aminoácidos no naturales que contienen aminas aromáticas con reactivos que contienen carbonilo

Las cadenas laterales de los aminoácidos que se producen naturalmente carecen de sitios altamente nucleófilos. Por tanto, la incorporación de un aminoácido no natural con una cadena lateral que contiene sitios nucleófilos, que incluye, a modo de ejemplo solo, un aminoácido que contiene un grupo de amina aromática o un grupo de amina aromática sustituida, hace posible la alquilación específica del sitio de esta cadena lateral mediante la adición nucleófila a un reactivo que contiene carbonilo, que incluye un reactivo que contiene aldehído, seguido de una reacción de reducción. Esta reacción de alquilación reductora genera un enlace de amina, que incluye aminas secundarias y terciarias.

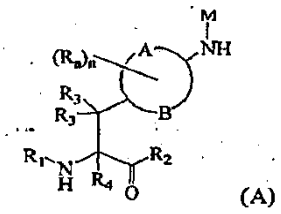
Pueden realizarse procedimientos de derivatización y/o modificación adicional con polipéptidos naturalmente generados o polipéptidos químicamente sintetizados que se han purificado antes de la etapa de alquilación reductora o después de la etapa de alquilación reductora. Además, los procedimientos de derivatización y/o modificación adicional pueden realizarse en polímeros sintéticos, polisacáridos o polinucleótidos que se han purificado antes o después de tales modificaciones.

La modificación postraduccional de polipéptidos basada en alquilación reductora de un polipéptido que contiene amina aromática con un reactivo que contiene aldehído tiene distintas ventajas. Primera, las aminas aromáticas pueden alquilarse reductoramente con compuestos que contienen carbonilo, que incluyen aldehídos y cetonas, en un intervalo de pH de aproximadamente 4 a aproximadamente 10 (y en otras realizaciones en un intervalo de pH de aproximadamente 4 a aproximadamente 7) con un agente reductor, tal como $NaBCNH_3$, para generar enlaces amina secundaria o terciaria. Otros agentes reductores que pueden usarse incluyen, pero no se limitan a, TCEP, Na_2S, $Na_2S_2O_4$, $LiAlH_4$, B_2H_6 y $NaBH_4$. Segunda, bajo estas condiciones de reacción, la química es selectiva para aminoácidos no naturales, ya que las cadenas laterales de aminoácidos que se producen naturalmente no son reactivas. Esto permite la derivatización específica del sitio de polipéptidos que han incorporado aminoácidos no naturales que contienen restos de amina aromática o restos de aldehído protegidos, que incluyen, a modo de ejemplo, proteínas recombinantes. Tales polipéptidos y proteínas derivatizados pueden así prepararse como productos homogéneos definidos. Tercera, las condiciones suaves necesarias para efectuar la reacción de un resto de amina aromática sobre un aminoácido, que se ha incorporado en un polipéptido, con un reactivo que contiene aldehído generalmente no destruyen irreversiblemente la estructura terciaria del polipéptido (exceptuando, por supuesto, si el fin de la reacción es destruir tal estructura terciaria). Similarmente, las condiciones suaves necesarias para efectuar la reacción de un resto de aldehído sobre un aminoácido, que se ha incorporado en un polipéptido y desprotegido, con un reactivo que contiene amina aromática generalmente no destruyen irreversiblemente la estructura terciaria del polipéptido (exceptuando, por supuesto, si el fin de la reacción es destruir tal estructura terciaria). Cuarta, la reacción se produce rápidamente a temperatura ambiente, que permite el uso de muchos tipos de polipéptidos o reactivos que de otro modo serían inestables a mayores temperaturas. Quinta, la reacción se produce fácilmente en condiciones acuosas, permitiendo de nuevo el uso de polipéptidos y reactivos incompatibles (a cualquier grado) con disoluciones no acuosas. Sexta, la reacción se produce fácilmente incluso cuando la relación de polipéptido o aminoácido con respecto a reactivo es estequiométrica, similar a estequiométrica o casi estequiométrica, de manera que es innecesario añadir exceso de reactivo o polipéptido para obtener una cantidad útil de producto de reacción. Séptima, la amina resultante puede producirse regioselectivamente y/o regioespecíficamente, dependiendo del diseño de las porciones de amina y carbonilo de los reactantes. Finalmente, la alquilación reductora de aminas aromáticas con reactivos que contienen aldehído, y la aminación reductora de aldehídos con reactivos que contienen aminas aromáticas, genera enlaces amina, que incluyen amina secundaria y

terciaria, que son estables en condiciones biológicas.

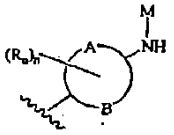
A modo de ejemplo solo, los siguientes aminoácidos no naturales pueden alquilarse reductoramente con reactivos que contienen aldehído descritos en el presente documento,

(a)



(A)

en la que:



está seleccionado del grupo que consiste en un anillo de arilo monocíclico, un anillo de arilo bicíclico, un anillo de arilo multicíclico, un anillo de heteroarilo monocíclico, un anillo de heteroarilo bicíclico y un anillo de heteroarilo multicíclico;

A es independientemente CR_a o N;

B es independientemente CR_a, N, O o S;

cada R_a está seleccionado independientemente del grupo que consiste en H, halógeno, alquilo, $-NO_2$, -CN, alquilo sustituido, $-N(R')_2$, $-C(O)_kR'$, $-C(O)N(R')_2$, -OR' y $-S(O)_kR'$ en la que k es 1, 2 ó 3; y n es 0, 1, 2, 3, 4, 5 ó 6;

R_1 es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y

R_2 es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;

cada uno de R_3 y R_4 es independientemente H, halógeno, alquilo inferior o alquilo inferior sustituido, o R_3 y R_4 o dos grupos R_3 forman opcionalmente un cicloalquilo o un heterocicloalquilo;

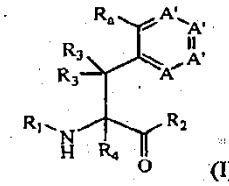
M es H o $-CH_2R_5$, o el resto M-N-C(R_5) puede formar una estructura de anillo de 4 a 7 miembros;

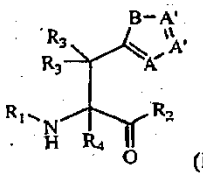
R_5 es alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alquinilo, alquinilo sustituido, alcoxi, alcoxi sustituido, alquilalcoxi, alquilalcoxi sustituido, poli(óxido de alquileno), poli(óxido de alquileno) sustituido, cicloalquilo, cicloalquilo sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido, -C(O)R", -C(O)OR", $-C(O)N(R")_2$, $-C(O)NHCH(R")_2$, -(alquilen o alquilen sustituido)-$N(R")_2$, -(alquenilen o alquenilen sustituido)-$N(R")_2$, -(alquilen o alquilen sustituido)-(arilo o arilo sustituido), -(alquenilen o alquenilen sustituido)-(arilo o arilo sustituido), -(alquilen o alquilen sustituido)-$ON(R")_2$, -(alquilen o alquilen sustituido)-C(O)SR", -(alquilen o alquilen sustituido)-S-S-(arilo o arilo sustituido), en las que cada R" es independientemente hidrógeno, alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alcoxi, alcoxi sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido o -C(O)OR';

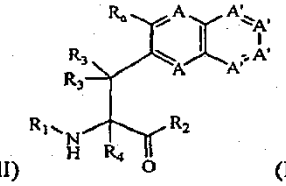
o R_5 es L-X, en la que X se selecciona del grupo que consiste en una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos; y L es opcional, y cuando está presente es un conector

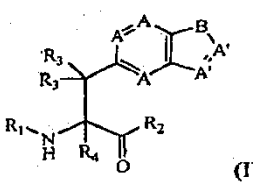
seleccionado del grupo que consiste en alquileno, alquileno sustituido, alquenileno, alquenileno sustituido, -O-, -O-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-O-, -C(O)-, -C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)-, -C(O)N(R')-, -C(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)N(R'), -OC(O)N(R')-, -OC(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-OC(O)N(R')-, -N(R')C(O)-, -NR'C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-NR'C(O)-, -S-, -S-(alquilen o alquilen sustituido)-, $-S(O)_k-$ en la que k es 1, 2 ó 3, $-S(O)_k$(alquilen o alquilen sustituido)-, -C(S)-, -C(S)-(alquilen o alquilen sustituido)-, -CSN(R')-, -CSN(R')-(alquilen o alquilen sustituido)-, -N(R')CO-(alquilen o alquilen sustituido)-, -N(R')C(O)O-, -(alquilen o alquilen sustituido)-O-N=CR'-, -(alquilen o alquilen sustituido)-C(O)NR'-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-$S(O)_k$-(alquilen o alquilen sustituido)-S-, -(alquilen o alquilen sustituido)-S-S-, $-S(O)_kN(R')-$, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, $-N(R')S(O)_kN(R')-$, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, $-C(R')=N-N=$, $-C(R')_2-N=N-$ y $-C(R')_2-N(R')-N(R')-$;
o R_5 y cualquier R_a forman opcionalmente un cicloalquilo o un heterocicloalquilo; y cada R' es independientemente H, alquilo o alquilo sustituido.

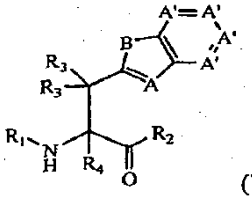
Tales aminoácidos no naturales incluyen, pero no se limitan a, los aminoácidos con las siguientes estructuras:


(I)


(II)


(III)


(IV)


(V),

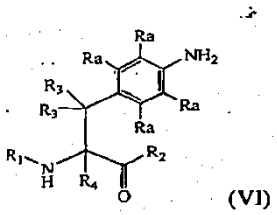
en las que cada A' está seleccionado independientemente de CR_a, N, o

y hasta

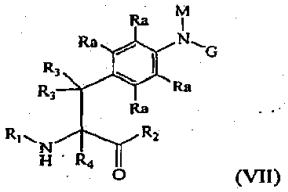
dos A' pueden ser

seleccionándose el A' restante de CR_a o N.

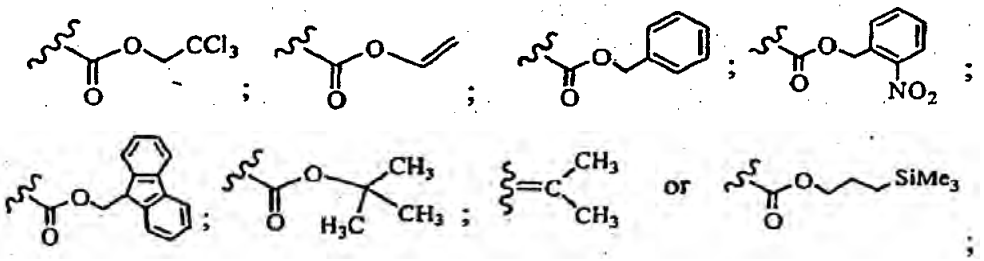
Además, tales aminoácidos no naturales incluyen, pero no se limitan a, los aminoácidos con las siguientes estructuras:


(VI)

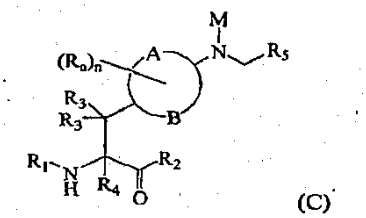
y


(VII)

en la que G es un grupo protector de amina, que incluye, pero no se limita a,

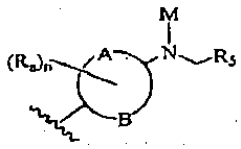


(b)



(C)

en la que:



está seleccionado del grupo que consiste en un anillo de arilo monocíclico, un anillo de arilo bicíclico, un anillo de arilo multicíclico, un anillo de heteroarilo monocíclico, un anillo de heteroarilo bicíclico y un anillo de heteroarilo multicíclico;

A es independientemente CR_a o N;

B es independientemente CR_a, N, O o S;

cada R_a está seleccionado independientemente del grupo que consiste en H, halógeno, alquilo, $-NO_2$, -CN, alquilo sustituido, $-N(R')_2$, $-C(O)_kR'$, $-C(O)N(R')_2$, -OR' y $-S(O)_kR'$ en la que k es 1, 2 ó 3; y n es 0, 1, 2, 3, 4, 5 ó 6;

R_1 es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y

R_2 es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;

cada uno de R_3 y R_4 es independientemente H, halógeno, alquilo inferior o alquilo inferior sustituido, o R_3 y R_4 o dos grupos R_3 forman opcionalmente un cicloalquilo o un heterocicloalquilo;

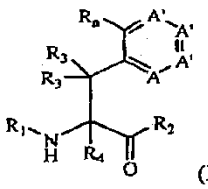
M es H o $-CH_2R_5$; o el resto M-N-C(R_5) puede formar una estructura de anillo de 4 a 7 miembros;

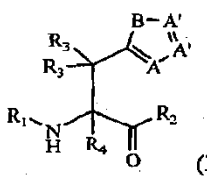
R_5 es alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alquinilo, alquinilo sustituido, alcoxi, alcoxi sustituido, alquilalcoxi, alquilalcoxi sustituido, poli(óxido de alquileno), poli(óxido de alquileno) sustituido, cicloalquilo, cicloalquilo sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido, -C(O)R", -C(O)OR", $-C(O)N(R")_2$, $-C(O)NHCH(R")_2$, -(alquilen o alquilen sustituido)-$N(R")_2$, -(alquenilen o alquenilen sustituido)-$N(R")_2$, -(alquilen o alquilen sustituido)-(arilo o arilo sustituido), -(alquenilen o alquenilen sustituido)-(arilo o arilo sustituido), -(alquilen o alquilen sustituido)-$ON(R")_2$, -(alquilen o alquilen sustituido)-C(O)SR", -(alquilen o alquilen sustituido)-S-S-(arilo o arilo sustituido), en las que cada R" es independientemente hidrógeno, alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alcoxi, alcoxi sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido o -C(O)OR';

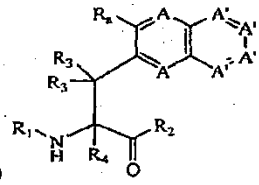
o R_5 es L-X, en la que X se selecciona del grupo que consiste en una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono;

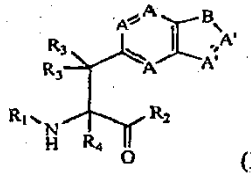
un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos; y L es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno, alquileno sustituido, alquenileno, alquenileno sustituido, -O- -O-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-O-, -C(O)-, -C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)-, -C(O)N(R')-, -C(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)N(R'), -OC(O)N(R')-, -OC(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-OC(O)N(R')-, -N(R')C(O)-, -NR'C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-NR'C(O)-, -S-, -S-(alquilen o alquilen sustituido)-, $-S(O)_k-$ en la que k es 1, 2 ó 3, $-S(O)_k$(alquilen o alquilen sustituido)-, -C(S)-, -C(S)-(alquilen o alquilen sustituido)-, -CSN(R')-, -CSN(R')-(alquilen o alquilen sustituido)-, -N(R')CO-(alquilen o alquilen sustituido)-, -N(R')C(O)O-, -(alquilen o alquilen sustituido)-O-N=CR'-, -(alquilen o alquilen sustituido)-C(O)NR'-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-$S(O)_k$-(alquilen o alquilen sustituido)-S-, -(alquilen o alquilen sustituido)-S-S-, -$S(O)_kN$(R')-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')$S(O)_kN$(R')-, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, $-C(R')_2-N=N-$ y $-C(R')_2-N(R')-N(R')-$;
o R_5 y cualquier R_a forman opcionalmente un cicloalquilo o un heterocicloalquilo; y cada R' es H, alquilo o alquilo sustituido.

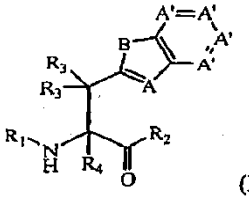
Tales aminoácidos no naturales incluyen, pero no se limitan a, los aminoácidos con las siguientes estructuras:


(XII)

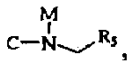

(XIII)


(XIV)

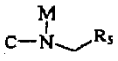

(XV)


(XVI),

en las que cada A' está seleccionado independientemente de CR_a, N o



y hasta dos A' pueden ser



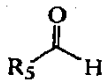
seleccionándose el A' restante de CR_a o N.

Tales reacciones de alquilación reductora modifican postraduccionalmente polipéptidos de aminoácidos no naturales que contienen aminas aromáticas en polipéptidos de aminoácidos no naturales que contienen aminoácidos no naturales que contienen aminas aromáticas mono-alquilados o di-alquilados.

Los tipos de polipéptidos que comprenden tales aminoácidos no naturales que contienen aminas aromáticas son prácticamente ilimitados, en tanto que el aminoácido no natural que contiene amina aromática se localice sobre el polipéptido de manera que el reactivo que contiene aldehído pueda reaccionar con el grupo de amina aromática y no cree un aminoácido no natural modificado resultante que destruya la estructura terciaria del polipéptido (exceptuando, por supuesto, si tal destrucción es el fin de la reacción).

A modo de ejemplo solo, los reactivos que contienen carbonilo que son reactivos con los aminoácidos no naturales que contienen aminas aromáticas descritos en el presente documento, y que también pueden usarse para modificar

adicionalmente los polipéptidos de aminoácidos no naturales que contienen aminas aromáticas, incluyen compuestos que contienen aldehído con la siguiente estructura

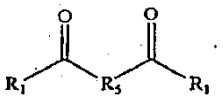


en la que

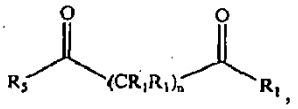
R_5 es alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alquinilo, alquinilo sustituido, alcoxi, alcoxi sustituido, alquilalcoxi, alquilalcoxi sustituido, poli(óxido de alquileno), poli(óxido de alquileno) sustituido, cicloalquilo, cicloalquilo sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido, -C(O)R", -C(O)OR", -C(O)N(R")$_2$, -C(O)NHCH(R")$_2$, -(alquilen o alquilen sustituido)-N(R")$_2$, -(alquenilen o alquenilen sustituido)-N(R")$_2$, -(alquilen o alquilen sustituido)-(arilo o arilo sustituido), -(alquenilen o alquenilen sustituido)-(arilo o arilo sustituido), -(alquilen o alquilen sustituido)-ON(R")$_2$, -(alquilen o alquilen sustituido)-C(O)SR", -(alquilen o alquilen sustituido)-S-S-(arilo o arilo sustituido), en las que cada R" es independientemente hidrógeno, alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alcoxi, alcoxi sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido, o -C(O)OR';

o R_5 es L-X, en la que X se selecciona del grupo que consiste en una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos; y L es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno, alquileno sustituido, alquenileno, alquenileno sustituido, -O-, -O-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-O-, -C(O)-, -C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)-, -C(O)N(R')-, -C(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)N(R'), -OC(O)N(R')-, -OC(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-OC(O)N(R')-, -N(R')C(O)-, -NR'C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-NR'C(O)-, -S-, -S-(alquilen o alquilen sustituido)-, -S(O)$_k$- en la que k es 1, 2 ó 3, -S(O)$_k$(alquilen o alquilen sustituido)-, -C(S)-, -C(S)-(alquilen o alquilen sustituido)-, -CSN(R')-, -CSN(R')-(alquilen o alquilen sustituido)-, -N(R')CO-(alquilen o alquilen sustituido)-, -N(R')C(O)O-, -(alquilen o alquilen sustituido)-O-N=CR'-, -(alquilen o alquilen sustituido)-C(O)NR'-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-S(O)$_k$-(alquilen o alquilen sustituido)-S-, -(alquilen o alquilen sustituido)-S-S-, -S(O)$_k$N(R')-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')S(O)$_k$N(R')-, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, -C(R')$_2$-N=N- y -C(R')$_2$-N(R')-N(R')-; cada R' es independientemente H, alquilo o alquilo sustituido.

Además, a modo de ejemplo solo, los reactivos que contienen carbonilo que son reactivos con los aminoácidos no naturales que contienen aminas aromáticas descritos en el presente documento y que pueden usarse para modificar adicionalmente los polipéptidos de aminoácidos no naturales que contienen aminas aromáticas, son compuestos que contienen dicarbonilo, que incluye dicetonas, cetoaldehídos y dialdehídos, con las siguientes estructuras,



o



en las que

cada R_1 está seleccionado independientemente de H, alquilo opcionalmente sustituido, alqueno opcionalmente sustituido, alquino opcionalmente sustituido, cicloalquilo opcionalmente sustituido, heterociclo opcionalmente sustituido, arilo opcionalmente sustituido o heteroarilo opcionalmente sustituido;

R_5 es alquileno, alquileno sustituido, alquenileno, alquenileno sustituido, alquinileno, alquinileno sustituido, poli(óxido de alquileno), poli(óxido de alquileno) sustituido, cicloalquileno, cicloalquileno sustituido, arileno, arileno sustituido, heteroarileno, heteroarileno sustituido, heterocicloalquileno, heterocicloalquileno sustituido, -C(O)R"-, -C(O)OR"-, -C(O)N(R")-, -(alquilen o alquilen sustituido)-N(R")-, -(alquenilen o alquenilen sustituido)-N(R")-, -(alquilen o alquilen sustituido)-ON(R")-, -(alquilen o alquilen sustituido)-C(O)SR"-, en las que cada R" es independientemente hidrógeno, alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alcoxi, alcoxi sustituido, alquileno sustituido, alquenileno, alquenileno sustituido, alquinileno, alquinileno sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo o aralquilo sustituido;

o R_5 es L-X, en la que X se selecciona del grupo que consiste en una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos; y L es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno, alquileno sustituido, alquenileno, alquenileno sustituido, -O-, -O-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-O-, -C(O)-, -C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)-, -C(O)N(R')-, -C(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)N(R'), -OC(O)N(R')-, -OC(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-OC(O)N(R')-, -N(R')C(O)-, -NR'C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-NR'C(O)-, -S-, -S-(alquilen o alquilen sustituido)-, $-S(O)_k-$ en la que k es 1, 2 ó 3, $-S(O)_k$(alquilen o alquilen sustituido)-, -C(S)-, -C(S)-(alquilen o alquilen sustituido)-, -CSN(R')-, -CSN(R')-(alquilen o alquilen sustituido)-, -N(R')CO-(alquilen o alquilen sustituido)-, -N(R')C(O)O-, -(alquilen o alquilen sustituido)-O-N=CR'-, -(alquilen o alquilen sustituido)-C(O)NR'-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-$S(O)_k$-(alquilen o alquilen sustituido)-S-, -(alquilen o alquilen sustituido)-S-S-, $-S(O)_kN(R')-$, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')$S(O)_kN(R')-$, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, $-C(R')_2-N=N-$ y $-C(R')_2-N(R')-N(R')-$; cada R' es independientemente H, alquilo o alquilo sustituido; n es 1, 2 ó 3.

Los tipos de reactivos que contienen carbonilo que pueden reaccionar con aminoácidos no naturales que contienen aminas aromáticas son prácticamente ilimitados en tanto que el aminoácido no natural que contiene amina aromática se localice sobre el polipéptido de manera que el reactivo que contiene carbonilo pueda reaccionar con el grupo de amina aromática y no cree un aminoácido no natural modificado resultante que destruya la estructura terciaria del polipéptido (exceptuando, por supuesto, si tal destrucción es el fin de la reacción). Tales reactivos que contienen carbonilo incluyen, pero no se limitan a, reactivos que contienen al menos un resto de aldehído y reactivos que contienen al menos dos restos de aldehído.

Adicionalmente, la reacción entre un resto de aldehído y un resto de amina aromática es fácil y tales reacciones de alquilación reductora tienen al menos una de las siguientes características: (i) se produce en un intervalo de pH de aproximadamente 4 a aproximadamente 7, (ii) genera un enlace de amina que es estable en condiciones biológicas; (iii) es específica del sitio; (iv) no destruye irreversiblemente la estructura terciaria de un polipéptido; (v) se produce rápidamente a temperatura ambiente; (vi) se produce fácilmente en condiciones acuosas; (vii) se produce fácilmente

cuando la relación del aminoácido no natural que comprende el resto de amina aromática con respecto al reactante que contiene aldehído es estequiométrica, similar a estequiométrica o casi estequiométrica; y (viii) es regioselectiva y/o regioespecífica. La naturaleza ortogonal de las reacciones de alquilación reductora produce la regioselectividad y/o regioespecificidad, permitiendo así la modificación postraduccional específica del sitio de polipéptidos de aminoácidos no naturales sin afectar otros aminoácidos en el polipéptido que contiene el (los) aminoácido(s) no natural(es).

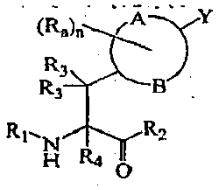
Realizaciones ilustrativas de procedimientos para alquilar reductoramente un aminoácido no natural que contiene amina aromática en un polipéptido se presentan en las FIG. 20-34. Ciertas realizaciones incluyen una única alquilación reductora de un aminoácido no natural que contiene amina aromática en un polipéptido con un reactivo que contiene carbonilo, que incluye un reactivo que contiene aldehído, produciendo un resto de amina secundaria, mientras que otras realizaciones incluyen dos alquilaciones reductoras de un aminoácido no natural que contiene amina aromática sobre un polipéptido con un reactivo que contiene carbonilo, que incluye un reactivo que contiene aldehído, produciendo un resto de amina terciaria. Adicionalmente, ciertas realizaciones incluyen una única alquilación reductora de un aminoácido no natural que contiene amina aromática sobre un polipéptido con un reactivo que contiene al menos dos grupos carbonilo, produciendo así un resto de amina secundaria. Todavía otras realizaciones incluyen alquilaciones reductoras dobles de un aminoácido no natural que contiene amina aromática sobre un polipéptido con un reactivo que contiene al menos dos grupos carbonilo, produciendo así un resto de amina terciaria. En estas realizaciones ilustrativas, un reactivo que contiene aldehído se añade a una disolución tamponada (pH de aproximadamente 4 a aproximadamente 7) de un polipéptido de aminoácidos no naturales que contienen aminas aromáticas y un agente reductor tal como, a modo de ejemplo solo, cianoborohidruro de sodio. La reacción avanza a temperatura ambiente, y el polipéptido de aminoácidos no naturales que contienen aminas aromáticas alquilado resultante puede purificarse por HPLC, FPLC o cromatografía de exclusión por tamaño.

En otras realizaciones, químicas de múltiples conectores pueden reaccionar específicamente en el sitio con un polipéptido de aminoácidos no naturales que contienen aminas aromáticas. En una realización, los procedimientos de conector descritos en el presente documento utilizan conectores que contienen la funcionalidad carbonilo, que incluyen una funcionalidad aldehído, sobre al menos un extremo del conector (mono, bi- o multi-funcional). La alquilación reductora de un polipéptido que contiene amina aromática con un conector derivatizado de aldehído genera un enlace de amina estable. Los conectores bi- y/o multi-funcionales, también conocidos como conectores heterofuncionales (por ejemplo, carbonilo, que incluye un aldehído, con una, o más, de otras químicas de enlace) permiten la conexión específica del sitio de diferentes moléculas (por ejemplo, otros polipéptidos, ácidos polinucleicos, polímeros o moléculas pequeñas) al polipéptido de aminoácidos no naturales, mientras que conectores mono-funcionales, también conocidos como conectores homofuncionales (sustituidos con carbonilo, que incluyen sustituidos con aldehído, en todos los extremos) facilitan la dimerización u oligomerización específica del sitio del polipéptido de aminoácidos no naturales de amina aromática. Combinando esta estrategia de conectores con la tecnología de traducción *in vivo* descrita en el presente documento es posible especificar las estructuras tridimensionales de proteínas químicamente elaboradas.

B. Procedimientos para modificar postraduccionalmente polipéptidos de aminoácidos no naturales usando dos etapas de modificación postraduccional: formación de aminoácidos no naturales que contienen aminas aromáticas seguido de reacciones de alquilación reductora de aminoácidos no naturales que contienen aminas aromáticas con reactivos que contienen carbonilo

Los aminoácidos que contienen aminas aromáticas pueden incorporarse traduccionalmente en polipéptidos antes de la alquilación reductora con reactivos que contienen carbonilo, que incluyen reactivos que contienen aldehído. Alternativamente, precursores de aminoácidos que contienen aminas aromáticas, tales como aminoácidos que contienen restos aromáticos sustituidos, pueden incorporarse traduccionalmente en polipéptidos y posteriormente transformarse en aminoácidos que contienen aminas aromáticas antes de la alquilación reductora con reactivos que contienen carbonilo. El último procedimiento implica dos modificaciones postraduccionales, mientras que el primero solo implica una única modificación postraduccional. Los procedimientos de derivatización y/o modificación adicional pueden realizarse con polipéptidos naturalmente generados o polipéptidos químicamente sintetizados que pueden purificarse antes, o después, de estos procedimientos de modificación. Además, los procedimientos de derivatización y/o modificación adicional pueden realizarse en polímeros sintéticos, polisacáridos o polinucleótidos que puede purificarse antes, o después, de estos procedimientos de modificación.

A modo de ejemplo solo, los siguientes aminoácidos no naturales pueden reducirse para generar aminoácidos que contienen aminas aromáticas



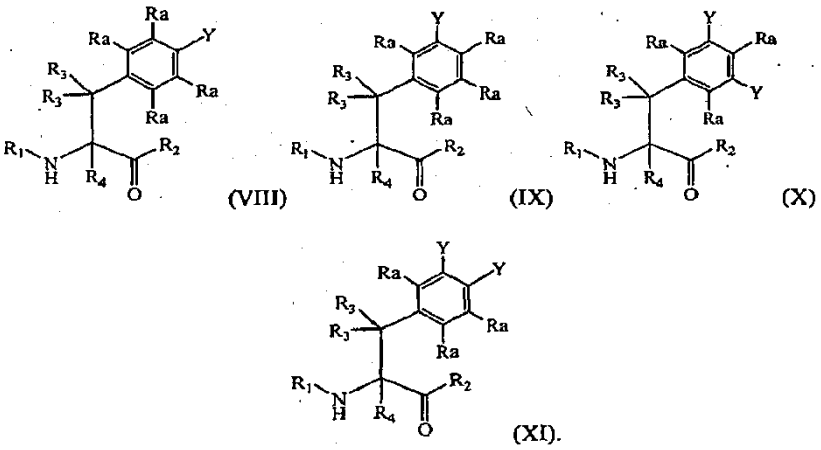
(B)

en la que:

está seleccionado del grupo que consiste en un anillo de arilo monocíclico, un anillo de arilo bicíclico, un anillo de arilo multicíclico, un anillo de heteroarilo monocíclico, un anillo de heteroarilo bicíclico y un anillo de heteroarilo multicíclico;

A	es independientemente CR_a o N;
B	es independientemente CR_a, N, O o S;
cada R_a	está seleccionado independientemente del grupo que consiste en H, halógeno, alquilo, $-NO_2$, -CN, alquilo sustituido, $-N(R')_2$, $-C(O)_kR'$, $-C(O)N(R')_2$, -OR' y $-S(O)_kR'$ en la que k es 1, 2 ó 3; y n es 0, 1, 2, 3, 4, 5 ó 6;
R_1	es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y
R_2	es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;
cada uno de R_3 y R_4	es independientemente H, halógeno, alquilo inferior o alquilo inferior sustituido, o R_3 y R_4 o dos grupos R_3 forman opcionalmente un cicloalquilo o un heterocicloalquilo;
Y	es $-NH-NH_2$, -NH-NHR', -CR'=NR', $-NO_2$ o $-N_3$, y
cada R'	es independientemente H, alquilo o alquilo sustituido.

Tales aminoácidos no naturales incluyen los aminoácidos con la siguiente estructura:



(VIII) (IX) (X)

(XI).

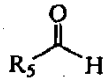
Realizaciones ilustrativas para generar reductoramente un aminoácido no natural que contiene amina aromática sobre un polipéptido natural, polímero sintético, polisacárido, polinucleótido o polipéptido químicamente sintetizado se presentan en las FIG. 14 y 15. Ciertas realizaciones incluyen reducción de sustituyentes de imina en aminoácidos no naturales que contienen amina primaria aromática, sustituyentes de imina en aminoácidos no naturales que contienen amina secundaria aromática, reducción de sustituyentes de amina de hidracina en aminoácidos no naturales que contienen amina primaria aromática, reducción de sustituyentes de amina de hidracina en aminoácidos no naturales que contienen amina secundaria aromática, reducción de sustituyentes nitro en aminoácidos no naturales que contienen amina primaria aromática y reducción de sustituyentes azido en aminoácidos no naturales que contienen amina primaria aromática. En estas realizaciones ilustrativas, restos aromáticos sustituidos se reducen en una disolución tamponada (pH de aproximadamente 4 a aproximadamente 7), y los agentes reductores usados incluyen, pero no se limita a, TCEP, Na_2S, $Na_2S_2O_4$, $LiAlH_4$, B_2H_6, $NaBH_4$ o $NaBCNH_3$. La reacción avanza a temperatura ambiente y el polipéptido natural de aminoácido no natural que contiene amina aromática resultante, polímero sintético, polisacárido, polinucleótido o polipéptido químicamente sintetizado puede purificarse por HPLC, FPLC o cromatografía de exclusión por tamaño.

La alquilación reductora de tal aminoácido que contiene amina aromática generado a partir de la reducción de aminoácidos que contienen restos aromáticos sustituidos es como se ha descrito en la sección *A: Procedimientos para modificar postraduccionalmente polipéptidos de aminoácidos no naturales usando una única etapa de modificación postraduccional: Reacciones de alquilación reductora de aminoácidos no naturales que contienen aminas aromáticas con reactivos que contienen aldehído.* Realizaciones ilustrativas de la alquilación reductora del aminoácido que contiene amina aromática generado mediante reducciones de una amina protegida (o enmascarada) se presentan en las FIG. 15 y 34. La etapa de alquilación reductora sigue la primera etapa postraduccional que genera reductoramente aminoácidos no naturales que contienen aminas aromáticas sobre polipéptidos. Las alquilaciones reductoras son las segundas reacciones postraduccionales que así modifican los polipéptidos de aminoácidos no naturales que contienen aminas aromáticas en polipéptidos de aminoácidos no naturales con

aminoácidos no naturales alquilados que contienen aminas aromáticas. Tales modificaciones postraduccionales o modificaciones post-incorporación también pueden aplicarse a aminoácidos que contienen aminas aromáticas incorporados sobre polímeros sintéticos, polisacáridos, polinucleótidos o polipéptidos químicamente sintetizados.

Los tipos de polipéptidos que comprenden tales aminoácidos no naturales que contienen aminas aromáticas son prácticamente ilimitados, en tanto que el aminoácido no natural que contiene amina aromática se localice sobre el polipéptido de manera que el reactivo que contiene carbonilo, que incluye reactivos que contienen aldehído, pueda reaccionar con el grupo de amina aromática y no cree un aminoácido no natural modificado resultante que destruya la estructura terciaria del polipéptido (exceptuando, por supuesto, si tal destrucción es el fin de la reacción).

A modo de ejemplo solo, los reactivos que contienen aldehído que son reactivos con los aminoácidos no naturales que contienen aminas aromáticas descritos en el presente documento, y también pueden usarse para modificar adicionalmente polipéptidos de aminoácidos no naturales que contienen aminas aromáticas, son compuestos con la siguiente estructura



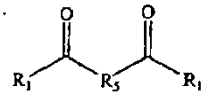
en la que

R_5 es alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alquinilo, alquinilo sustituido, alcoxi, alcoxi sustituido, alquilalcoxi, alquilalcoxi sustituido, poli(óxido de alquileno), poli(óxido de alquileno) sustituido, cicloalquilo, cicloalquilo sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido, -C(O)R'', -C(O)OR'', -$C(O)N(R'')_2$, -$C(O)NHCH(R'')_2$, -(alquilen o alquilen sustituido)-$N(R'')_2$, -(alquenilen o alquenilen sustituido)-$N(R'')_2$, -(alquilen o alquilen sustituido)-(arilo o arilo sustituido), -(alquenilen o alquenilen sustituido)-(arilo o arilo sustituido), -(alquilen o alquilen sustituido)-$ON(R'')_2$, -(alquilen o alquilen sustituido)-C(O)SR'', -(alquilen o alquilen sustituido)-S-S-(arilo o arilo sustituido), en las que cada R'' es independientemente hidrógeno, alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alcoxi, alcoxi sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido o -C(O)OR';

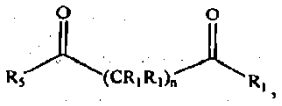
o R_5 es L-X, en la que X se selecciona del grupo que consiste en una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante; un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos; y L es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno, alquileno sustituido, alquenileno, alquenileno sustituido, -O-, -O-(alquilen o alquilen sustituido)-, (alquilen o alquilen sustituido)-O-, -C(O)-, -C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)-, -C(O)N(R')-, -C(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)N(R'), -OC(O)N(R')-, -OC(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-OC(O)N(R')-, -N(R')C(O)-, -NR'C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-NR'C(O)-, -S-, -S-(alquilen o alquilen sustituido)-, -$S(O)_k$- en la que k es 1, 2 ó 3, -$S(O)_k$(alquilen o alquilen sustituido)-, -C(S)-, -C(S)-(alquilen o alquilen sustituido)-, -CSN(R')-, -CSN(R')-(alquilen o alquilen sustituido)-, -N(R')CO-(alquilen o alquilen sustituido)-, -N(R')C(O)O-, -(alquilen o alquilen sustituido)-O-N=CR'-, -(alquilen o alquilen sustituido)-C(O)NR'-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-$S(O)_k$-(alquilen o alquilen sustituido)-S-, -(alquilen o alquilen sustituido)-S-S-, -$S(O)_kN(R')$-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')$S(O)_kN(R')$-, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, -$C(R')_2$-N=N- y -$C(R')_2$-N(R')-N(R')-; cada R' es independientemente H, alquilo o alquilo sustituido.

Además, a modo de ejemplo solo, los reactivos que contienen carbonilo que son reactivos con los aminoácidos no

naturales que contienen aminas aromáticas descritos en el presente documento y que pueden usarse para modificar adicionalmente polipéptidos de aminoácidos no naturales que contienen aminas aromáticas, son compuestos que contienen dicarbonilo, que incluyen dicetonas, cetoaldehídos y dialdehídos, con las siguientes estructuras



o



en las que

cada R_1 está seleccionado independientemente de H, alquilo opcionalmente sustituido, alqueno opcionalmente sustituido, alquino opcionalmente sustituido, cicloalquilo opcionalmente sustituido, heterociclo opcionalmente sustituido, arilo opcionalmente sustituido o heteroarilo opcionalmente sustituido;

R_5 es alquileno, alquileno sustituido, alquenileno, alquenileno sustituido, alquinileno, alquinileno sustituido, poli(óxido de alquileno), poli(óxido de alquileno) sustituido, cicloalquileno, cicloalquileno sustituido, arileno, arileno sustituido, heteroarileno, heteroarileno sustituido, heterocicloalquileno, heterocicloalquileno sustituido, -C(O)R"-, -C(O)OR"-, -C(O)N(R")-, -(alquilen o alquilen sustituido)-N(R")-, -(alquenilen o alquenilen sustituido)-N(R")-, -(alquilen o alquilen sustituido)-ON(R")-, -(alquilen o alquilen sustituido)-C(O)SR"-, en las que cada R" es independientemente hidrógeno, alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alcoxi, alcoxi sustituido, alquileno sustituido, alquenileno, alquenileno sustituido, alquinileno, alquinileno sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo o aralquilo sustituido;

o R_5 es L-X, en la que X se selecciona del grupo que consiste en una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos; y L es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno, alquileno sustituido, alquenileno, alquenileno sustituido, -O-, -O-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-O-, -C(O)-, -C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)-, -C(O)N(R')-, -C(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)N(R'), -OC(O)N(R')-, -OC(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-OC(O)N(R')-, -N(R')C(O)-, -NR'C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-NR'C(O)-, -S-, -S-(alquilen o alquilen sustituido)-, $-S(O)_k-$ en la que k es 1, 2 ó 3, $-S(O)_k$(alquilen o alquilen sustituido)-, -C(S)-, -C(S)-(alquilen o alquilen sustituido)-, -CSN(R')-, -CSN(R')-(alquilen o alquilen sustituido)-, -N(R')CO-(alquilen o alquilen sustituido)-, -N(R')C(O)O-, -(alquilen o alquilen sustituido)-O-N=CR'-, -(alquilen o alquilen sustituido)-C-(O)NR'-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-$S(O)_k$-(alquilen o alquilen sustituido)-S-, -(alquilen o alquilen sustituido)-S-S-, $-S(O)_kN(R')-$, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, $-N(R')S(O)_kN(R')-$, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, $-C(R')_2$-N=N- y $-C(R')_2$-N(R')-N(R')-; cada R' es independientemente H, alquilo o alquilo sustituido; n es 1, 2 ó 3.

Los tipos de reactivos que contienen carbonilo que pueden reaccionar con aminoácidos no naturales que contienen aminas aromáticas son prácticamente ilimitados en tanto que el aminoácido no natural que contiene amina aromática se localice sobre el polipéptido de manera que el reactivo que contiene carbonilo, que incluye reactivos que contienen aldehído, pueda reaccionar con el grupo de amina aromática y no cree un aminoácido no natural

modificado resultante que destruya la estructura terciaria del polipéptido (exceptuando, por supuesto, si tal destrucción es el fin de la reacción). Tales reactivos que contienen carbonilo incluyen, pero no se limitan a, reactivos que contienen al menos un resto carbonilo y/o reactivos que contienen al menos dos restos carbonilo.

Adicionalmente, la reacción entre un resto carbonilo, a modo de ejemplo solo un resto de aldehído, y un resto de amina aromática es fácil y tales reacciones de alquilación reductora tienen al menos una de las siguientes características: (i) se produce en un intervalo de pH de aproximadamente 4 a aproximadamente 7, (ii) genera un enlace de amina que es estable en condiciones biológicas; (iii) es específica del sitio; (iv) no destruye irreversiblemente la estructura terciaria de un polipéptido; (v) se produce rápidamente a temperatura ambiente; (vi) se produce fácilmente en condiciones acuosas; (vii) se produce fácilmente cuando la relación del aminoácido no natural que comprende el resto de amina aromática con respecto al reactante que contiene aldehído es estequiométrica, similar a estequiométrica o casi estequiométrica; y (viii) es regioselectiva y/o regioespecífica. La naturaleza ortogonal de las reacciones de alquilación reductora produce la regioselectividad y/o regioespecificidad, permitiendo así la modificación postraduccional específica del sitio de polipéptidos de aminoácidos no naturales sin afectar otros aminoácidos en el polipéptido de aminoácidos no naturales.

Ciertas realizaciones incluyen una única alquilación reductora del resto de amina aromática con un reactivo que contiene carbonilo que produce un resto de amina secundaria, mientras que otras realizaciones incluyen dos alquilaciones reductoras del resto de amina aromática con dos reactivos que contienen carbonilo que producen un resto de amina terciaria. Los reactivos que contienen carbonilo en esta realización pueden ser idénticos o diferentes. Adicionalmente, ciertas realizaciones incluyen una única alquilación reductora del resto de amina aromática con un reactivo que contiene al menos dos grupos carbonilo, produciendo así un resto de amina secundaria. Todavía otras realizaciones incluyen dos alquilaciones reductoras, también denominada alquilación reductora doble, del resto de amina aromática con un reactivo que contiene al menos dos grupos carbonilo, produciendo así un resto de amina terciaria cíclico. En estas realizaciones ilustrativas, el reactivo que contiene carbonilo se añade a una disolución tamponada (pH de aproximadamente 4 a aproximadamente 7) del aminoácido no natural que contiene amina aromática sobre un polipéptido natural, polímero sintético, polisacárido, polinucleótido o polipéptido químicamente sintetizado y un agente reductor tal como, a modo de ejemplo solo, $NaBCNH_3$, TCEP, Na_2S, $Na_2S_2O_4$, $LiAlH_4$, B_2H_6 y $NaBH_4$. La reacción avanza a temperatura ambiente y el polipéptido de aminoácidos no naturales que contienen aminas aromáticas alquilado resultante puede purificarse por HPLC, FPLC o cromatografía de exclusión por tamaño.

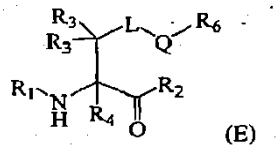
En otras realizaciones, químicas de múltiples conectores pueden reaccionar específicamente en el sitio con un polipéptido de aminoácidos no naturales que contienen aminas aromáticas. En una realización, los procedimientos de conector descritos en el presente documento utilizan conectores que contienen la funcionalidad carbonilo, que incluye funcionalidad aldehído, sobre al menos un extremo del conector (mono, bi- o multi-funcional). La alquilación reductora de un conector derivatizado con carbonilo con un polipéptido que contiene amina aromática genera un enlace de amina estable. Los conectores bi- y/o multi-funcionales, también conocidos como conectores heterofuncionales (por ejemplo, aldehído con una, o más, otras químicas de enlace) permiten la conexión específica del sitio de diferentes moléculas (por ejemplo, otros polipéptidos, ácidos polinucleicos, polímeros o moléculas pequeñas) al polipéptido de aminoácidos no naturales, mientras que los conectores mono-funcionales, también conocidos como conectores homofuncionales (por ejemplo, sustituidos con aldehído en todos los extremos) facilitan la dimerización u oligomerización específica del sitio del polipéptido de aminoácidos no naturales de amina aromática. Combinando esta estrategia de conectores con la tecnología de traducción *in vivo* descrita en el presente documento es posible especificar las estructuras tridimensionales de polipéptidos químicamente elaborados.

C. Procedimientos para modificar postraduccionalmente polipéptidos de aminoácidos no naturales usando dos etapas de modificación postraduccional: desprotección de aminoácidos no naturales que contienen aldehído protegido seguido de aminación reductora de los aminoácidos no naturales que contienen aldehído con reactivos que contienen aminas aromáticas

La incorporación de aminoácidos no naturales con cadena lateral que contiene electrófilo enmascarado o protegido en polipéptidos permite la aminación reductora específica del sitio de las cadenas laterales. Tales aminoácidos no naturales incluyen grupos carbonilo enmascarado o protegido, tales como, a modo de ejemplo solo, aminoácidos que contienen grupos aldehído enmascarados o grupos aldehído protegidos, y en el que la aminación reductora específica del sitio se lleva a cabo, tras el desenmascaramiento o desprotección del aldehído, mediante adición nucleófila de una amina aromática a la cadena lateral que contiene aldehído disponible. Esta reacción de aminación reductora genera un enlace de amina, que incluye aminas secundarias y terciarias. Los procedimientos de derivatización y/o modificación adicional pueden realizarse con polipéptidos naturalmente generados o polipéptidos químicamente sintetizados que se han purificado antes de la etapa de aminación reductora o después de la etapa de aminación reductora. Además, los procedimientos de derivatización y/o modificación adicional pueden realizarse sobre polímeros sintéticos, polisacáridos o polinucleótidos que pueden purificarse antes, o después, de estos procedimientos de modificación.

A modo de ejemplo solo, los siguientes aminoácidos no naturales son el tipo de aminoácidos que contienen restos de aldehído enmascarados/protegidos que pueden incorporarse en polipéptidos, a continuación desenmascararse/desprotegerse antes de aminar reductoramente la cadena lateral que contiene aldehído disponible

para generar un enlace de amina.



(E)

en la que:

L es opcional, y cuando está presente es alquileno inferior, alquileno inferior sustituido, cicloalquileno inferior, cicloalquileno inferior sustituido, alquenileno inferior, alquenileno inferior sustituido, alquinileno, heteroalquileno inferior, heteroalquileno sustituido, heterocicloalquileno inferior, heterocicloalquileno inferior sustituido, arileno, arileno sustituido, heteroarileno, heteroarileno sustituido, alcarileno, alcarileno sustituido, aralquileno o aralquileno sustituido;

Q es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno inferior, alquileno inferior sustituido, alquenileno inferior, alquenileno inferior sustituido, heteroalquileno inferior, heteroalquileno inferior sustituido, -O-(alquilen o alquilen sustituido)-, -S-(alquilen o alquilen sustituido)-, en la que k es 1, 2 ó 3, -$S(O)_k$(alquilen o alquilen sustituido)-, -C(O)-(alquilen o alquilen sustituido)-, -C(S)-(alquilen o alquilen sustituido)-, -NR'-(alquilen o alquilen sustituido)-, -CON(R'')-(alquilen o alquilen sustituido)-, -CSN(R')-(alquilen o alquilen sustituido)-, -N(R')CO-(alquilen o alquilen sustituido)- y en las que R' es independientemente H, alquilo o alquilo sustituido;

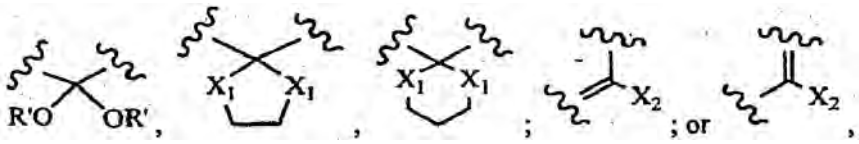
R_1 es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido; y

R_2 es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;

cada uno de R_3 y R_4 es independientemente H, halógeno, alquilo inferior o alquilo inferior sustituido, o R_3 y R_4 o dos grupos R_3, forman opcionalmente un cicloalquilo o heterocicloalquilo;

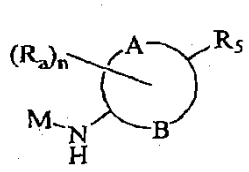
cada R_a está seleccionado independientemente del grupo que consiste en H, halógeno, alquilo, -NO_2, -CN, alquilo sustituido, -$N(R')_2$, -$C(O)_kR'$, -$C(O)N(R')_2$, -OR' y -$S(O)_kR'$ en la que k es 1, 2 ó 3;

R_6 es un aldehído protegido o un aldehído enmascarado, en el que el grupo protector incluye, pero no se limita a,

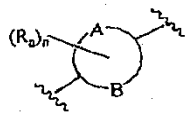


en las que X_1 está seleccionado independientemente del grupo que consiste en -O-, -S-, -N(H)-, -N(R)-, -N(Ac)- y -N(OMe)-; X_2 es -OR, -OAc, -SR, -$N(R)_2$, -N(R)(Ac), -N(R)(OMe) o N_3, y en las que cada R' y R es independientemente H, alquilo o alquilo sustituido. Tales aminoácidos no naturales también pueden estar en forma de una sal, o pueden incorporarse en un polipéptido de aminoácidos no naturales, polímero, o un polinucleótido. Tales aminoácidos no naturales también pueden incorporarse en un polipéptido de aminoácidos no naturales y entonces modificarse postraduccionalmente desprotegiéndose para formar un grupo aldehído *"in situ"*, seguido de aminación reductora del aldehído con un reactivo que contiene aminas aromáticas. Además, aminoácidos no naturales que tienen la estructura de fórmula (E) también pueden incorporarse en un polímero o un polinucleótido y desprotegerse para formar un grupo aldehído *"in situ"*, seguido de aminación reductora del aldehído con un reactivo que contiene aminas aromáticas.

A modo de ejemplo solo, los reactivos que contienen aminas aromáticas que son reactivos con los aminoácidos no naturales que contienen aldehído desenmascarado/desprotegido descritos en el presente documento son compuestos con la siguiente estructura:



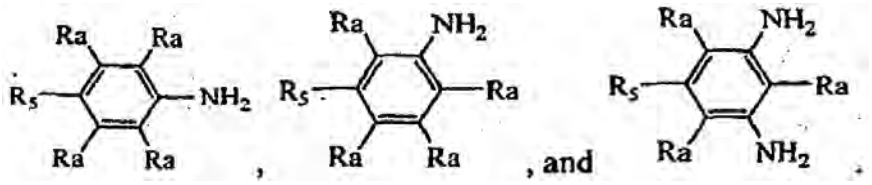
en la que



está seleccionado del grupo que consiste en un anillo de arilo monocíclico, un anillo de arilo bicíclico, un anillo de arilo multicíclico, un anillo de heteroarilo monocíclico, un anillo de heteroarilo bicíclico y un anillo de heteroarilo multicíclico;

A es independientemente CR_a o N;
B es independientemente CR_a, N, O o S;
cada R_a está seleccionado independientemente del grupo que consiste en H, halógeno, alquilo, $-NO_2$, -CN, alquilo sustituido, $-N(R')_2$, $-C(O)_kR'$, $-C(O)N(R')_2$, -OR' y $-S(O)_kR'$ en la que k es 1, 2 ó 3; n es 0, 1, 2, 3, 4, 5 ó 6;
M es H o $-CH_2R_5$;
R_5 es alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alquinilo, alquinilo sustituido, alcoxi, alcoxi sustituido, alquilalcoxi, alquilalcoxi sustituido, poli(óxido de alquileno), poli(óxido de alquileno) sustituido, cicloalquilo, cicloalquilo sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo; aralquilo sustituido, -C(O)R", -C(O)OR", $-C(O)N(R")_2$, $-C(O)NHCH(R")_2$, -(alquilen o alquilen sustituido)-$N(R")_2$, -(alquenilen o alquenilen sustituido)-$N(R")_2$, -(alquilen o alquilen sustituido)-(arilo o arilo sustituido), -(alquenilen o alquenilen sustituido)-(arilo o arilo sustituido), -(alquilen o alquilen sustituido)-$ON(R")_2$, -(alquilen o alquilen sustituido)-C(O)SR", -(alquilen o alquilen sustituido)-S-S-(arilo o arilo sustituido), en las que cada R" es independientemente hidrógeno, alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alcoxi, alcoxi sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido o -C(O)OR';
o R_5 es L-X, en la que X se selecciona del grupo que consiste en una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos; y L es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno, alquileno sustituido, alquenileno, alquenileno sustituido, -O-, -O-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-O-, -C(O)-, -C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)-, -C(O)N(R')-, -C(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)N(R'), -OC(O)N(R')-, -OC(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-OC(O)N(R')-, -N(R')C(O)-, -NR'C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-NR'C(O)-, -S-, -S-(alquilen o alquilen sustituido)-, $-S(O)_k$- en la que k es 1, 2 ó 3, $-S(O)_k$(alquilen o alquilen sustituido)-, -C(S)-, -C(S)-(alquilen o alquilen sustituido)-, -CSN(R')-, -CSN(R')-(alquilen o alquilen sustituido)-, -N(R')CO-(alquilen o alquilen sustituido)-, -N(R')C(O)O-, -(alquilen o alquilen sustituido)-O-N=CR'-, -(alquilen o alquilen sustituido)-C(O)NR'-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-$S(O)_k$-(alquilen o alquilen sustituido)-S-, -(alquilen o alquilen sustituido)-S-S-, $-S(O)_kN(R')$-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, $-N(R')S(O)_kN(R')$-, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, $-C(R')_2$-N=N- y $-C(R')_2$-N(R')-N(R')-; cada R' es independientemente H, alquilo o alquilo sustituido.

Tales reactivos incluyen compuestos que tienen las siguientes estructuras



Estos compuestos también pueden usarse para modificar adicionalmente polipéptidos de aminoácidos no naturales que contienen aldehído (naturalmente o químicamente sintetizados), polímeros sintéticos, polisacáridos o polinucleótidos.

Estas reacciones de aminación reductora modifican postraduccionalmente polipéptidos de aminoácidos no naturales que contienen aldehído desenmascarado/desprotegido en polipéptidos de aminoácidos no naturales que contienen aminoácido no natural que contiene amina aromática. Los tipos de polipéptidos que comprenden tales aminoácidos no naturales que contienen aminas aromáticas son prácticamente ilimitados en tanto que el aminoácido no natural que contiene aldehído desenmascarado/desprotegido se localice sobre el polipéptido de forma que el reactivo que contiene amina aromática pueda reaccionar con el grupo aldehído disponible y no cree un aminoácido no natural modificado resultante que destruya la estructura terciaria del polipéptido (exceptuando, por supuesto, si tal destrucción es el fin de la reacción).

Adicionalmente, la reacción entre un resto de aldehído y un resto de amina aromática es fácil y tales reacciones de aminación reductora tienen al menos una de las siguientes características: (i) se produce en un intervalo de pH de aproximadamente 4 a aproximadamente 7, (ii) genera un enlace de amina que es estable en condiciones biológicas; (iii) es específica del sitio; (iv) no destruye irreversiblemente la estructura terciaria de un polipéptido; (v) se produce rápidamente a temperatura ambiente; (vi) se produce fácilmente en condiciones acuosas; (vii) se produce fácilmente cuando la relación del aminoácido no natural que comprende el resto de aldehído con respecto al reactante que contiene la amina aromática es estequiométrica, similar a estequiométrica o casi estequiométrica; y (viii) es regioselectiva y/o regioespecífica. La naturaleza ortogonal de las reacciones de aminación reductora produce la regioselectividad y/o regioespecificidad, permitiendo así modificación postraduccional específica del sitio de polipéptidos de aminoácidos no naturales sin afectar otros aminoácidos en el polipéptido de aminoácidos no naturales.

Una realización ilustrativa de aminar reductoramente un aminoácido no natural que contiene aldehído desenmascarado o desprotegido sobre un polipéptido se presenta en la FIG. 18. Además, ciertas realizaciones incluyen una única aminación reductora de un aminoácido no natural que contiene aldehído desenmascarado o desprotegido sobre un polipéptido con un reactivo que contiene amina aromática, produciendo así un resto de amina secundaria. Adicionalmente, ciertas realizaciones incluyen una aminación reductora de un aminoácido no natural que contiene carbonilo sobre un polipéptido, en el que el resto carbonilo contiene al menos dos grupos aldehído desenmascarados o desprotegidos, con un reactivo que contiene amina aromática, produciendo así un resto de amina terciaria cíclico. Todavía otras realizaciones incluyen dos aminaciones reductoras de un aminoácido no natural que contiene aldehído desenmascarado o desprotegido sobre un polipéptido que contiene al menos dos grupos aldehído, con dos reactivos que contienen aminas aromáticas idénticos o dos diferentes, produciendo así dos restos de amina secundaria. En estas realizaciones ilustrativas, un reactivo que contiene amina aromática se añade a una disolución tamponada (pH de aproximadamente 4 a aproximadamente 7) de un polipéptido de aminoácidos no naturales que contienen aldehído desenmascarado o desprotegido y un agente reductor tal como, a modo de ejemplo solo, TCEP, Na_2S, $Na_2S_2O_4$, $LiAlH_4$, B_2H_6, $NaBH_4$ o $NaBCNH_3$. La reacción avanza a temperatura ambiente y el polipéptido de aminoácidos no naturales que contienen aldehído aminado resultante puede purificarse por HPLC, FPLC o cromatografía de exclusión por tamaño.

En otras realizaciones, químicas de múltiples conectores pueden reaccionar específicamente en el sitio con un polipéptido de aminoácidos no naturales que contienen aldehído desenmascarado o desprotegido. En una realización, los procedimientos de conector descritos en el presente documento utilizan conectores que contienen la funcionalidad de amina aromática sobre al menos un extremo del conector (mono, bi- o multi-funcional). La aminación reductora de un conector derivatizado con amina aromática con un polipéptido que contiene aldehído desenmascarado o desprotegido genera un enlace de amina estable. Los conectores bi- y/o multi-funcionales, también conocidos como conectores heterofuncionales (por ejemplo, amina aromática con una, o más, otras químicas de enlace) permiten la conexión específica del sitio de diferentes moléculas (por ejemplo, otros polipéptidos, ácidos polinucleicos, polímeros o moléculas pequeñas) al polipéptido de aminoácidos no naturales, mientras que los conectores mono-funcionales, también conocidos como conectores homofuncionales (sustituidos con amina aromática en todos los extremos) facilitan la dimerización u oligomerización específica del sitio del polipéptido de aminoácidos no naturales que contienen aldehído desenmascarado o desprotegido. Combinando esta estrategia de conectores con la tecnología de traducción *in vivo* descrita en el presente documento es posible especificar las estructuras tridimensionales de proteínas químicamente elaboradas.

D. Ejemplo de adición de funcionalidad: Polímeros macromoleculares acoplados a polipéptidos de aminoácidos no naturales

Pueden efectuarse diversas modificaciones a los polipéptidos de aminoácidos no naturales descritos en el presente documento usando las composiciones, procedimientos, técnicas y estrategias descritos en el presente documento. Estas modificaciones incluyen la incorporación de funcionalidad adicional sobre el componente de aminoácido no natural del polipéptido, que incluye, pero no se limita a, una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos. Como un ejemplo no limitante ilustrativo de las composiciones, procedimientos, técnicas y estrategias descritos en el presente documento, la siguiente descripción se basará en añadir polímeros macromoleculares al polipéptido de aminoácidos no naturales con el entendimiento de que las composiciones, procedimientos, técnicas y estrategias descritos de ellos también son aplicables (con modificaciones apropiadas, si fuera necesario, y las cuales podría hacer un experto en la materia con las divulgaciones en el presente documento) para añadir otras funcionalidades, aquellas enumeradas anteriormente.

Una amplia variedad de polímeros macromoleculares y otras moléculas pueden ligarse a los polipéptidos de aminoácidos no naturales descritos en el presente documento para modular propiedades biológicas del polipéptido de aminoácidos no naturales (o el polipéptido de aminoácidos naturales correspondiente), y/o proporcionan nuevas propiedades biológicas al polipéptido de aminoácidos no naturales (o el polipéptido de aminoácidos naturales correspondiente). Estos polímeros macromoleculares pueden ligarse al polipéptido de aminoácidos no naturales mediante el aminoácido no natural, o cualquier sustituyente funcional del aminoácido no natural, o cualquier sustituyente o grupo funcional añadido al aminoácido no natural.

Los polímeros solubles en agua pueden acoplarse a los aminoácidos no naturales incorporados en polipéptidos (naturales o sintéticos), polinucleótidos, polisacáridos o polímeros sintéticos. Los polímeros solubles en agua pueden acoplarse mediante un aminoácido no natural incorporado en el polipéptido o cualquier grupo funcional o sustituyente de un aminoácido no natural, o cualquier grupo funcional o sustituyente añadido a un aminoácido no natural. En algunos casos, los polipéptidos de aminoácidos no naturales descritos en el presente documento comprenden uno o más aminoácidos no naturales ligados a polímeros solubles en agua y uno o más aminoácidos que se producen naturalmente acoplados a polímeros solubles en agua. La unión covalente de polímeros hidrófilos a una molécula biológicamente activa representa un enfoque para aumentar la solubilidad en agua (tal como en un entorno fisiológico), biodisponibilidad, aumentar la semivida en suero, aumentar la semivida terapéutica, modular la inmunogenicidad, modular la actividad biológica, o prologar el tiempo en circulación de la molécula biológicamente activa, que incluye proteínas, péptidos y moléculas particularmente hidrófobas. Características importantes adicionales de tales polímeros hidrófilos pueden incluir biocompatibilidad, ausencia de toxicidad y ausencia de inmunogenicidad. Para el uso terapéutico de la preparación de producto final puede usarse un polímero farmacéuticamente aceptable.

Ejemplos de polímeros hidrófilos incluyen polialquiléteres y análogos tapados con alcoxi de los mismos (por ejemplo, polioxietilenglicol, polioxietileno/propilenglicol, y análogos tapados con metoxi o etoxi de los mismos, especialmente polioxietilenglicol, el último también se conoce como polietilenglicol o PEG); polivinilpirrolidonas; polivinilalquiléteres; polioxazolinas, polialquiloxazolinas y polihidroxialquiloxazolinas; poliacrilamidas, polialquilacrilamidas y polihidroxialquilacrilamidas (por ejemplo, polihidroxipropilmetacrilamida y derivados de la misma); polihidroxialquilacrilatos; ácidos polisiálicos y análogos de los mismos; secuencias de péptidos hidrófilas; polisacáridos y sus derivados, que incluyen dextrano y derivados de dextrano, por ejemplo, carboximetildextrano, sulfatos de dextrano, aminodextrano; celulosa y sus derivados, por ejemplo, carboximetilcelulosa, hidroxialquilcelulosas; quitina y sus derivados, por ejemplo, quitosano, succinilquitosano, carboximetilquitina, carboximetilquitosano; ácido hialurónico y sus derivados; almidones; alginatos; sulfato de condroitina; albúmina; pululano y carboximetilpululano; poliaminoácidos y derivados de los mismos, por ejemplo, ácidos poliglutámicos, polilisinas, ácidos poliaspárticos, poliaspartamidas; copolímeros de anhídrido maleico tales como: copolímero de

estireno-anhídrido maleico, copolímero de éter divinílico-anhídrido maleico; poli(alcoholes vinílicos); copolímeros de los mismos; terpolímeros de los mismos; mezclas de los mismos; y derivados de los anteriores. El polímero soluble en agua puede ser cualquier forma estructural que incluye, pero no se limita a, lineal, en horquilla o ramificada. En algunas realizaciones son particularmente útiles esqueletos de polímeros que son soluble en agua, con de aproximadamente 2 a aproximadamente 300 extremos. Derivados de polímeros multifuncionales incluyen, pero no se limitan a, polímeros lineales que tienen dos extremos, estando cada extremo unido a un grupo funcional que puede ser igual o diferente. En algunas realizaciones, el polímero acuoso comprende un resto de poli(etilenglicol). El peso molecular del polímero puede estar en un amplio intervalo, que incluye entre aproximadamente 100 Da y aproximadamente 100.000 Da o más. El peso molecular del polímero puede estar entre aproximadamente 100 Da y aproximadamente 100.000 Da, que incluye, pero no se limita a, 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 y 100 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 100 Da y aproximadamente 50.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 100 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 1.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 5.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 10.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, la molécula de poli(etilenglicol) es un polímero ramificado. El peso molecular del PEG de cadena ramificada puede estar entre aproximadamente 1.000 Da y aproximadamente 100.000 Da, que incluye 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 y 1.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 1.000 Da y aproximadamente 50.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 1.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 5.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 5.000 Da y aproximadamente 20.000 Da. Aquellos expertos habituales en la materia reconocerán que la lista anterior para esqueletos sustancialmente solubles en agua no es ni mucho menos exhaustiva y es simplemente ilustrativa, y que todos los materiales poliméricos que tienen las cualidades descritas anteriormente se contemplan como que son adecuados para su uso en procedimientos y composiciones descritos en el presente documento.

Como se ha descrito anteriormente, un ejemplo de un polímero hidrófilo es poli(etilenglicol), abreviado PEG, que se ha usado ampliamente en productos farmacéuticos, en implantes artificiales y en otras aplicaciones en las que la biocompatibilidad, ausencia de toxicidad y ausencia de inmunogenicidad son de importancia. Las realizaciones de polímero:polipéptido descritas en el presente documento usarán PEG como un polímero hidrófilo de ejemplo con el entendimiento de que otros polímeros hidrófilos pueden utilizarse similarmente en tales realizaciones.

El PEG es un polímero soluble en agua muy conocido que está comercialmente disponible o puede prepararse por polimerización por apertura de anillo de etilenglicol según procedimientos muy conocidos en la técnica (Sandler y Karo, Polymer Synthesis, Academic Press, New York, vol. 3, páginas 138-161). El PEG normalmente es transparente, incoloro, inodoro, soluble en agua, estable al calor, inerte a muchos agentes químicos, no se hidroliza o deteriora y es generalmente no tóxico. El poli(etilenglicol) se considera que es biocompatible, es decir, que el PEG puede coexistir con tejidos u organismos vivos sin causar daño. Más específicamente, el PEG es sustancialmente no inmunogénico, es decir, que el PEG no tiende a producir una respuesta inmunitaria en el cuerpo. Cuando se une a una molécula que tiene alguna función deseable en el cuerpo, tal como un agente biológicamente activo, el PEG tiende a enmascarar el agente y puede reducir o eliminar cualquier respuesta inmunitaria de manera que un organismo pueda tolerar la presencia del agente. Los conjugados de PEG tienden a no producir una respuesta inmunitaria sustancial o producen coagulación u otro efecto no deseable.

El término “PEG” se usa ampliamente para englobar cualquier molécula de polietilenglicol, sin consideración al tamaño o a modificación en un extremo del PEG, y puede representarse como ligado a un polipéptido de aminoácidos no naturales por la fórmula:

$$XO\text{-}(CH_2CH_2O)_n\text{-}CH_2CH_2\text{-}Y$$

en la que n es 2 a 10.000 y X es H o una modificación terminal, que incluye, pero no se limita a, un alquilo $C_{1\text{-}4}$, un grupo protector o un grupo funcional terminal. El término PEG incluye, pero no se limita a, poli(etilenglicol) en cualquiera de sus formas, que incluye PEG bifuncional, PEG de multibrazos, PEG derivatizado, PEG en horquilla, PEG ramificado (teniendo cada cadena un peso molecular de aproximadamente 1 kDa a aproximadamente 100 kDa, de aproximadamente 1 kDa a aproximadamente 50 kDa, o de aproximadamente 1 kDa a aproximadamente 20 kDa), PEG colgante (es decir, PEG o polímeros relacionados que tienen uno o más grupos funcionales colgantes con respecto al esqueleto del polímero), o PEG con enlaces degradables en su interior. En una realización, el PEG en el que n es de aproximadamente 20 a aproximadamente 2000 es adecuado para su uso en los procedimientos y

composiciones descritos en el presente documento. En algunas realizaciones, el polímero acuoso comprende un resto de poli(etilenglicol). El peso molecular del polímero puede estar en un amplio intervalo, que incluye entre aproximadamente 100 Da y aproximadamente 100.000 Da o más. El peso molecular del polímero puede estar entre aproximadamente 100 Da y aproximadamente 100.000 Da, que incluye, pero no se limita a, 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 y 100 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 100 Da y aproximadamente 50.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 100 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 1.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 5.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 10.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, la molécula de poli(etilenglicol) es un polímero ramificado. El peso molecular del PEG de cadena ramificada puede estar entre aproximadamente 1.000 Da y aproximadamente 100.000 Da, que incluye, pero no se limita a, 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 y 1.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 1.000 Da y aproximadamente 50.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 1.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 5.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 5.000 Da y aproximadamente 20.000 Da. Una amplia variedad de moléculas de PEG se describen en, que incluye el catálogo de Shearwater Polymers, Inc., el catálogo de Nektar Therapeutics.

Ejemplos específicos de grupos funcionales terminales en la bibliografía incluyen carbonato de N-succinimidilo (véanse, por ejemplo, las patentes de EE.UU. nº 5.281.698, 5.468.478), amina (véanse, por ejemplo, Buckmann y col., Makromol. Chem. 182:1379 (1981), Zalipsky y col. Eur. Polym. J. 19:1177 (1983)), hidrazida (véase, por ejemplo, Andresz y col., Makromol. Chem. 179:301 (1978)), propionato de succinimidilo y butanoato de succinimidilo (véanse, por ejemplo, Olson y col. en Poly(ethylene glycol) Chemistry & Biological Applications, pág 170-181, Harris & Zalipsky Eds., ACS, Washington, D.C., 1997; véase también la patente de EE.UU. nº 5.672.662), succinato de succinimidilo (véanse, por ejemplo, Abuchowski y col., Cancer Biochem. Biophys. 7:175 (1984) y Joppich y col., Makromol. Chem. 180:1381 (1979), éster de succinimidilo (véase, por ejemplo, la patente de EE.UU. nº 4.670.417), carbonato de benzotriazol (véase, por ejemplo, la patente de EE.UU. nº 5.650.234), éter de glicidilo (véanse, por ejemplo, Pitha y col., Eur. J Biochem. 94:11 (1979), Elling y col., Biotech. Appl. Biochem. 13:354 (1991), oxicarbonilimidazol (véase, por ejemplo, Beauchamp y col., Anal. Biochem. 131:25 (1983), Tondelli y col., J. Controlled Release 1:251 (1985)), carbonato de p-nitrofenilo (véanse, por ejemplo, Veronese y col., Appl. Biochem. Biotech., 11: 141 (1985); y Sartore y col., Appl. Biochem. Biotech., 27:45 (1991)), aldehído (véanse, por ejemplo, Harris y col., J. Polym. Sci. Chem. Ed. 22:341 (1984), patente de EE.UU. nº 5.824.784, patente de EE.UU. nº 5.252.714), maleimida (véanse, por ejemplo, Goodson y col., Bio/Technology 8:343 (1990), Romani y col. en Chemistry of Peptides and Proteins 2:29 (1984)), y Kogan, Synthetic Comm. 22:2417 (1992)), ortopiridildisulfuro (véase, por ejemplo, Woghiren y col., Bioconj. Chem. 4:314(1993)), acrilol (véase, por ejemplo, Sawhney y col., Macromolecules, 26:581 (1993)), vinilsulfona (véase, por ejemplo, la patente de EE.UU. nº 5.900.461).

En algunos casos, un PEG termina en un extremo con hidroxi o metoxi, es decir, X es H o CH_3 ("metoxi-PEG"). Alternativamente, el PEG puede terminar con un grupo reactivo, formando así un polímero bifuncional. Grupos reactivos típicos pueden incluir aquellos grupos reactivos que se usan comúnmente para reaccionar con los grupos funcionales encontrados en los 20 aminoácidos comunes (incluyendo grupos maleimida, carbonatos activados (incluyendo éster de p-nitrofenilo); ésteres activados (incluyendo N-hidroxisuccinimida, éster de p-nitrofenilo) y aldehídos), además de en grupos funcionales que son inertes a los 20 aminoácidos comunes, pero que reaccionan específicamente con grupos funcionales complementarios presentes en aminoácidos no naturales (incluyendo grupos amina aromática).

Se observa que el otro extremo del PEG, que se muestra en la fórmula anterior por Y, se unirá tanto directamente como indirectamente a un polipéptido (sintético o natural), polinucleótido, polisacárido o polímero sintético mediante un aminoácido no natural. Si Y es un grupo aldehído, entonces el reactivo de PEG que contiene aldehído puede reaccionar con un aminoácido no natural que contiene amina aromática en un polipéptido para formar un grupo PEG ligado al polipéptido mediante un enlace de amina. Ejemplos de condiciones de reacción apropiadas, procedimientos de purificación y reactivos se describen en toda esta memoria descriptiva y las figuras adjuntas. La FIG. 35 presenta un esquema que ilustra la comparación entre la PEGilación del extremo N y la PEGilación basada en la alquilación reductora de restos de amina aromática sobre aminoácidos que se han incorporado en péptidos. La figura ilustra que bajo las condiciones de reacción usadas la PEGilación de un péptido natural solo se consigue fácilmente en la amina terminal, mientras que bajo las condiciones de reacción usadas para la PEGilación de péptidos no naturales con aminoácidos no naturales incorporados que contienen restos aromáticos puede lograrse en el sitio de incorporación del aminoácido no natural en la secuencia, sin ninguna reacción en la amina terminal protonada. El sitio de

PEGilación depende del sitio de incorporación del aminoácido no natural que contiene el resto de amina aromática. Un sitio tal puede ser un sitio terminal, o puede estar en cualquier sitio dentro de la secuencia de péptidos.

A modo de ejemplo solo y no como una limitación a los tipos o clases de reactivos de PEG que pueden usarse con las composiciones, procedimientos, técnicas y estrategias descritos en el presente documento, la FIG. 36 presenta ejemplos ilustrativos de reactivos de PEG que contienen aldehído que pueden reaccionar con polipéptidos de aminoácidos no naturales que contienen aminas aromáticas para formar polipéptidos de aminoácidos no naturales ligados al grupo PEG mediante un enlace de amina. También se presentan en la FIG. 36 ejemplos de reactivos de PEG que contienen aldehído ramificados que pueden reaccionar con polipéptidos de aminoácidos no naturales que contienen aminas aromáticas para formar polipéptidos de aminoácidos no naturales ligados a grupos PEG.

También son particularmente útiles derivados heterobifuncionales cuando se desea unir diferentes moléculas a cada extremo del polímero de PEG. En algunas realizaciones, un extremo del PEG contiene un resto de aldehído que puede reaccionar con un grupo de amina aromática presente en un aminoácido no natural para formar un enlace de amina entre el PEG y el péptido, mientras que el otro extremo del PEG contiene otra funcionalidad que puede someterse adicionalmente a reacción mediante tratamiento con un agente apropiado. A modo de ejemplo, tal funcionalidad puede ser un grupo de amina disponible para actuar de nucleófilo en una variedad de electrófilos, que incluyen reactivos que contienen carbonilo.

Así, en algunas realizaciones, el polipéptido que comprende el aminoácido no natural está ligado a un polímero soluble en agua, tal como polietilenglicol (PEG), mediante la cadena lateral del aminoácido no natural. Los procedimientos y composiciones de aminoácidos no naturales descritos en el presente documento proporcionan un procedimiento altamente eficaz para la modificación selectiva de proteínas con derivados de PEG, que implica la incorporación selectiva de aminoácidos no naturales, que incluye, pero no se limita a, aquellos aminoácidos que contienen grupos funcionales o sustituyentes no encontrados en los 20 aminoácidos naturalmente incorporados, en proteínas en respuesta a un selector de codones y la posterior modificación de aquellos aminoácidos con un derivado de PEG adecuadamente reactivo. Metodologías de química conocidas de una amplia variedad son adecuadas para su uso con los procedimientos y composiciones de aminoácidos no naturales descritos en el presente documento para incorporar un polímero soluble en agua en la proteína.

El esqueleto del polímero puede ser lineal o ramificado. Los esqueletos del polímero ramificados son generalmente conocidos en la técnica. Normalmente, un polímero ramificado tiene un resto de núcleo de la rama central y una pluralidad de cadenas de polímero lineales ligadas al núcleo de la rama central. El PEG se usa en formas ramificadas que pueden prepararse mediante adición de óxido de etileno a diversos polioles, tales como glicerol, oligómeros de glicerol, pentaeritritol y sorbitol. El resto de la rama central también puede derivarse de varios aminoácidos, tales como lisina. El poli(etilenglicol) ramificado puede representarse en forma general como $R(\text{-PEG-OH})_m$ en la que R se deriva de un resto de núcleo, tal como glicerol, oligómeros de glicerol o pentaeritritol, y m representa el número de brazos. También pueden usarse moléculas de PEG de multibrazos, tales como aquellas descritas en las patentes de EE.UU. nº 5.932.462; 5.643.575; 5.229.490; 4.289.872; la solicitud de patente de EE.UU. 2003/0143596; los documentos WO 96/21469 y WO 93/21259, como esqueleto del polímero.

El PEG ramificado también puede estar en forma de un PEG en horquilla representado por $PEG(\text{-}YCHZ_2)_n$ en la que Y es un grupo de enlace y Z es un grupo terminal activado ligado a CH por una cadena de átomos de longitud definida. Todavía otra forma ramificada, el PEG colgante tiene grupo reactivos, tales como carboxilo, a lo largo del esqueleto de PEG en vez de en el extremo de las cadenas de PEG. La FIG. 36 muestra ejemplos no limitantes de reactivos de PEG que contienen aldehído ramificados usados en los procedimientos descritos en el presente documento.

Además de estas formas de PEG, el polímero también puede prepararse con enlaces débiles o degradables en el esqueleto. Por ejemplo, el PEG puede prepararse con enlaces éster en el esqueleto del polímero que están sometidos a hidrólisis. Como se muestra a continuación, esta hidrólisis produce la escisión del polímero en fragmentos de menor peso molecular:

$$\text{-PEG-}CO_2\text{-PEG-} + H_2O \rightarrow \text{PEG-}CO_2H + \text{HO-PEG-}$$

Se entiende por aquellos expertos en la materia que el término poli(etilenglicol) o PEG representa o incluye todas las formas conocidas en la técnica que incluyen las desveladas en el presente documento. El peso molecular del polímero puede estar en un amplio intervalo, que incluye entre aproximadamente 100 Da y aproximadamente 100.000 Da o más. El peso molecular del polímero puede estar entre aproximadamente 100 Da y aproximadamente 100.000 Da, que incluye, pero no se limita a, 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 y 100 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 100 Da y aproximadamente 50.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 100 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 1.000 Da y

aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 5.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 10.000 Da y aproximadamente 40.000 Da.

Con el fin de maximizar las propiedades deseadas del PEG, el peso molecular total y el estado de hidratación del polímero o polímeros de PEG unidos a la molécula biológicamente activa deben ser suficientemente altos para conferir las ventajosas características normalmente asociadas a la unión del polímero de PEG, tales como elevada solubilidad en agua y semivida en circulación, mientras que no se afecte adversamente la bioactividad de la molécula parental.

Los procedimientos y composiciones descritos en el presente documento pueden usarse para producir preparaciones sustancialmente homogéneas de conjugados de polímero:proteína. “Sustancialmente homogéneas” como se usa en el presente documento significa que se observa que las moléculas del conjugado 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 polipéptidos PEGilados “sustancialmente homogéneas” proporcionadas en el presente documento son aquellas que son suficientemente homogéneas para mostrar las ventajas de una preparación homogénea, por ejemplo, facilidad en la aplicación clínica en la predictibilidad de la farmacocinética de lote a lote.

También puede elegirse preparar una mezcla de moléculas de conjugado de polímero:proteína, y la ventaja proporcionada en el presente documento es que puede seleccionarse la proporción de conjugado de mono-polímero:proteína para incluir en la mezcla. Así, si se desea, puede prepararse una mezcla de diversas proteínas con diversos números de restos de polímeros unidos (es decir, di-, tri-, tetra-, etc.) y combinar dichos conjugados con el conjugado de mono-polímero:proteína preparado usando los procedimientos descritos en el presente documento, y tener una mezcla con una proporción predeterminada de conjugados de mono-polímero:proteína.

La proporción de moléculas de polietilenglicol con respecto a moléculas de proteínas variará, como lo harán sus concentraciones en la mezcla de reacción. En general, la relación óptima (en términos de eficiencia de reacción porque hay exceso mínimo de proteína o polímero sin reaccionar) puede determinarse por el peso molecular del polietilenglicol seleccionado y el número de grupos reactivos disponibles disponibles. En lo que se refiere al peso molecular, normalmente cuanto mayor sea el peso molecular del polímero, menor será el número de moléculas de polímero que puedan unirse a la proteína. Similarmente, la ramificación del polímero debe tenerse en cuenta cuando se optimizan estos parámetros. Generalmente, cuanto mayor sea el peso molecular (o cuantas más ramas haya), mayor será la relación de polímero:proteína.

Como se usa en el presente documento, y cuando se contemplan los conjugados de polímero hidrófilo:polipéptido/proteína, el término “cantidad terapéuticamente eficaz” se refiere adicionalmente a una cantidad que da el beneficio deseado a un paciente. La cantidad variará de un individuo a otro y dependerá de varios factores, que incluyen la condición física global del paciente y la afección subyacente que va a tratarse. Una cantidad terapéuticamente eficaz de las presentes composiciones puede determinarse fácilmente por un experto en la materia usando materiales y procedimientos públicamente disponibles.

El número de polímeros solubles en agua ligados a un polipéptido de aminoácidos no naturales modificado o no modificado (es decir, el grado de PEGilación o glucosilación) descrito en el presente documento puede ajustarse para proporcionar una característica farmacológica, farmacocinética o farmacodinámica alterada (que incluye, pero no se limita a, elevada o reducida) tal como semivida *in vivo*. En algunas realizaciones, la semivida del polipéptido aumenta al menos aproximadamente el 10, 20, 30, 40, 50, 60, 70, 80, 90 por ciento, dos veces, cinco veces, 10 veces, 50 veces, o al menos aproximadamente 100 veces con respecto a un polipéptido no modificado.

En una realización, un polipéptido que comprende un aminoácido no natural que contiene amina aromática se modifica con un derivado de PEG que contiene un resto de aldehído terminal que está directamente ligado al esqueleto del PEG. En otra realización, un polipéptido que comprende un aminoácido no natural que contiene amina aromática se modifica con un derivado de PEG ramificado que contiene un resto de aldehído terminal, teniendo cada cadena del PEG ramificado un MW que oscila de aproximadamente 10-40 kDa y, en otras realizaciones, de aproximadamente 5-20 kDa.

En algunas realizaciones, el derivado de PEG con aldehído terminal tendrá la estructura:

$$RO\text{-}(CH_2CH_2O)_n\text{-}O\text{-}(CH_2)_m\text{-}CH_2\text{-}NH\text{-}CH_2\text{-}C(O)H$$

en la que R es un alquilo simple (metilo, etilo, propilo, etc.), m es 2-10 y n es 100-1.000 (es decir, el peso molecular promedio está entre 5-40 kDa). El peso molecular del polímero puede estar en un amplio intervalo, que incluye, pero no se limita a, entre aproximadamente 100 Da y aproximadamente 100.000 Da o más. El peso molecular del polímero puede estar entre aproximadamente 100 Da y aproximadamente 100.000 Da, que incluye 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 y 100 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 100 Da y aproximadamente 50.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 100 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 1.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 5.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del polímero está entre aproximadamente 10.000 Da y aproximadamente 40.000 Da.

En otra realización, un polipéptido que comprende un aminoácido de amina aromática se modifica con un derivado de PEG ramificado que contiene un resto de aldehído terminal formando así un enlace de amina con cada cadena del PEG ramificado que tiene un MW que oscila de aproximadamente 10-40 kDa y, en otras realizaciones, de aproximadamente 5-20 kDa. El peso molecular del PEG de cadena ramificada puede estar entre aproximadamente 1.000 Da y aproximadamente 100.000 Da, que incluye, pero no se limita a, 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 y 1.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 1.000 Da y aproximadamente 50.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 1.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 5.000 Da y aproximadamente 40.000 Da. En algunas realizaciones, el peso molecular del PEG de cadena ramificada está entre aproximadamente 5.000 Da y aproximadamente 20.000 Da.

Están disponibles varias revisiones y monografías sobre la funcionalización y conjugación del PEG. Véanse, por ejemplo, Harris, Macromol. Chem. Phys. C25: 325-373 (1985); Scouten, Methods in Enzymology 135: 30-65 (1987); Wong y col., Enzyme Microb. Technol. 14: 866-874 (1992); Delgado y col., Critical Reviews in Therapeutic Drug Carrier Systems 9: 249-304 (1992); Zalipsky, Bioconjugate Chem. 6: 150-165 (1995).

También pueden encontrarse procedimientos para la activación de polímeros en el documento WO 94/17039, la patente de EE.UU. nº 5.324.844, los documentos WO 94/18247, WO 94/04193, la patente de EE.UU. nº 5.219.564, la patente de EE.UU. nº 5.122.614, el documento WO 90/13540, la patente de EE.UU. nº 5.281.698 y el documento WO 93/15189, y para la conjugación entre polímeros y enzimas activados que incluyen, pero no se limitan a, factor de coagulación VIII (documento WO 94/15625), hemoglobina (documento WO 94/09027), molécula portadora de oxígeno (patente de EE.UU. nº 4.412.989), ribonucleasa y superóxido dismutasa (Veronese y col., App. Biochem. Biotech. 11: 141-52 (1985)).

Si fuera necesario, los polipéptidos de aminoácidos no naturales PEGilados descritos en el presente documento obtenidos a partir de cromatografía hidrófoba pueden purificarse adicionalmente por uno o más procedimientos conocidos para aquellos expertos en la materia que incluyen, pero no se limitan a, cromatografía de afinidad; cromatografía de intercambio aniónico o catiónico (usando, que incluye, pero no se limita a, DEAE-SEPHAROSE); cromatografía sobre sílice; HPLC de fase inversa; filtración en gel (usando, que incluye, pero no se limita a, SEPHADEX G-75); cromatografía de interacción hidrófoba; cromatografía de exclusión por tamaño, cromatografía de quelato de metal; ultrafiltración/diafiltración; precipitación con etanol; precipitación con sulfato de amonio; cromatoenfoque; cromatografía de desplazamiento; procedimientos electroforéticos (que incluyen, pero no se limitan a, isoelectroenfoque preparativo), solubilidad diferencial (que incluye, pero no se limita a, precipitación con sulfato de amonio) o extracción. El peso molecular aparente puede estimarse por GPC por comparación con patrones de proteína globular (Preneta AZ, PROTEIN, PURIFICATION METHODS, A PRACTICAL APPROACH (Harris & Angal, Eds.) IRL Press 1989,293-306). La pureza del conjugado de polipéptido de aminoácidos no naturales:PEG puede evaluarse por degradación proteolítica (que incluye, pero no se limita a, escisión con tripsina), seguido de análisis de espectrometría de masas. Pepinsky RB. Y col., J. Pharmcol. & Exp. Ther. 297(3):1059-66 (2001).

La PEGilación de polipéptidos naturales por alquilación reductora del resto de amina terminal requiere más tiempo que el requerido para la PEGilación de un resto de amina aromática sobre un aminoácido no natural. Adicionalmente, el último procedimiento de PEGilación es específico del sitio y puede producirse en cualquier sitio en el polipéptido en el que el aminoácido no natural se ha incorporado. La FIG. 37 muestra que la PEGilación de una amina terminal de un péptido MT-9 usando relaciones estequiométricas o casi estequiométricas, a pH 5, requiere al menos doce horas para que se produzca aproximadamente el 50 % de PEGilación. A diferencia, la FIG. 38 muestra un ejemplo de la PEGilación de un péptido que contiene un aminoácido que contiene un resto de amina aromática, que incluye pAF2 (p-amino-fenilalanina). En la FIG. 38, el péptido MT-9 está próximo a la PEGilación completa con PEG de 20k-aldehído o PEG de 40 k-aldehído en relaciones estequiométricas o casi estequiométricas, a pH 4, después de reaccionar durante aproximadamente una hora.

Esto demuestra las beneficiosas características de la PEGilación usando alquilación reductora de un aminoácido que contiene un resto de amina aromática, tal como selectividad (solo PEGilación en la amina aromática que contiene aminoácido); tiempo de reacción rápido (aproximadamente una hora frente a doce horas); relaciones estequiométricas o casi estequiométricas en vez de un exceso de reactivo de PEG-aldehído (usando así menos

reactivo); y especificidad (a diferencia de solo la PEGilación del extremo N, la PEGilación mediante alquilación reductora de un aminoácido que contiene un resto de amina aromática puede producirse en cualquier sitio en la secuencia de péptidos).

Un ejemplo no limitante de la PEGilación de hGH ("hormona de crecimiento humana") se muestra en la FIG. 39, en la que el residuo de tirosina en la posición 35 de la proteína hGH se ha sustituido traduccionalmente con un aminoácido que contiene un resto de amina aromática, que incluye pAF2 (p-amino-fenilalanina), que a continuación se alquila reductoramente con un PEG que contiene un resto de aldehído. También se muestra en la FIG. 39 un ejemplo no limitante de la PEGilación de IFNα, en la que un aminoácido que contiene un resto de amina aromática, que incluye pero no se limita a pAF2, se incorpora en la secuencia de IFNα y se alquila posteriormente reductoramente con un PEG ramificado que contiene un resto de aldehído.

Un polímero soluble en agua ligado a un aminoácido no natural de un polipéptido descrito en el presente documento puede derivatizarse o sustituirse adicionalmente sin limitación.

E. Otros ejemplos de modificaciones usando derivatización específica del sitio

Un ejemplo no limitante de la derivatización específica del sitio selectiva de polipéptidos/proteínas se muestra en las FIG. 20-34, en las que la alquilación reductora del resto de amina aromática, que incluye pAF2 (p-amino-fenilalanina), se produce preferencialmente con respecto a otros grupos de la cadena lateral tales como, a modo de ejemplo, lisina, o cualquier derivatización del grupo terminal. Tal derivatización específica del sitio selectiva puede permitir la modificación de polipéptidos/proteínas para diseñar agonistas y/o antagonistas, PEGilación específica del sitio de polipéptidos/proteínas, diseño de profármacos, glucosilación de polipéptidos/proteínas, dimerización de polipéptidos/proteínas, conjugados de fármacos de molécula pequeña de polipéptidos/proteínas y conjugados de fármacos de molécula pequeña de anticuerpos.

F. Potenciamiento de la afinidad por albúmina de suero

También pueden fusionarse diversas moléculas con los polipéptidos de aminoácidos no naturales descritos en el presente documento para modular la semivida en suero. En algunas realizaciones, las moléculas se ligan o fusionan con los polipéptidos de aminoácidos no naturales modificados o no modificados descritos en el presente documento para potenciar la afinidad por albúmina de suero endógena en un animal.

Por ejemplo, en algunos casos, se hace una fusión recombinante de un polipéptido y una secuencia de unión a albúmina. Secuencias de unión a albúmina a modo de ejemplo incluyen el dominio de unión a albúmina de la proteína G estreptocócica (véanse, por ejemplo, Makrides y col., J. Pharmacol. Exp. Ther. 277:534-542 (1996) y Sjolander y col., J, Immunol. Methods 201:115-123 (1997)), o péptidos de unión a albúmina tales como aquellos descritos en, por ejemplo, Dennis, y col., J. Biol. Chem. 277:35035-35043 (2002).

En otras realizaciones, los polipéptidos de aminoácidos no naturales modificados o no modificados descritos en el presente documento se acilaron con ácidos grasos. En algunos casos, los ácidos grasos promueven la unión a albúmina de suero. Véase, por ejemplo, Kurtzhals y col., Biochem. J. 312:725-731 (1995).

En otras realizaciones, los polipéptidos de aminoácidos no naturales modificados o no modificados descritos en el presente documento se fusionan directamente con albúmina de suero (incluyendo albúmina de suero humano). Aquellos expertos en la materia reconocerán que también puede ligarse una amplia variedad de otras moléculas a polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describe en el presente documento, para modular la unión a albúmina de suero u otros componentes del suero.

G. Glucosilación de polipéptidos de aminoácidos no naturales descritos en el presente documento

Los procedimientos y composiciones descritos en el presente documento incluyen polipéptidos que incorporan uno o más aminoácidos no naturales que llevan residuos de sacárido. Los residuos de sacárido pueden ser tanto naturales (que incluyen, pero no se limitan a, N-acetilglucosamina) como no naturales (que incluyen, pero no se limitan a, 3-fluorogalactosa). Los sacáridos pueden ligarse a los aminoácidos no naturales tanto por un enlace glucosídico ligado en N u O (que incluye N-acetilgalactosa-L-serina) como un enlace no natural (que incluye, pero no se limita a, un enlace de heterociclo, que incluye un heterociclo que contiene nitrógeno, o el glucósido ligado en C o S correspondiente).

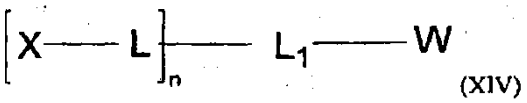
Los restos de sacárido (que incluyen, pero no se limitan a, glucosilo) pueden añadirse a los polipéptidos de aminoácidos no naturales tanto *in vivo* como *in vitro*. En algunas realizaciones, un polipéptido que comprende un aminoácido no natural que contiene amina aromática se modifica con un sacárido derivatizado con un grupo aldehído para generar el polipéptido glucosilado correspondiente ligado mediante un enlace de amina. En otras realizaciones, un polipéptido que comprende un aminoácido no natural que contiene aldehído protegido que, tras la desprotección, se modifica con un sacárido derivatizado con un grupo de amina aromática para generar el polipéptido glucosilado correspondiente ligado mediante un enlace de amina. Una vez unido al aminoácido no

natural, el sacárido puede elaborarse adicionalmente mediante tratamiento con glucosiltransferasas y otras enzimas para generar un oligosacárido unido al polipéptido de aminoácidos no naturales. Véase, por ejemplo, H. Liu y col. J. Am. Chem. Soc. 125: 1702-1703 (2003).

H. Uso de grupos de enlace y aplicaciones, que incluyen dímeros y multímeros de polipéptidos

Además de añadir funcionalidad directamente al polipéptido de aminoácidos no naturales, la porción de aminoácido no natural del polipéptido puede primero modificarse con una molécula de conector multifuncional (por ejemplo, bi-, tri, tetra-) que luego se modifica posteriormente adicionalmente. Es decir, al menos un extremo de la molécula de conector multifuncional reacciona con al menos un aminoácido no natural en un polipéptido y al menos el otro extremo del conector multifuncional está disponible para la funcionalización adicional. Si todos los extremos del conector multifuncional son idénticos, entonces (dependiendo de las condiciones estequiométricas), pueden formarse homomultímeros del polipéptido de aminoácidos no naturales. Si los extremos del conector multifuncional tienen distintas reactividades químicas, entonces al menos un extremo del grupo del conector multifuncional se unirá al polipéptido de aminoácidos no naturales y el otro extremo puede reaccionar posteriormente con una funcionalidad diferente, que incluye a modo de ejemplo solo: una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos.

El grupo del conector multifuncional tiene la estructura general:



(XIV)

en la que:

cada X es un aldehído o una amina aromática;
cada L está seleccionado independientemente del grupo que consiste en alquileno, alquileno sustituido, alquenileno, alquenileno sustituido, -O-, -O-(alquilen o alquilen sustituido)-, -S-, -S-(alquilen o alquilen sustituido)-, $-S(O)_k-$ en la que k es 1, 2 ó 3, $-S(O)_k$(alquilen o alquilen sustituido)-, -C(O)-, -C(O)-(alquilen o alquilen sustituido)-, -C(S)-, -C(S)-(alquilen o alquilen sustituido)-, -N(R')-, -NR'-(alquilen o alquilen sustituido)-, -C(O)N(R')-, -CON(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)NR'C(O)O-(alquilen o alquilen sustituido)-, -O-CON(R')-(alquilen o alquilen sustituido)-, -CSN(R')-, -CSN(R')-(alquilen o alquilen sustituido)-, -N(R')CO-(alquilen o alquilen sustituido)-, -N(R')C(O)O-, -N(R')C(O)O-(alquilen o alquilen sustituido)-, $-S(O)_kN(R')-$, -N(R')C(O)N(R')-, -N(R')C(O)N(R')-(alquilen o alquilen sustituido)-, -N(R')C(S)N(R')-, $-N(R')S(O)_kN(R')-$, N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, $-C(R')_2-N=N-$ y $-C(R')_2-N(R')-N(R')-$;
en las que R' es independientemente H, alquilo o alquilo sustituido;
L_1 es opcional, y cuando está presente, es $-C(R')_p-NR'-C(O)O-$(alquilen o alquilen sustituido)- en la que p es 0, 1, o 2;
W es un aldehído o una amina aromática; y n es 1 a 3.

Los procedimientos y composiciones descritos en el presente documento también proporcionan combinaciones de polipéptidos, tales como homodímeros, heterodímeros, homomultímeros o heteromultímeros (es decir, trímeros, tetrámeros, etc.). A modo de ejemplo solo, la siguiente descripción se basa en los miembros de la familia de supergenes de GH, sin embargo, los procedimientos, técnicas y composiciones descritos en esta sección pueden aplicarse a prácticamente cualquier otro polipéptido que pueda proporcionar beneficio en forma de dímeros y multímeros, que incluyen a modo de ejemplo solo: alfa-1 antitripsina, angiostatina, factor antihemolítico, anticuerpo, apolipoproteína, apoproteína, factor natriurético auricular, polipéptido natriurético auricular, péptido auricular, quimiocina C-X-C, T39765, NAP-2, ENA-78, gro-a, gro-b, gro-c, IP-10, GCP-2, NAP-4, SDF-1, PF4, MIG, calcitonina, ligando c-kit, citocina, quimiocina CC, proteína 1 quimioatrayente de monocitos, proteína 2 quimioatrayente de

monocitos, proteína 3 quimioatrayente de monocitos, proteína 1 alfa inflamatoria de monocitos, proteína 1 beta inflamatoria de monocitos, RANTES, 1309, R83915, R91733, HCCl, T58847, D31065, T64262, CD40, ligando CD40, ligando c-kit, colágeno, factor estimulante de colonias (CSF), factor 5a del complemento, inhibidor del complemento, receptor 1 del complemento, citocina, péptido 78 activante de neutrófilos epiteliales, MIP-16, MCP-1, factor de crecimiento epidérmico (EGF), péptido activante de neutrófilos epiteliales, eritropoyetina (EPO), toxina exfoliante, factor IX, factor VII, factor VIII, factor X, factor de crecimiento de fibroblastos (FGF), fibrinógeno, fibronectina, proteína de cuatro haces helicoidales, G-CSF, glp-1, GM-CSF, glucocerebrosidasa, gonadotropina, factor de crecimiento, receptor de factores de crecimiento, grf, proteína hedgehog, hemoglobina, factor de crecimiento de hepatocitos (hGF), hirudina, hormona de crecimiento humana (hGH), albúmina de suero humano, ICAM-1, receptor de ICAM-1, LFA-1, receptor de LFA-1, insulina, factor de crecimiento similar a la insulina (IGF), IGF-I, IGF-II, interferón (IFN), IFN-alfa, IFN-beta, IFN-gamma, cualquier molécula similar a interferón o miembro de la familia de IFN, interleucina (IL), IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, factor de crecimiento de queratinocitos (KGF), lactoferrina, factor inhibidor de la leucemia, luciferasa, neurturina, factor inhibidor de neutrófilos (NIF), oncostatina M, proteína osteogénica, producto oncogénico, paracitonina, hormona paratiroidea, PD-ECSF, PDGF, hormona peptídica, pleiotropina, proteína A, proteína G, pth, endotoxina A pirogénica, exotoxina B pirogénica, exotoxina C pirogénica, pyy, relaxina, renina, SCF, proteína biosintética pequeña, receptor I del complemento soluble, I-CAM 1 soluble, receptor de interleucina soluble, receptor de TNF soluble, somatomedina, somatostatina, somatotropina, estreptocinasa, superantígenos, enterotoxina estafilocócica, SEA, SEB, SEC1, SEC2, SEC3, SED, SEE, receptor de hormonas esteroideas, superóxido dismutasa, toxina del síndrome del choque tóxico, timosina alfa 1, activador tisular del plasminógeno, factor de crecimiento tumoral (TGF), factor de necrosis tumoral, factor de necrosis tumoral alfa, factor de necrosis tumoral beta, receptor del factor de necrosis tumoral (TNFR), proteína VLA-4, proteína VCAM-1, factor de crecimiento endotelial vascular (VEGF), urocinasa, 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 corticosterona.

Así, englobado dentro de los procedimientos, técnicas y composiciones descritos en el presente documento está un polipéptido miembro de la familia de supergenes de GH que contiene uno o más aminoácidos no naturales unidos a otro miembro de la familia de supergenes de GH o variante del mismo o cualquier otro polipéptido que sea un miembro no de la familia de supergenes de GH o variante del mismo, tanto directamente al esqueleto del polipéptido como mediante un conector. Debido a su elevado peso molecular en comparación con los monómeros, los conjugados de dímeros o multímeros de miembros de la familia de supergenes de GH pueden presentar propiedades nuevas o deseables, que incluyen diferente farmacológica, farmacocinética, farmacodinámica, semivida terapéutica modulada o semivida en plasma modulada con respecto al miembro de la familia de supergenes de GH monomérico. En algunas realizaciones, los dímeros de miembros de la familia de supergenes de GH descritos en el presente documento modularán la dimerización del receptor de miembros de la familia de supergenes de GH. En otras realizaciones, los dímeros o multímeros de miembros de la familia de supergenes de GH descritos en el presente documento actuarán de antagonista, agonista o modulador de receptores de miembros de la familia de supergenes de GH.

En algunas realizaciones, los polipéptidos del miembro de la familia de supergenes de GH están ligados directamente, que incluye, pero no se limita a, mediante un enlace de amida Asn-Lys o enlace de disulfuro Cys-Cys. En algunas realizaciones, los polipéptidos del miembro de la familia de supergenes de GH ligados, y/o el miembro no de la familia de supergenes de GH ligado, comprenderán diferentes aminoácidos no naturales para facilitar la dimerización, incluyendo un primer miembro de la familia de supergenes de GH, y/o el miembro no de la familia de supergenes de GH ligado, comprendiendo el polipéptido un aminoácido no natural que contiene amina aromática conjugado con un segundo polipéptido del miembro de la familia de supergenes de GH que comprende un aminoácido no natural que contiene aldehído y los polipéptidos se hacen reaccionar mediante alquilación reductora, formando un enlace de amina entre los dos.

Alternativamente, los dos polipéptidos del miembro de la familia de supergenes de GH, y/o los miembros no de la familia de supergenes de GH ligados, se ligan mediante un conector. Puede usarse cualquier conector hetero- u homo-bifuncional para ligar los dos miembros de la familia de supergenes de GH, y/o los miembros no de la familia de supergenes de GH ligados, polipéptidos que pueden tener la misma secuencia primaria o diferente. En algunos casos, el conector usado para unir los polipéptidos del miembro de la familia de supergenes de GH, y/o del miembro no de la familia de supergenes de GH ligado, juntos puede ser un reactivo de PEG bifuncional.

En algunas realizaciones, los procedimientos y composiciones descritos en el presente documento proporcionan conectores bifuncionales solubles en agua que tienen una estructura de mancuerna que incluye: a) resto que contiene aldehído sobre al menos un primer extremo de un polímero esqueleto; y b) al menos un segundo grupo funcional sobre un segundo extremo del esqueleto del polímero. El segundo grupo funcional puede ser igual o diferente del primer grupo funcional. El segundo grupo funcional, en algunas realizaciones, no es reactivo con el primer grupo funcional. Los procedimientos y composiciones descritos en el presente documento proporcionan, en algunas realizaciones, compuestos solubles en agua que comprenden al menos un brazo de una estructura molecular ramificada. Por ejemplo, la estructura molecular ramificada puede ser dendrítica.

En algunas realizaciones, los procedimientos y composiciones descritos en el presente documento proporcionan

multímeros que comprenden uno o más miembros de la familia de supergenes de GH formados mediante reacciones con polímeros activados solubles en agua que tienen la estructura:

$$R\text{-}(CH_2CH_2O)_n\text{-}O\text{-}(CH_2)_m\text{-}X$$

en la que n es de aproximadamente 5 a aproximadamente 3.000, m es 2-10, X puede ser un resto que contiene carbonilo (incluyendo un aldehído) y R es un grupo de encapuchado, un grupo funcional o un grupo saliente que puede ser igual o diferente de X. R puede ser, por ejemplo, un grupo funcional seleccionado del grupo que consiste en hidroxilo, hidroxilo protegido, alcoxilo, éster de N-hidroxisuccinimidilo, éster de 1-benzotriazolilo, carbonato de N-hidroxisuccinimidilo, carbonato de 1-benzotriazolilo, acetal, aldehído, hidratos de aldehído, alquenilo, acrilato, metacrilato, acrilamida, sulfona activa, amina, aminooxi, amina protegida, hidrazida, hidrazida protegida, tiol protegido, ácido carboxílico, ácido carboxílico protegido, isocianato, isotiocianato, maleimida, vinilsulfona, ditiopiridina, vinilpiridina, yodoacetamida, epóxido, glioxales, dionas, mesilatos, tosilatos y tresilato, alqueno y cetona.

I. Ejemplo de funcionalidad de adición: Facilitar las propiedades de aislamiento de un polipéptido

Un polipéptido de aminoácidos que se producen naturalmente o no naturales puede ser difícil de aislar de una muestra por varios motivos, que incluyen, pero no se limitan a, la solubilidad o características de unión del polipéptido. Por ejemplo, en la preparación de un polipéptido para uso terapéutico, un polipéptido tal puede aislarse de un sistema recombinante que se ha manipulado para producir en exceso el polipéptido. Sin embargo, debido a la solubilidad o características de unión del polipéptido, el conseguir un nivel de pureza deseado frecuentemente ha demostrado ser difícil. Los procedimientos, composiciones, técnicas y estrategias descritos en el presente documento proporcionan una solución a esta situación.

Usando los procedimientos, composiciones, técnicas y estrategias descritos en el presente documento, un experto en la materia puede producir un polipéptido de aminoácidos no naturales que contienen amina alquilada que es homólogo al polipéptido deseado, en el que el polipéptido de aminoácidos no naturales que contienen amina alquilada tiene características de aislamiento mejoradas. En una realización, un polipéptido de aminoácidos no naturales homólogo se produce biosintéticamente. En otra realización o realización adicional, el aminoácido no natural ha incorporado en su estructura uno de los aminoácidos no naturales descritos en el presente documento. En otra realización o realización adicional, el aminoácido no natural se incorpora en una posición interna o terminal y adicionalmente se incorpora de manera específica del sitio.

En una realización, el aminoácido no natural resultante, como se ha producido biosintéticamente, ya tiene las características de aislamiento mejoradas deseadas. En otras realizaciones o realizaciones adicionales, el aminoácido no natural comprende un enlace de amina a un grupo que proporciona las características de aislamiento mejoradas. En otras realizaciones o realizaciones adicionales, el aminoácido no natural se modifica adicionalmente para formar un polipéptido de aminoácidos no naturales que contienen amina alquilada, en el que la modificación proporciona un enlace de amina a un grupo que proporciona las características de aislamiento mejoradas. En algunas realizaciones, un grupo tal se liga directamente al aminoácido no natural, y en otras realizaciones, un grupo tal se liga mediante un grupo conector al aminoácido no natural. En ciertas realizaciones, un grupo tal se conecta al aminoácido no natural por una única reacción química, en otras realizaciones se requiere una serie de reacciones químicas para conectar un grupo tal al aminoácido no natural. Preferentemente, el grupo que confiere características de aislamiento mejoradas se liga de manera específica del sitio al aminoácido no natural en el polipéptido de aminoácidos no naturales y no se liga a un aminoácido que se produce naturalmente bajo las condiciones de reacción utilizadas.

En otras realizaciones o realizaciones adicionales, el polipéptido de aminoácidos no naturales resultante es homólogo a los miembros de la familia de supergenes de GH, sin embargo, los procedimientos, técnicas y composiciones descritos en esta sección pueden aplicarse a prácticamente cualquier otro polipéptido que pueda beneficiarse de las características de aislamiento mejoradas, que incluyen a modo de ejemplo solo: alfa-1 antitripsina, angiostatina, factor antihemolítico, anticuerpo, apolipoproteína, apoproteína, factor natriurético auricular, polipéptido natriurético auricular, péptido auricular, quimiocina C-X-C, T39765, NAP-2, ENA-78, gro-a, gro-b, gro-c, IP-10, GCP-2, NAP-4, SDF-1, PF4, MIG, calcitonina, ligando c-kit, citocina, quimiocina CC, proteína 1 quimioatrayente de monocitos, proteína 2 quimioatrayente de monocitos, proteína 3 quimioatrayente de monocitos, proteína 1 alfa inflamatoria de monocitos, proteína 1 beta inflamatoria de monocitos, RANTES, 1309, R83915, R91733, HCC1, T58847, D31065, T64262, CD40, ligando CD40, ligando c-kit, colágeno, factor estimulante de colonias (CSF), factor 5a del complemento, inhibidor del complemento, receptor 1 del complemento, citocina, péptido 78 activante de neutrófilos epiteliales, MIP-16, MCP-1, factor de crecimiento epidérmico (EGF), péptido activante de neutrófilos epiteliales, eritropoyetina (EPO), toxina exfoliante, factor IX, factor VII, factor VIII, factor X, factor de crecimiento de fibroblastos (FGF), fibrinógeno, fibronectina, proteína de cuatro haces helicoidales, G-CSF, glp-1, GM-CSF, glucocerebrosidasa, gonadotropina, factor de crecimiento, receptor de factores de crecimiento, grf, proteína hedgehog, hemoglobina, factor de crecimiento de hepatocitos (hGF), hirudina, hormona de crecimiento humana (hGH), albúmina de suero humano, ICAM-1, receptor de ICAM-1, LFA-1, receptor de LFA-1, insulina, factor de crecimiento similar a la insulina (IGF), IGF-I, IGF-II, interferón (IFN), IFN-alfa, IFN-beta, IFN-gamma, cualquier molécula similar a interferón o miembro de la familia de IFN, interleucina (IL), IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-

8, IL-9, IL-10, IL-11, IL-12, factor de crecimiento de queratinocitos (KGF), lactoferrina, factor inhibidor de la leucemia, luciferasa, neurturina, factor inhibidor de neutrófilos (NIF), oncostatina M, proteína osteogénica, producto oncogénico, paracitonina, hormona paratiroidea, PD-ECSF, PDGF, hormona peptídica, pleiotropina, proteína A, proteína G, pth, endotoxina A pirogénica, exotoxina B pirogénica, exotoxina C pirogénica, pyy, relaxina, renina, SCF, proteína biosintética pequeña, receptor I del complemento soluble, I-CAM 1 soluble, receptor de interleucina soluble, receptor de TNF soluble, somatomedina, somatostatina, somatotropina, estreptocinasa, superantígenos, enterotoxina estafilocócica, SEA, SEB, SEC1, SEC2, SEC3, SED, SEE, receptor de hormonas esteroideas, superóxido dismutasa, toxina del síndrome del choque tóxico, timosina alfa 1, activador tisular del plasminógeno, factor de crecimiento tumoral (TGF), factor de necrosis tumoral, factor de necrosis tumoral alfa, factor de necrosis tumoral beta, receptor del factor de necrosis tumoral (TNFR), proteína VLA-4, proteína VCAM-1, factor de crecimiento endotelial vascular (VEGF), urocinasa, 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 corticosterona.

En otras realizaciones o realizaciones adicionales, el grupo que confiere características de aislamiento mejoradas mejora la solubilidad en agua del polipéptido; en otras realizaciones, el grupo mejora las propiedades de unión del polipéptido; en otras realizaciones, el grupo proporciona nuevas propiedades de unión al polipéptido (incluyendo, a modo de ejemplo solo, un grupo de biotina o un grupo de unión a biotina). En realizaciones en las que el grupo mejora la solubilidad en agua del polipéptido, el grupo está seleccionado de los polímeros solubles en agua descritos en el presente documento, que incluyen a modo de ejemplo solo, cualquiera de los grupos de polímeros de PEG descritos en el presente documento.

J. Ejemplo de adición de funcionalidad: Detectar la presencia de un polipéptido

Un polipéptido de aminoácidos que se producen naturalmente o no naturales puede ser difícil de detectar en una muestra (incluyendo una muestra *in vivo* y una muestra *in vitro*) por varios motivos, que incluyen ausencia de un reactivo o marca que pueda unirse fácilmente al polipéptido. Los procedimientos, composiciones, técnicas y estrategias descritos en el presente documento proporcionan una solución a esta situación.

Usando los procedimientos, composiciones, técnicas y estrategias descritos en el presente documento, un experto en la materia puede producir un polipéptido de aminoácidos no naturales que contienen amina alquilada que es homólogo al polipéptido deseado, en el que el polipéptido de aminoácidos no naturales que contienen amina alquilada permite la detección del polipéptido en una muestra *in vivo* y una muestra *in vitro*. En una realización, un polipéptido de aminoácidos no naturales homólogo se produce biosintéticamente. En otra realización o realización adicional, el aminoácido no natural ha incorporado en su estructura uno de los aminoácidos no naturales descritos en el presente documento. En otra realización o realización adicional, el aminoácido no natural se incorpora en una posición interna o terminal y adicionalmente se incorpora de manera específica del sitio.

En una realización, el aminoácido no natural resultante, como se ha producido biosintéticamente, ya tiene las características de detección deseadas. En otras realizaciones o realizaciones adicionales, el aminoácido no natural comprende un enlace de amina a un grupo que proporciona las características de detección mejoradas. En otras realizaciones o realizaciones adicionales, el aminoácido no natural se modifica adicionalmente para formar un polipéptido de aminoácidos no naturales que contienen amina alquilada, en el que la modificación proporciona un enlace de amina a un grupo que proporciona las características de detección mejoradas. En algunas realizaciones, un grupo tal se liga directamente al aminoácido no natural, y en otras realizaciones, un grupo tal se liga mediante un grupo conector al aminoácido no natural. En ciertas realizaciones, un grupo tal se conecta al aminoácido no natural por una única reacción química, en otras realizaciones se requiere una serie de reacciones químicas para conectar un grupo tal al aminoácido no natural. Preferentemente, el grupo que confiere características de detección mejoradas se liga de manera específica del sitio al aminoácido no natural en el polipéptido de aminoácidos no naturales y no se liga a un aminoácido que se produce naturalmente bajo las condiciones de reacción utilizadas.

En otras realizaciones o realizaciones adicionales, el polipéptido de aminoácidos no naturales resultante es homólogo a los miembros de la familia de supergenes de GH, sin embargo, los procedimientos, técnicas y composiciones descritos en esta sección pueden aplicarse a prácticamente cualquier otro polipéptido que necesite ser detectado en una muestra *in vivo* y una muestra *in vivo*, que incluyen a modo de ejemplo solo: alfa-1 antitripsina, angiostatina, factor antihemolítico, anticuerpo, apolipoproteína, apoproteína, factor natriurético auricular, polipéptido natriurético auricular, péptido auricular, quimiocina C-X-C, T39765, NAP-2, ENA-78, gro-a, gro-b, gro-c, IP-10, GCP-2, NAP-4, SDF-1, PF4, MIG, calcitonina, ligando c-kit, citocina, quimiocina CC, proteína 1 quimioatrayente de monocitos, proteína 2 quimioatrayente de monocitos, proteína 3 quimioatrayente de monocitos, proteína 1 alfa inflamatoria de monocitos, proteína 1 beta inflamatoria de monocitos, RANTES, 1309, R83915, R91733, HCC1, T58847, D31065, T64262, CD40, ligando CD40, ligando c-kit, colágeno, factor estimulante de colonias (CSF), factor 5a del complemento, inhibidor del complemento, receptor 1 del complemento, citocina, péptido 78 activante de neutrófilos epiteliales, MIP-16, MCP-1, factor de crecimiento epidérmico (EGF), péptido activante de neutrófilos epiteliales, eritropoyetina (EPO), toxina exfoliante, factor IX, factor VII, factor VIII, factor X, factor de crecimiento de fibroblastos (FGF), fibrinógeno, fibronectina, proteína de cuatro haces helicoidales, G-CSF, glp-1, GM-CSF, glucocerebrosidasa, gonadotropina, factor de crecimiento, receptor de factores de crecimiento, grf, proteína

hedgehog, hemoglobina, factor de crecimiento de hepatocitos (hGF), hirudina, hormona de crecimiento humana (hGH), albúmina de suero humano, ICAM-1, receptor de ICAM-1, LFA-1, receptor de LFA-1, insulina, factor de crecimiento similar a la insulina (IGF), IGF-I, IGF-II, interferón (IFN), IFN-alfa, IFN-beta, IFN-gamma, cualquier molécula similar a interferón o miembro de la familia de IFN, interleucina (IL), IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL- 10, IL-11, IL-12, factor de crecimiento de queratinocitos (KGF), lactoferrina, factor inhibidor de la leucemia, luciferasa, neurturina, factor inhibidor de neutrófilos (NIF), oncostatina M, proteína osteogénica, producto oncogénico, paracitonina, hormona paratiroidea, PD-ECSF, PDGF, hormona peptídica, pleiotropina, proteína A, proteína G, pth, endotoxina A pirogénica, exotoxina B pirogénica, exotoxina C pirogénica, pyy, relaxina, renina, SCF, proteína biosintética pequeña, receptor I del complemento soluble, I-CAM 1 soluble, receptor de interleucina soluble, receptor de TNF soluble, somatomedina, somatostatina, somatotropina, estreptocinasa, superantígenos, enterotoxina estafilocócica, SEA, SEB, SEC1, SEC2, SEC3, SED, SEE, receptor de hormonas esteroideas, superóxido dismutasa, toxina del síndrome del choque tóxico, timosina alfa 1, activador tisular del plasminógeno, factor de crecimiento tumoral (TGF), factor de necrosis tumoral, factor de necrosis tumoral alfa, factor de necrosis tumoral beta, receptor del factor de necrosis tumoral (TNFR), proteína VLA-4, proteína VCAM-1, factor de crecimiento endotelial vascular (VEGF), urocinasa, 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 corticosterona.

En otras realizaciones o realizaciones adicionales, el grupo que confiere características de detección mejoradas está seleccionado del grupo que consiste en una marca; un colorante; una marca de afinidad; una marca de fotoafinidad; una marca de espín; un fluoróforo; un resto radiactivo; un resto que incorpora un átomo pesado; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un cromóforo; un agente de transferencia de energía; una marca detectable; y cualquier combinación de los anteriores.

K. Ejemplo de adición de funcionalidad: Mejora de las propiedades terapéutico de un polipéptido

Un polipéptido de aminoácidos que se producen naturalmente o no naturales podrá proporcionar un cierto beneficio terapéutico a un paciente con un trastorno, enfermedad o afección particular. Un beneficio terapéutico tal dependerá de varios factores, que incluyen a modo de ejemplo solo: el perfil de seguridad del polipéptido, y la farmacocinética, farmacológica y/o farmacodinámica del polipéptido (por ejemplo, solubilidad en agua, biodisponibilidad, semivida en suero, semivida terapéutica, inmunogenicidad, actividad biológica o tiempo en circulación). Además, puede ser ventajoso proporcionar funcionalidad adicional al polipéptido, tal como un compuesto citotóxico o fármaco unido, o puede desearse unir polipéptidos adicionales para formar los homo- y heteromultímeros descritos en el presente documento. Tales modificaciones no destruyen preferentemente la actividad y/o estructura terciaria del polipéptido original. Los procedimientos, composiciones, técnicas y estrategias descritos en el presente documento proporcionan soluciones a estas cuestiones.

Usando los procedimientos, composiciones, técnicas y estrategias descritos en el presente documento, un experto en la materia puede producir un polipéptido de aminoácidos no naturales que contienen amina alquilada que es homólogo al polipéptido deseado, en el que el polipéptido de aminoácidos no naturales que contienen amina alquilada tiene características terapéuticas mejoradas. En una realización, un polipéptido de aminoácidos no naturales homólogo se produce biosintéticamente. En otra realización o realización adicional, el aminoácido no natural ha incorporado en su estructura uno de los aminoácidos no naturales descritos en el presente documento. En otra realización o realización adicional, el aminoácido no natural se incorpora en una posición interna o terminal y adicionalmente se incorpora de manera específica del sitio.

En una realización, el aminoácido no natural resultante, como se ha producido biosintéticamente, ya tiene las características terapéuticas mejoradas deseadas. En otras realizaciones o realizaciones adicionales, el aminoácido no natural comprende un enlace de amina a un grupo que proporciona las características terapéuticas mejoradas. En otras realizaciones o realizaciones adicionales, el aminoácido no natural se modifica adicionalmente para formar un polipéptido de aminoácidos no naturales que contienen amina alquilada, en el que la modificación proporciona un enlace de amina a un grupo que proporciona las características terapéuticas mejoradas. En algunas realizaciones, un grupo tal se liga directamente al aminoácido no natural, y en otras realizaciones, un grupo tal se liga mediante un grupo conector al aminoácido no natural. En ciertas realizaciones, un grupo tal se conecta al aminoácido no natural por una única reacción química, en otras realizaciones se requiere una serie de reacciones químicas para conectar un grupo tal al aminoácido no natural. Preferentemente, el grupo que confiere características terapéuticas mejoradas se liga de manera específica del sitio al aminoácido no natural en el polipéptido de aminoácidos no naturales y no se liga a un aminoácido que se produce naturalmente bajo las condiciones de reacción utilizadas.

En otras realizaciones o realizaciones adicionales, el polipéptido de aminoácidos no naturales resultante es homólogo a los miembros de la familia de supergenes de GH, sin embargo, los procedimientos, técnicas y composiciones descritos en esta sección pueden aplicarse a prácticamente cualquier otro polipéptido que pueda beneficiarse de características terapéuticas mejoradas, que incluyen a modo de ejemplo solo: alfa-1 antitripsina, angiostatina, factor antihemolítico, anticuerpo, apolipoproteína, apoproteína, factor natriurético auricular, polipéptido natriurético auricular, péptido auricular, quimiocina C-X-C, T39765, NAP-2, ENA-78, gro-a, gro-b, gro-c, IP-10, GCP-

2, NAP-4, SDF-1, PF4, MIG, calcitonina, ligando c-kit, citocina, quimiocina CC, proteína 1 quimioatrayente de monocitos, proteína 2 quimioatrayente de monocitos, proteína 3 quimioatrayente de monocitos, proteína 1 alfa inflamatoria de monocitos, proteína 1 beta inflamatoria de monocitos, RANTES, 1309, R83915, R91733, HCCI, T58847, D31065, T64262, CD40, ligando CD40, ligando c-kit, colágeno, factor estimulante de colonias (CSF), factor 5a del complemento, inhibidor del complemento, receptor 1 del complemento, citocina, péptido 78 activante de neutrófilos epiteliales, MIP-16, MCP-1, factor de crecimiento epidérmico (EGF), péptido activante de neutrófilos epiteliales, eritropoyetina (EPO), toxina exfoliante, factor IX, factor VII, factor VIII, factor X, factor de crecimiento de fibroblastos (FGF), fibrinógeno, fibronectina, proteína de cuatro haces helicoidales, G-CSF, glp-1, GM-CSF, glucocerebrosidasa, gonadotropina, factor de crecimiento, receptor de factores de crecimiento, grf, proteína hedgehog, hemoglobina, factor de crecimiento de hepatocitos (hGF), hirudina, hormona de crecimiento humana (hGH), albúmina de suero humano, ICAM-1, receptor de ICAM-1, LFA-1, receptor de LFA-1, insulina, factor de crecimiento similar a la insulina (IGF), IGF-I, IGF-II, interferón (IFN), IFN-alfa, IFN-beta, IFN-gamma, cualquier molécula similar a interferón o miembro de la familia de IFN, interleucina (IL), IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, factor de crecimiento de queratinocitos (KGF), lactoferrina, factor inhibidor de la leucemia, luciferasa, neurturina, factor inhibidor de neutrófilos (NIF), oncostatina M, proteína osteogénica, producto oncogénico, paracitonina, hormona paratiroidea, PD-ECSF, PDGF, hormona peptídica, pleiotropina, proteína A, proteína G, pth, endotoxina A pirogénica, exotoxina B pirogénica, exotoxina C pirogénica, pyy, relaxina, renina, SCF, proteína biosintética pequeña, receptor I del complemento soluble, I-CAM 1 soluble, receptor de interleucina soluble, receptor de TNF soluble, somatomedina, somatostatina, somatotropina, estreptocinasa, superantígenos, enterotoxina estafilocócica, SEA, SEB, SEC1, SEC2, SEC3, SED, SEE, receptor de hormonas esteroideas, superóxido dismutasa, toxina del síndrome del choque tóxico, timosina alfa 1, activador tisular del plasminógeno, factor de crecimiento tumoral (TGF), factor de necrosis tumoral, factor de necrosis tumoral alfa, factor de necrosis tumoral beta, receptor del factor de necrosis tumoral (TNFR), proteína VLA-4, proteína VCAM-1, factor de crecimiento endotelial vascular (VEGF), urocinasa, 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 corticosterona.

En otras realizaciones o realizaciones adicionales, el grupo que confiere características terapéuticas mejoradas mejora la solubilidad en agua del polipéptido; en otras realizaciones, el grupo mejora las propiedades de unión del polipéptido; en otras realizaciones, el grupo proporciona nuevas propiedades de unión al polipéptido (incluyendo, a modo de ejemplo solo, un grupo de biotina o un grupo de unión a biotina). En realizaciones en las que el grupo mejora la solubilidad en agua del polipéptido, el grupo está seleccionado de los polímeros solubles en agua descritos en el presente documento, que incluyen a modo de ejemplo solo, los grupos de polímeros de PEG. En otras realizaciones o realizaciones adicionales, el grupo es un compuesto citotóxico, mientras que en otras realizaciones el grupo es un fármaco. En otras realizaciones, el fármaco o compuesto citotóxico ligado puede escindirse del polipéptido de aminoácidos no naturales de manera que se administre el fármaco o compuesto citotóxico a una localización terapéutica deseada. En otras realizaciones, el grupo es un segundo polipéptido, que incluye a modo de ejemplo, un polipéptido de aminoácidos no naturales que contienen amina alquilada, que incluye adicionalmente a modo de ejemplo, un polipéptido que tiene la misma estructura de aminoácido que el primer polipéptido de aminoácidos no naturales.

En otras realizaciones o realizaciones adicionales, el polipéptido de aminoácidos no naturales que contienen amina alquilada es un polipéptido de aminoácidos no naturales que contienen amina alquilada modificado. En otras realizaciones o realizaciones adicionales, el polipéptido de aminoácidos no naturales que contienen amina alquilada aumenta la biodisponibilidad del polipéptido con respecto al polipéptido de aminoácidos que se producen naturalmente homólogo. En otras realizaciones o realizaciones adicionales, el polipéptido de aminoácidos no naturales que contienen amina alquilada aumenta el perfil de seguridad del polipéptido con respecto al polipéptido de aminoácidos que se producen naturalmente homólogo. En otras realizaciones o realizaciones adicionales, el polipéptido de aminoácidos no naturales que contienen amina alquilada aumenta la solubilidad en agua del polipéptido con respecto al polipéptido de aminoácidos que se producen naturalmente homólogo. En otras realizaciones o realizaciones adicionales, el polipéptido de aminoácidos no naturales que contienen amina alquilada aumenta la semivida terapéutica del polipéptido con respecto al polipéptido de aminoácidos que se producen naturalmente homólogo. En otras realizaciones o realizaciones adicionales, el polipéptido de aminoácidos no naturales que contienen amina alquilada aumenta la semivida en suero del polipéptido con respecto al polipéptido de aminoácidos que se producen naturalmente homólogo. En otras realizaciones o realizaciones adicionales, el polipéptido de aminoácidos no naturales que contienen amina alquilada prolonga el tiempo en circulación del polipéptido con respecto al polipéptido de aminoácidos que se producen naturalmente homólogo. En otras realizaciones o realizaciones adicionales, el polipéptido de aminoácidos no naturales que contienen amina alquilada modula la actividad biológica del polipéptido con respecto al polipéptido de aminoácidos que se producen naturalmente homólogo. En otras realizaciones o realizaciones adicionales, el polipéptido de aminoácidos no naturales que contienen amina alquilada modula la inmunogenicidad del polipéptido con respecto al polipéptido de aminoácidos que se producen naturalmente homólogo.

XI. Usos terapéuticos de polipéptidos modificados

Por comodidad, los polipéptidos no naturales “modificados o no modificados” descritos en esta sección se han

descrito genéricamente y/o con ejemplos específicos. Sin embargo, los polipéptidos no naturales "modificados o no modificados" descritos en esta sección no deben limitarse a solo las descripciones genéricas o ejemplo específico proporcionado en esta sección, sino que los polipéptidos no naturales "modificados o no modificados" descritos en esta sección se aplican igualmente de bien a todos los polipéptidos no naturales "modificados o no modificados" que comprenden al menos un aminoácido que se encuentra dentro del alcance de las fórmulas (A)-(E) y (I)-(XVI), que incluye cualquier sub-fórmula o compuesto específico que se encuentra dentro del alcance de las fórmulas (A)-(E) y (I)-(XVI) que se describen en la memoria descriptiva, reivindicaciones y figuras en el presente documento.

Los polipéptidos de aminoácidos no naturales modificados o no modificados descritos en el presente documento, que incluyen homo- y hetero-multímeros de los mismos, encuentran múltiples usos, que incluyen, pero no se limitan a: usos terapéuticos, de diagnóstico, basados en ensayos, industriales, cosméticos, de biología vegetal, medioambientales, de producción de energía y/o militares. Como una ilustración no limitante, se proporcionan los siguiente usos terapéuticos de polipéptidos de aminoácidos no naturales modificados o no modificados.

Los polipéptidos de aminoácidos no naturales modificados o no modificados descritos en el presente documento son útiles para tratar una amplia variedad de trastornos, condiciones o enfermedades. La administración de los productos de polipéptidos de aminoácidos no naturales modificados o no modificados descritos en el presente documento produce cualquiera de las actividades demostradas por preparaciones de polipéptidos comercialmente disponibles en seres humanos. Las cantidades promedio del producto de polipéptidos de aminoácidos no naturales modificados o no modificados pueden variarse y en particular deben basarse en las recomendaciones y prescripción de un médico cualificado. La cantidad exacta del polipéptido de aminoácidos no naturales modificado o no modificado es una cuestión de preferencia sujeta a factores tales como el tipo exacto de afección que está tratándose, la afección del paciente que está tratándose, además de los otros componentes en la composición. La cantidad que va a administrarse puede determinarse fácilmente por un experto en la materia basándose en la terapia con el polipéptido de aminoácidos no naturales modificado o no modificado.

A. Administración y composiciones farmacéuticas

Los polipéptidos de aminoácidos no naturales "modificados o no modificados" descritos en el presente documento, que incluye homo- y hetero-multímeros de los mismos, encuentran múltiples usos, que incluyen, pero no se limitan a: usos terapéuticos, de diagnóstico, basados en ensayos, industriales, cosméticos, de biología vegetal, medioambientales, de producción de energía y/o militares. Como una ilustración no limitante, se proporcionan los siguientes usos terapéuticos de polipéptidos de aminoácidos no naturales "modificados o no modificados".

Los polipéptidos de aminoácidos no naturales "modificados o no modificados" descritos en el presente documento son útiles para tratar una amplia variedad de trastornos. La administración de los productos de polipéptidos de aminoácidos no naturales "modificados o no modificados" descritos en el presente documento produce cualquiera de las actividades demostradas por preparaciones de polipéptidos comercialmente disponibles en seres humanos. Las cantidades promedio del producto de polipéptidos de aminoácidos no naturales modificados o no modificados pueden variarse y en particular deben basarse en las recomendaciones y prescripción de un médico cualificado. La cantidad exacta del polipéptido de aminoácidos no naturales "modificado o no modificado" es una cuestión de preferencia sujeta a factores tales como el tipo exacto de afección que está tratándose, la condición del paciente que está tratándose, además de los otros componentes en la composición. La cantidad que va a administrarse puede determinarse fácilmente por un experto en la materia basándose en la terapia con el polipéptido de aminoácidos no naturales "modificado o no modificado".

Los polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describen en el presente documento (que incluyen, pero no se limitan a, sintetasas, proteínas que comprenden uno o más aminoácidos no naturales, etc.) se emplean opcionalmente para usos terapéuticos, que incluyen, pero no se limitan a, en combinación con un vehículo farmacéutico adecuado. Tales composiciones, por ejemplo, comprenden una cantidad terapéuticamente eficaz de los polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describe en el presente documento, y un vehículo o excipiente farmacéuticamente aceptable. Un vehículo o excipiente tal incluye solución salina, solución salina tamponada, dextrosa, agua, glicerol, etanol, y/o combinaciones de los mismos. La formulación se prepara para adaptarse al modo de administración. En general, los procedimientos de administración de proteínas son muy conocidos en la técnica y pueden aplicarse a la administración de los polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describe en el presente documento.

Composiciones terapéuticas que comprenden uno o más de los polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describen en el presente documento se prueban opcionalmente en uno o más modelos animales *in vitro* y/o *in vivo* apropiados de enfermedad, para confirmar la eficacia, metabolismo del tejido, y para estimar dosificaciones, según procedimientos muy conocidos en la técnica. En particular, las dosificaciones pueden determinarse inicialmente por actividad, estabilidad u otras medidas adecuadas de homólogos de aminoácidos no naturales con respecto a naturales (que incluyen, pero no se limitan a, comparación de un polipéptido modificado para incluir uno o más aminoácidos no naturales con un polipéptido de aminoácidos naturales), es decir, en un ensayo relevante.

La administración es por cualquiera de las vías normalmente usadas para introducir una molécula en contacto íntimo con sangre o células de tejido. Los polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describen en el presente documento, se administran en cualquier manera adecuada, opcionalmente con uno o más vehículos farmacéuticamente aceptables. Están disponibles procedimientos adecuados de administración de los polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describen en el presente documento, a un paciente, y, aunque puede usarse más de una vía para administrar una composición particular, una vía particular puede proporcionar frecuentemente una acción o reacción más inmediata y más eficaz que otra vía.

Los vehículos farmacéuticamente aceptables se determinan en parte por la composición particular que se administra, además de por el procedimiento particular usado para administrar la composición. Por consiguiente, hay una amplia variedad de formulaciones adecuadas de composiciones farmacéuticas descritas en el presente documento.

Los polipéptidos de aminoácidos no naturales descritos en el presente documento y las composiciones que comprenden tales polipéptidos pueden administrarse por cualquier vía convencional adecuada para proteínas o péptidos, que incluyen parenteralmente, por ejemplo, inyecciones que incluyen subcutáneamente o intravenosamente o cualquier otra forma de inyecciones o infusiones. Las composiciones de polipéptidos farmacéuticas que comprenden polipéptidos de aminoácidos no naturales descritos en el presente documento pueden administrarse por varias vías que incluyen medios orales, intravenosos, intraperitoneales, intramusculares, transdérmicos, subcutáneos, tópicos, sublinguales o rectales. Las composiciones que comprenden polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describen en el presente documento, también pueden administrarse mediante liposomas. Tales vías de administración y formulaciones apropiadas son generalmente conocidas para aquellos expertos en la materia. Los polipéptidos de aminoácidos no naturales descritos en el presente documento pueden usarse solos o en combinación con otros componentes adecuados, que incluyen, pero no se limitan a, un vehículo farmacéutico.

Los polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describen en el presente documento, solos o en combinación con otros componentes adecuados, también pueden prepararse en formulaciones en aerosol (es decir, pueden “nebulizarse”) para administrarse mediante inhalación. Las formulaciones en aerosol pueden colocarse en propulsores aceptables presurizados, tales como diclorodifluorometano, propano, nitrógeno y similares.

Formulaciones adecuadas para administración parenteral, tales como, por ejemplo, por vías intraarticular (en las articulaciones), intravenosa, intramuscular, intradérmica, intraperitoneal y subcutánea, incluyen disoluciones para inyección estéril isotónica, acuosas y no acuosas, que pueden contener antioxidantes, tampones, bacteriostáticos y solutos que convierten la formulación en isotónica con la sangre del receptor previsto, y suspensiones estériles acuosas y no acuosas que pueden incluir agentes de suspensión, solubilizantes, espesantes, estabilizadores y conservantes. Las formulaciones de ácido nucleico envasadas pueden presentarse en recipientes cerrados de dosis unitaria o multi-dosis, tales como ampollas y viales.

La administración parenteral y la administración intravenosa son procedimientos preferidos de administración. En particular, las vías de administración ya en uso para terapéuticos homólogos de aminoácidos naturales (que incluyen, pero no se limitan a, aquellos normalmente usados para EPO, IFN, GH, G-CSF, GM-CSF, IFN, interleucinas, anticuerpos, y/o cualquier otra proteína farmacéuticamente administrada), junto con formulaciones actualmente en uso, proporcionan vías preferidas de administración y formulación para los polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describen en el presente documento.

La dosis administrada a un paciente, en el contexto de las composiciones y procedimientos descritos en el presente documento, es suficiente para tener una respuesta terapéutica beneficiosa en el paciente con el tiempo. La dosis se determina por la eficacia de la formulación particular, y la actividad, estabilidad o semivida en suero de los polipéptidos de aminoácidos no naturales, modificados o no modificados, empleados y la afección del paciente, además del peso corporal o área superficial del paciente que va a tratarse. El tamaño de la dosis también se determina por la existencia, naturaleza y grado de cualquier efecto secundario adverso que acompañe la administración de una formulación particular, o similares, en un paciente particular.

En la determinación de la cantidad eficaz de formulación que va a administrarse en el tratamiento o profilaxis de la enfermedad (que incluye cánceres, enfermedades hereditarias, diabetes, SIDA, o similares), el médico evalúa los niveles en plasma circulantes, toxicidades de la formulación, progresión de la enfermedad, y/o si es relevante, la producción de anticuerpos anti-polipéptido de aminoácidos no naturales.

La dosis administrada, por ejemplo, a un paciente de 70 kilogramos, está normalmente en el intervalo equivalente a dosificaciones de proteínas terapéuticas actualmente usadas, ajustadas para la actividad o semivida alterada en suero de la composición relevante. Las formulaciones farmacéuticas descritas en el presente documento pueden complementar condiciones de tratamiento por cualquier terapia convencional conocida, que incluye administración de anticuerpos, administración de vacunas, administración de agentes citotóxicos, polipéptidos de aminoácidos

naturales, ácidos nucleicos, análogos de nucleótidos, modificadores de la respuesta biológica, y similares.

Para administración, las formulaciones farmacéuticas descritas en el presente documento se administran a una tasa determinada por la DL50 o DE50 de la formulación relevante, y/u observación de cualquier efecto secundario de los polipéptidos de aminoácidos no naturales, modificados o no modificados, a diversas concentraciones, que incluyen como se aplica a la masa y salud general del paciente. La administración puede llevarse a cabo mediante dosis únicas o divididas.

Si un paciente que recibe infusión de una formulación desarrolla fiebre, escalofríos o dolores musculares, recibe la dosis apropiada de aspirina, ibuprofeno, acetaminofeno u otro fármaco que controle el dolor/fiebre. Pacientes que experimentan reacciones a la infusión tales como fiebre, dolores musculares y escalofríos son previamente medicados 30 minutos antes de las futuras infusiones con tanto aspirina, acetaminofeno, o, que incluye, pero no se limita a, difenhidramina. Se usa meperidina para escalofríos y dolores musculares más graves que no responden rápidamente a antipiréticos y antihistamínicos. La infusión de células se ralentiza o se interrumpe dependiendo de la gravedad de la reacción.

Los polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describen en el presente documento, pueden administrarse directamente a un sujeto mamífero. La administración es por cualquiera de las vías normalmente usadas para introducir un polipéptido a un sujeto. Los polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describen en el presente documento, incluyen aquellos adecuados para administración oral, rectal, tópica, inhalación (incluyendo mediante un aerosol), bucal (incluyendo sub-lingual), vaginal, parenteral (incluyendo subcutánea, intramuscular, intradérmica, intraarticular, intrapleural, intraperitoneal, intracerebral, intraarterial o intravenosa), tópica (es decir, tanto la piel como las superficies de la mucosa, que incluyen superficies de las vías respiratorias) y transdérmica, aunque la vía más adecuada en cualquier caso dado dependerá de la naturaleza y gravedad de la afección que está tratándose. La administración puede ser tanto local como sistémica. Las formulaciones pueden presentarse en recipientes cerrados de dosis unitaria o multi-dosis, tales como ampollas y viales. Los polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describen en el presente documento, pueden prepararse en una mezcla en una forma inyectable de dosificación unitaria (incluyendo disolución, suspensión o emulsión) con un vehículo farmacéuticamente aceptable. Los polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describen en el presente documento, también pueden administrarse por infusión continua (usando, que incluye minibombas tales como bombas osmóticas), bolo único o formulaciones de liberación prolongada de liberación lenta.

Formulaciones adecuadas para administración incluyen disoluciones acuosas y no acuosas, disoluciones estériles isotónicas, que pueden contener antioxidantes, tampones, bacteriostáticos y solutos que convierten la formulación en isotónica, y suspensiones estériles acuosas y no acuosas que pueden incluir agentes de suspensión, solubilizantes, espesantes, estabilizadores y conservantes. Las disoluciones y suspensiones pueden prepararse a partir de polvos estériles, gránulos y comprimidos del tipo previamente descrito.

El secado por congelación es una técnica comúnmente empleada para presentar proteínas que sirve para eliminar agua de la preparación de proteínas de interés. El secado por congelación, o liofilización, es un procedimiento por el cual el material que va a secarse se congela primero y a continuación el hielo o disolvente congelado se elimina por sublimación en un entorno a vacío. Un excipiente puede incluirse en pre-formulaciones liofilizadas para potenciar la estabilidad durante el procedimiento de secado por congelación y/o para mejorar la estabilidad del producto liofilizado tras el almacenamiento. Pikal, M. Biopharm. 3(9)26-30 (1990) y Arakawa y col. Pharm. Res. 8(3):285-291 (1991).

El secado por pulverización de productos farmacéuticos también es conocido para aquellos expertos habituales en la materia. Por ejemplo, véase Broadhead, J. y col., “The Spray Drying of Pharmaceuticals” en Drug Dev. Ind. Pharm, 18 (11 & 12), 1169-1206 (1992). Además de productos farmacéuticos de molécula pequeña, una variedad de materiales biológicos se han secado por pulverización y estos incluyen: enzimas, sueros, plasma, microorganismos y levaduras. El secado por pulverización es una técnica útil debido a que puede convertir una preparación farmacéutica líquida en un polvo fino, sin polvo o aglomerado en un procedimiento de una etapa. La técnica básica comprende las cuatro siguientes etapas: a) atomización de la disolución de alimentación en un espray; b) contacto espray-aire; c) secado del espray; y d) separación del producto secado del aire de secado. Las patentes de EE.UU. nº 6.235.710 y 6.001.800 describen la preparación de eritropoyetina recombinante secando por pulverización.

Las composiciones farmacéuticas descritas en el presente documento pueden comprender un vehículo, excipiente o estabilizador farmacéuticamente aceptable. Los vehículos farmacéuticamente aceptables se determinan en parte por la composición particular que se administra, además de por el procedimiento particular usado para administrar la composición. Por consiguiente, hay una amplia variedad de formulaciones adecuadas de composiciones farmacéuticas (incluyendo vehículos, excipientes o estabilizadores farmacéuticamente aceptables opcionales) para los polipéptidos de aminoácidos no naturales, modificados o no modificados, descritos en el presente documento, (véanse, por ejemplo, en Remington: The Science and Practice of Pharmacy, Nineteenth Ed (Easton, Pa.: Mack Publishing Company, 1995); Hoover, John E., Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, Pennsilvania 1975; Liberman, H.A. y Lachman, L., Eds., Pharmaceutical Dosage Forms, Marcel Decker, New York,

N.Y., 1980; and Pharmaceutical Dosage Forms and Drug Delivery Systems, Seventh Ed. (Lippincott Williams & Wilkins, 1999)). Vehículos adecuados incluyen tampones que contienen succinato, fosfato, borato, HEPES, citrato, imidazol, acetato, bicarbonato, y otros ácidos orgánicos; antioxidantes que incluyen ácido ascórbico; polipéptidos de bajo peso molecular que incluyen aquellos de menos de aproximadamente 10 residuos; proteínas que incluyen albúmina de suero, gelatina o inmunoglobulinas; polímeros hidrófilos que incluyen polivinilpirrolidona; aminoácidos que incluyen glicina, glutamina, asparagina, arginina, histidina o derivados de histidina, metionina, glutamato o lisina; monosacáridos, disacáridos y otros hidratos de carbono, que incluyen trehalosa, sacarosa, glucosa, manosa o dextrinas; agentes quelantes que incluyen EDTA y edentato disódico; iones de metales divalentes que incluyen cinc, cobalto o cobre; alcoholes de azúcar que incluyen manitol o sorbitol; contraiones formadores de sal que incluyen sodio; y/o tensioactivos no iónicos, que incluyen Tween™ (incluyendo Tween 80 (polisorbato 80) y Tween 20 (polisorbato 20), Pluronics™ y otros ácidos plurónicos, que incluyen y otros ácidos plurónicos, que incluyen el ácido plurónico F68 (Poloxamer 188), o PEG. Tensioactivos adecuados incluyen, por ejemplo, poliéteres basados en poli(óxido de etileno)-poli(óxido de propileno)-poli(óxido de etileno), es decir, (PEO-PPO-PEO), o poli(óxido de propileno)-poli(óxido de etileno)-poli(óxido de propileno), es decir, (PPO-PEO-PPO), o una combinación de los mismos. PEO-PPO-PEO y PPO-PEO-PPO están comercialmente disponibles bajo los nombres comerciales Pluronics™, R-Pluronics™, Tetronics™ y R-Tetronics™ (BASF Wyandotte Corp., Wyandotte, Mich.) y se describen adicionalmente en la patente de EE.UU. nº 4.820.352. Otros polímeros de bloque de etileno/polipropileno pueden ser tensioactivos adecuados. Puede usarse un tensioactivo o una combinación de tensioactivos para estabilizar polipéptidos de aminoácidos no naturales PEGilados contra uno o más estreses que incluyen, pero no se limitan a, el estrés que resulta de la agitación. Algunos de los anteriores pueden denominarse “agentes de carga”. Algunos también pueden denominarse “modificadores de la tonicidad”. También pueden aplicarse conservantes antimicrobianos para la estabilidad del producto y eficacia antimicrobiana; conservantes adecuados incluyen, pero no se limitan a, alcohol bencílico, cloruro de benzalconio, metacresol, metil/propilparabeno, cresol y fenol, o una combinación de los mismos.

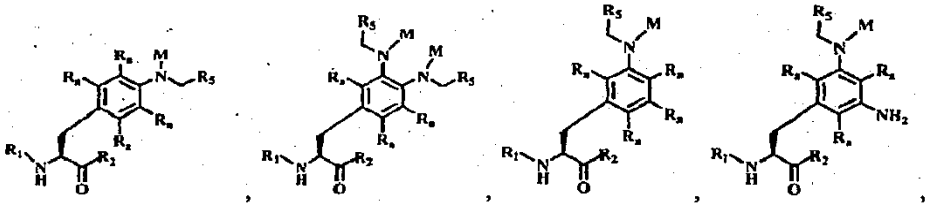
Los polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describen en el presente documento, que incluyen aquellos ligados a polímeros solubles en agua tales como PEG, también pueden administrarse por o como parte de sistemas de liberación sostenida. Las composiciones de liberación sostenida incluyen, que incluyen, pero no se limitan a, matrices de polímeros semi-permeables en forma de artículos moldeados, que incluyen películas, o microcápsulas. Las matrices de liberación sostenida incluyen desde materiales biocompatibles tales como poli(2-hidroximetacrilato de etilo) (Langer y col., J. Biomed. Mater. Res., 15: 267-277 (1981); Langer, Chem. Tech., 12: 98-105 (1982), etileno-acetato de vinilo (Langer y col., arriba) o ácido poli-D-(-)-3-hidroxibutírico (documento EP 133.988), polilactidas (ácido poliláctico) (patente de EE.UU. nº 3.773.919; EP 58.481), poliglicolida (polímero de ácido glicólico), polilactida-co-glicolida (copolímeros de ácido láctico y ácido glicólico), polianhídridos, copolímeros de ácido L-glutámico y L-glutamato de gamma-etilo (U. Sidman y col., Biopolymers 22, 547-556 (1983), poli(orto)ésteres, polipéptidos, ácido hialurónico, colágeno, sulfato de condroitina, ácidos carboxílicos, ácidos grasos, fosfolípidos, polisacáridos, ácidos nucleicos, poliaminoácidos, aminoácidos tales como fenilalanina, tirosina, isoleucina, polinucleótidos, polivinilpropileno, polivinilpirrolidona y silicona. Las composiciones de liberación sostenida también incluyen un compuesto liposómicamente atrapado. Liposomas que contienen el compuesto se preparan mediante procedimientos en sí conocidos: documento DE 3.218.121; Eppstein y col., Proc. Natl. Acad. Sci. U.S.A., 82: 3688-3692 (1985); Hwang y col., Proc. Natl. Acad. Sci. U.S.A., 77: 4030-4034 (1980); documentos EP 52.322; EP 36.676; EP 143.949; solicitud de patente japonesa 83-1118008; patentes de EE.UU. nº 4.485.045, 4.619.794, 5.021.234 y 4.544.545; y EP 102.324.

Los polipéptidos liposómicamente atrapados pueden prepararse mediante procedimientos descritos en, por ejemplo, el documento DE 3.218.121; Eppstein y col., Proc. Natl. Acad. Sci. U.S.A., 82: 3688-3692 (1985); Hwang y col., Proc. Natl. Acad. Sci. U.S.A., 77: 4030-4034 (1980); documentos EP 52.322; EP 36.676; EP 143.949; solicitud de patente japonesa 83-118008; patentes de EE.UU. nº 4.485.045, 4.619.794, 5.021.234 y 4.544.545; y EP 102.324. La composición y tamaño de los liposomas son muy conocidos o pueden determinarse fácilmente empíricamente por un experto en la materia. Algunos ejemplos de liposomas como se describen en, por ejemplo, Park JW, y col., Proc. Natl. Acad. Sci. USA 92:1327-1331 (1995); Lasic D y Papahadjopoulos D (eds): MEDICAL APPLICATIONS OF LIPOSOMES (1998); Drummond DC, y col., Liposomal drug delivery systems for cancer therapy, en Teicher B (ed): CANCER DRUG DISCOVERY AND DEVELOPMENT (2002); Park JW y col., Clin. Cancer Res. 8:1172-1181 (2002); Nielsen UB y col., Biochim. Biophys. Acta 1591(1-3):109-118 (2002); Mamot C y col., Cancer Res. 63: 3154-3161 (2003).

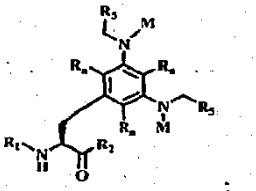
La dosis administrada a un paciente en el contexto de las composiciones, formulaciones y procedimientos descritos en el presente documento debe ser suficiente para producir una respuesta beneficiosa en el sujeto con el tiempo. Generalmente, la cantidad farmacéuticamente eficaz total de los polipéptidos de aminoácidos no naturales, modificados o no modificados, como se describe en el presente documento, administrada parenteralmente por dosis está en el intervalo de aproximadamente 0,01 µg/kg/día a aproximadamente 100 µg/kg, o aproximadamente 0,05 mg/kg a aproximadamente 1 mg/kg, de peso corporal del paciente, aunque esto está sujeto a discreción terapéutica. La frecuencia de dosificación también está sujeta a discreción terapéutica, y puede ser más frecuente o menos frecuente que los productos comercialmente disponibles autorizados para su uso en seres humanos. Generalmente, un conjugado de polímero:polipéptido, que incluye a modo de ejemplo solo, un polipéptido PEGilado, como se describe en el presente documento, puede administrarse por cualquiera de las vías de administración descritas

anteriormente. La invención se define por las reivindicaciones.

En el presente documento se describe un procedimiento de tratamiento de un trastorno, afección o enfermedad que comprende administrar una cantidad terapéuticamente eficaz de un polipéptido que comprende al menos un aminoácido no natural seleccionado del grupo que consiste en:



y



en la que

cada R_a está seleccionado independientemente del grupo que consiste en H, halógeno, alquilo, - -CN, alquilo sustituido, -N(R')$_2$, -C(O)$_k$R', -C(O)N(R')$_2$, -OR' y -S(O)$_k$R' en la que k es 1, 2 ó 3;

M es H o -CH_2R_5; o el resto M-N-C(R_5) puede formar una estructura de anillo de 4 a 7 miembros;

R_1 es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido;

R_2 es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;

R_5 es alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alquinilo, alquinilo sustituido, alcoxi, alcoxi sustituido, alquilalcoxi, alquilalcoxi sustituido, poli(óxido de alquileno), poli(óxido de alquileno) sustituido, cicloalquilo, cicloalquilo sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido, -C(O)R'', -C(O)OR'', -C(O)N(R'')$_2$, -C(O)NHCH(R'')$_2$, -(alquilen o alquilen sustituido)-N(R'')$_2$, -(alquenilen o alquenilen sustituido)-N(R'')$_2$, -(alquilen o alquilen sustituido)-(arilo o arilo sustituido), -(alquenilen o alquenilen sustituido)-(arilo o arilo sustituido), -(alquilen o alquilen sustituido)-ON(R'')$_2$, -(alquilen o alquilen sustituido)-C(O)SR'', -(alquilen o alquilen sustituido)-S-S-(arilo o arilo sustituido), en las que cada R'' es independientemente hidrógeno, alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alcoxi, alcoxi sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido o -C(O)OR';

o R_5 es L-X, en la que X se selecciona del grupo que consiste en una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable; biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del

complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos; y L es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno, alquileno sustituido, alquenileno, alquenileno sustituido, -O-, -O-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-O-, -C(O)-, -C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)-, -C(O)N(R')-, -C(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)N(R'), -OC(O)N(R')-, -OC(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-OC(O)N(R')-, -N(R')C(O)-, -NR'C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-NR'C(O)-, -S-, -S-(alquilen o alquilen sustituido)-, $-S(O)_k-$ en la que k es 1, 2 ó 3, $-S(O)_k$(alquilen o alquilen sustituido)-, -C(S)-, -C(S)-(alquilen o alquilen sustituido)-, -CSN(R')-, -CSN(R')-(alquilen o alquilen sustituido)-, -N(R')CO-(alquilen o alquilen sustituido)-, -N(R')C(O)O-, -(alquilen o alquilen sustituido)-O-N=CR'-, -(alquilen o alquilen sustituido)-C(O)NR'-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-$S(O)_k$-(alquilen o alquilen sustituido)-S-, -(alquilen o alquilen sustituido)-S-S-, $-S(O)_kN(R')-$, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, $-N(R')S(O)_kN(R')-$, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, $-C(R')_2-N=N-$ y $-C(R')_2-N(R')-N(R')-$;

o R_5 y cualquier R_a forman opcionalmente un cicloalquilo o un heterocicloalquilo; y

cada R' es independientemente H, alquilo o alquilo sustituido con la condición de que si R_1 es H, entonces R_2 no es OH, o si

R_2 es OH, entonces R_1 no es H.

En el presente documento se describen en otras realizaciones un procedimiento de tratamiento de un trastorno, afección o enfermedad, que comprende además administrar un vehículo farmacéuticamente aceptable con la cantidad terapéuticamente eficaz del polipéptido. En otras realizaciones se describe un procedimiento de tratamiento de un trastorno, afección o enfermedad, en el que R_1 y R_2 son ambos polipéptidos. En algunas realizaciones se describe un procedimiento de tratamiento de un trastorno, afección o enfermedad, en el que X es un polímero soluble en agua. En otras realizaciones se describe un procedimiento de tratamiento de un trastorno, afección o enfermedad, en el que X es un derivado de polietilenglicol. En otra realización se describe un procedimiento de tratamiento de un trastorno, afección o enfermedad, en el que X es un compuesto citotóxico.
En algunas realizaciones se describe un procedimiento de tratamiento de un trastorno, afección o enfermedad, en el que X es un fármaco.

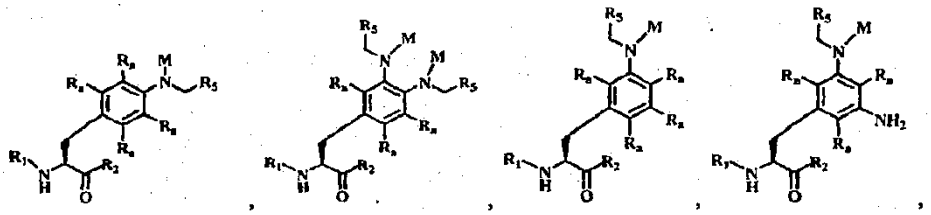
En algunas realizaciones se describe un procedimiento de tratamiento de un trastorno, afección o enfermedad, en el que X es un segundo polipéptido. En otras realizaciones es un procedimiento de tratamiento de un trastorno, afección o enfermedad, en el que el segundo polipéptido es un péptido que contiene un polipéptido de aminoácidos no naturales.

En algunas realizaciones se describe un procedimiento de tratamiento de un trastorno, afección o enfermedad, en el que el polipéptido es una proteína homóloga a una proteína terapéutica seleccionada del grupo que consiste en: alfa-1 antitripsina, angiostatina, factor antihemolítico, anticuerpo, fragmento de anticuerpo, apolipoproteína, apoproteína, factor natriurético auricular, polipéptido natriurético auricular, péptido auricular, quimiocina C-X-C, T39765, NAP-2, ENA-78, gro-a, gro-b, gro-c, IP-10, GCP-2, NAP-4, SDF-1, PF4, MIG, calcitonina, ligando c-kit, citocina, quimiocina CC, proteína 1 quimioatrayente de monocitos, proteína 2 quimioatrayente de monocitos, proteína 3 quimioatrayente de monocitos, proteína 1 alfa inflamatoria de monocitos, proteína 1 beta inflamatoria de monocitos, RANTES, 1309, R83915, R91733, HCC1, T58847, D31065, T64262, CD40, ligando CD40, ligando c-kit, colágeno, factor estimulante de colonias (CSF), factor 5a del complemento, inhibidor del complemento, receptor 1 del complemento, citocina, péptido 78 activante de neutrófilos epiteliales, MIP-16, MCP-1, factor de crecimiento epidérmico (EGF), péptido activante de neutrófilos epiteliales, eritropoyetina (EPO), toxina exfoliante, factor IX, factor VII, factor VIII, factor X, factor de crecimiento de fibroblastos (FGF), fibrinógeno, fibronectina, proteína de cuatro haces helicoidales, G-CSF, glp-1, GM-CSF, glucocerebrosidasa, gonadotropina, factor de crecimiento, receptor de factores de crecimiento, grf, proteína hedgehog, hemoglobina, factor de crecimiento de hepatocitos (hGF), hirudina, hormona de crecimiento humana (hGH), albúmina de suero humano, ICAM-1, receptor de ICAM-1, LFA-1, receptor de LFA-1, insulina, factor de crecimiento similar a la insulina (IGF), IGF-I, IGF-II, interferón (IFN), IFN-alfa, IFN-beta, IFN-gamma, interleucina (IL), IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, factor de crecimiento de queratinocitos (KGF), lactoferrina, factor inhibidor de la leucemia, luciferasa, neurturina, factor inhibidor de neutrófilos (NIF), oncostatina M, proteína osteogénica, producto oncogénico, paracitonina, hormona paratiroidea, PD-ECSF, PDGF, hormona peptídica, pleiotropina, proteína A, proteína G, pth, endotoxina A pirogénica, exotoxina B pirogénica, exotoxina C pirogénica, pyy, relaxina, renina, SCF, proteína biosintética pequeña, receptor I del complemento soluble, I-CAM 1 soluble, receptor de interleucina soluble, receptor de TNF soluble, somatomedina, somatostatina, somatotropina, estreptocinasa, superantígenos, enterotoxina estafilocócica, SEA, SEB, SEC1, SEC2, SEC3, SED, SEE, receptor de hormonas esteroideas, superóxido dismutasa, toxina del síndrome del choque tóxico, timosina alfa 1, activador tisular del plasminógeno, factor de crecimiento tumoral (TGF), factor de necrosis tumoral, factor de necrosis tumoral alfa, factor de necrosis tumoral beta, receptor del factor de necrosis tumoral (TNFR), proteína VLA-4, proteína VCAM-1, factor de crecimiento endotelial vascular (VEGF), urocinasa, mos, ras, raf, met,

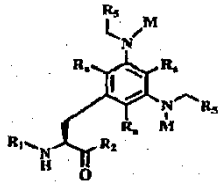
p53, tat, fos, myc, jun, myb, rel, receptor de estrógenos, receptor de progesterona, receptor de testosterona, receptor de aldosterona, receptor de LDL y corticosterona.

En otras realizaciones se describe un procedimiento de tratamiento de un trastorno, afección o enfermedad, en el que el al menos un aminoácido no natural se incorpora en un sitio específico dentro del polipéptido. En algunas realizaciones se describe un procedimiento de tratamiento de un trastorno, afección o enfermedad, en el que el polipéptido se sintetiza por un ribosoma. En algunas realizaciones se describe un procedimiento de tratamiento de un trastorno, afección o enfermedad, en el que el polipéptido que comprende al menos un aminoácido no natural es estable en disolución acuosa durante al menos 1 mes. En otras realizaciones se describe un procedimiento de tratamiento de un trastorno, afección o enfermedad, en el que el polipéptido que comprende al menos un aminoácido no natural es estable durante al menos 2 semanas. En algunas realizaciones se describe un procedimiento de tratamiento de un trastorno, afección o enfermedad, en el que el polipéptido que comprende al menos un aminoácido no natural es estable durante al menos 5 días.

Adicionalmente, en el presente documento se describe un procedimiento de detección de la presencia de un polipéptido en un paciente, comprendiendo el procedimiento administrar una cantidad eficaz de un polipéptido de aminoácidos no naturales homólogo que comprende al menos un aminoácido no natural seleccionado del grupo que consiste en:



y



;

en el que

cada R_a está seleccionado independientemente del grupo que consiste en H, halógeno, alquilo, $-NO_2$, -CN, alquilo sustituido, $-N(R')_2$, $-C(O)_kR'$, $-C(O)N(R')_2$, -OR' y $-S(O)_kR'$ en la que k es 1, 2 ó 3;

M es H o $-CH_2R_5$; o el resto M-N-C(R_5) puede formar una estructura de anillo de 4 a 7 miembros;

R_1 es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido;

R_2 es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;

R_5 es alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alquinilo, alquinilo sustituido, alcoxi, alcoxi sustituido, alquilalcoxi, alquilalcoxi sustituido, poli(óxido de alquileno), poli(óxido de alquileno) sustituido, cicloalquilo, cicloalquilo sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido, -C(O)R", -C(O)OR", $-C(O)N(R")_2$, $-C(O)NHCH(R")_2$, -(alquilen o alquilen sustituido)-$N(R")_2$; -(alquenilen o alquenilen sustituido)-$N(R")_2$, -(alquilen o alquilen sustituido)-(arilo o arilo sustituido), -(alquenilen o alquenilen sustituido)-(arilo o arilo sustituido), -(alquilen o alquilen sustituido)-$ON(R")_2$, -(alquilen o alquilen sustituido)-C(O)SR", -(alquilen o alquilen sustituido)-S-S-(arilo o arilo sustituido), en las que cada R" es independientemente hidrógeno, alquilo, alquilo sustituido, alquenilo, alquenilo sustituido, alcoxi, alcoxi sustituido, arilo, arilo sustituido, heteroarilo, heteroarilo sustituido, heterociclo, heterociclo sustituido, alcarilo, alcarilo sustituido, aralquilo, aralquilo sustituido, o -C(O)OR';

o R_5 es L-X, en la que X se selecciona del grupo que consiste en una marca; un colorante; un polímero; un polímero soluble en agua; un derivado de polietilenglicol; un fotorreticulante; un compuesto citotóxico; un fármaco; una marca de afinidad; una marca de fotoafinidad; un compuesto reactivo; una resina; una segunda proteína o polipéptido o análogo de polipéptido; un anticuerpo o fragmento de anticuerpo; un quelante de metal; un cofactor; un ácido graso; un hidrato de carbono; un polinucleótido; un ADN; un ARN; un polinucleótido antisentido; un sacárido, un dendrímero soluble en agua, una ciclodextrina, un biomaterial; una nanopartícula; una marca de espín; un fluoróforo, un resto que contiene metal; un resto radiactivo; un grupo funcional novedoso; un grupo que interacciona covalentemente o no covalentemente con otras moléculas; un resto fotoenjaulado; un resto excitable por radiación actínica; un ligando; un resto fotoisomerizable;

biotina; un análogo de biotina; un resto que incorpora un átomo pesado; un grupo químicamente escindible; un grupo fotoescindible; una cadena lateral alargada; un azúcar ligado a carbono; un agente activo para rédox; un aminotioácido; un resto tóxico; un resto isotópicamente marcado; una sonda biofísica; un grupo fosforescente; un grupo quimioluminiscente; un grupo con densidad electrónica; un grupo magnético; un grupo intercalante; un cromóforo; un agente de transferencia de energía; un agente biológicamente activo; una marca detectable; una molécula pequeña; un ácido nucleico inhibidor; un radionucleótido; un agente de captura de neutrones; un derivado de biotina; punto(s) cuántico(s); un nanotransmisor; un radiotransmisor; una abzima, un activador del complejo activado, un virus, un adyuvante, un aglicano, un alérgeno, una angiostatina, una antihormona, un antioxidante, un aptámero, un ARN guía, una saponina, un vector lanzadera, una macromolécula, un mimótope, un receptor, una micela inversa, y cualquier combinación de los mismos; y L es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno, alquileno sustituido, alquenileno, alquenileno sustituido, -O-, -O-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-O-, -C(O)-, -C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)-, -C(O)N(R')-, -C(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-C(O)N(R'), -OC(O)N(R')-, -OC(O)N(R')-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-OC(O)N(R')-, -N(R')C(O)-, -NR'C(O)-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-NR'C(O)-, -S-, -S-(alquilen o alquilen sustituido)-, $-S(O)_k-$ en la que k es 1, 2 ó 3, $-S(O)_k$(alquilen o alquilen sustituido)-, -C(S)-, -C(S)-(alquilen o alquilen sustituido)-, -CSN(R')-, -CSN(R')-(alquilen o alquilen sustituido)-, -N(R')CO-(alquilen o alquilen sustituido)-, -N(R')C(O)O-, -(alquilen o alquilen sustituido)-O-N=CR'-, -(alquilen o alquilen sustituido)-C(O)NR'-(alquilen o alquilen sustituido)-, -(alquilen o alquilen sustituido)-$S(O)_k$-(alquilen o alquilen sustituido)-S-, -(alquilen o alquilen sustituido)-S-S-, $-S(O)_kN(R')-$, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, $-N(R')S(O)_kN(R')-$, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, $-C(R')_2-N=N-$ y $-C(R')_2-N(R')-N(R')-$;

o R_5 y cualquier R_a forman opcionalmente un cicloalquilo o un heterocicloalquilo;
cada R' es independientemente H, alquilo o alquilo sustituido; y con la condición de que si R_1 es H, entonces R_2 no es OH, o si R_2 es OH, entonces R_1 no es H.

En el presente documento se describe un procedimiento de detección de la presencia de un polipéptido en un paciente en el que el al menos un aminoácido no natural se incorpora en un sitio específico dentro del polipéptido. En algunas realizaciones se describe un procedimiento de detección de la presencia de un polipéptido en un paciente, en el que el aminoácido no natural se incorpora usando un sistema de traducción. En algunas realizaciones se describe un procedimiento de detección de la presencia de un polipéptido en un paciente, en el que el aminoácido no natural se incorpora en el polipéptido usando un sistema de traducción y un sistema de modificación postraduccional. En otras realizaciones se describe un procedimiento de detección de la presencia de un polipéptido en un paciente, en el que el al menos un aminoácido no natural es estable en disolución acuosa durante al menos 1 mes. En algunas realizaciones se describe un procedimiento de detección de la presencia de un polipéptido en un paciente, en el que el al menos un aminoácido no natural es estable durante al menos 2 semanas. En otras realizaciones se describe un procedimiento de detección de la presencia de un polipéptido en un paciente, en el que el al menos un aminoácido no natural es estable durante al menos 5 días.

En algunas realizaciones se describe un procedimiento de detección de la presencia de un polipéptido en un paciente, en el que el polipéptido es una proteína homóloga a una proteína terapéutica seleccionada del grupo que consiste en: alfa-1 antitripsina, angiostatina, factor antihemolítico, anticuerpo, apolipoproteína, apoproteína, factor natriurético auricular, polipéptido natriurético auricular, péptido auricular, quimiocina C-X-C, T39765, NAP-2, ENA-78, gro-a, gro-b, gro-c, IP-10, GCP-2, NAP-4, SDF-1, PF4, MIG, calcitonina, ligando c-kit, citocina, quimiocina CC, proteína 1 quimioatrayente de monocitos, proteína 2 quimioatrayente de monocitos, proteína 3 quimioatrayente de monocitos, proteína 1 alfa inflamatoria de monocitos, proteína 1 beta inflamatoria de monocitos, RANTES, 1309, R83915, R91733, HCC1, T58847, D31065, T64262, CD40, ligando CD40, ligando c-kit, colágeno, factor estimulante de colonias (CSF), factor 5a del complemento, inhibidor del complemento, receptor 1 del complemento, citocina, péptido 78 activante de neutrófilos epiteliales, MIP-16, MCP-1, factor de crecimiento epidérmico (EGF), péptido activante de neutrófilos epiteliales, eritropoyetina (EPO), toxina exfoliante, factor IX, factor VII, factor VIII, factor X, factor de crecimiento de fibroblastos (FGF), fibrinógeno, fibronectina, proteína de cuatro haces helicoidales, G-CSF, glp-1, GM-CSF, glucocerebrosidasa, gonadotropina, factor de crecimiento, receptor de factores de crecimiento, grf, proteína hedgehog, hemoglobina, factor de crecimiento de hepatocitos (hGF), hirudina, hormona de crecimiento humana (hGH), albúmina de suero humano, ICAM-1, receptor de ICAM-1, LFA-1, receptor de LFA-1, insulina, factor de crecimiento similar a la insulina (IGF), IGF-I, IGF-II, interferón (IFN), IFN-alfa, IFN-beta, IFN-gamma, interleucina (IL), IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, factor de crecimiento de queratinocitos (KGF), lactoferrina, factor inhibidor de la leucemia, luciferasa, neurturina, factor inhibidor de neutrófilos (NIF), oncostatina M, proteína osteogénica, producto oncogénico, paracitonina, hormona paratiroidea, PD-ECSF, PDGF, hormona peptídica, pleiotropina, proteína A, proteína G, pth, endotoxina A pirogénica, exotoxina B pirogénica, exotoxina C pirogénica, pyy, relaxina, renina, SCF, proteína biosintética pequeña, receptor I del complemento soluble, I-CAM 1 soluble, receptor de interleucina soluble, receptor de TNF soluble, somatomedina, somatostatina, somatotropina, estreptocinasa, superantígenos, enterotoxina estafilocócica, SEA, SEB, SEC1, SEC2, SEC3, SED, SEE, receptor de

hormonas esteroideas, superóxido dismutasa, toxina del síndrome del choque tóxico, timosina alfa 1, activador tisular del plasminógeno, factor de crecimiento tumoral (TGF), factor de necrosis tumoral, factor de necrosis tumoral alfa, factor de necrosis tumoral beta, receptor del factor de necrosis tumoral (TNFR), proteína VLA-4, proteína VCAM-1, factor de crecimiento endotelial vascular (VEGF), urocinasa, 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 corticosterona.

EJEMPLOS

Los siguientes ejemplos describen procedimientos para la síntesis de aminoácidos que contienen aminas aromáticas.

Ejemplo 1a: Síntesis de 3-cloro-4-amino-fenilalanina

Se sintetizó 3-cloro4-amino-fenilalanina según el procedimiento en la FIG. 11.

Ejemplo 1b: Síntesis de 3-yodo-4-amino-fenilalanina

Se sintetizó 3-yodo-4-amino-fenilalanina según el procedimiento en la FIG. 11.

Ejemplo 1c: Síntesis de 3-metoxi-4-amino-fenilalanina

Se sintetizó 3-metoxi-4-amino-fenilalanina según el procedimiento en la FIG. 12.

Ejemplo 1d: Síntesis de 3-fluoro-4-amino-fenilalanina

Se sintetizó 3-fluoro-4-amino-fenilalanina según el procedimiento en la FIG. 12.

Ejemplo 1e: Síntesis de N-metil-p-amino-fenilalanina

Se sintetizó N-metil-p-amino-fenilalanina según el procedimiento en la FIG. 13.

Ejemplo 1f: Síntesis de N-etil-p-amino-fenilalanina

Se sintetizó N-etil-p-amino-fenilalanina según el procedimiento en la FIG. 13.

Los siguientes ejemplos describen procedimientos de clonación y expresión de un polipéptido modificado, seguido de modificación postraduccional

Ejemplo 2: PEGilación de hGH

Este ejemplo detalla la clonación y expresión de un polipéptido modificado en *E. coli*. Se usa un sistema de traducción introducido que comprende un ARNt ortogonal (O-ARNt) y una aminoacil ARNt sintetasa ortogonal (O-RS) para expresar el polipéptido que contiene un aminoácido no natural. La O-RS aminoacila preferencialmente el O-ARNt con un aminoácido no natural. A su vez, el sistema de traducción inserta el aminoácido no natural en el polipéptido, en respuesta a un selector de codones codificado. La transformación de *E. coli* con plásmidos que contienen el gen modificado y el par ortogonal de aminoacil ARNt sintetasa/ARNt (específico para el aminoácido no natural deseado) permite la incorporación específica del sitio del aminoácido no natural en el polipéptido. La *E. coli* transformada, cultivada a 37 ºC en medio que contiene entre 0,01 - 100 mM del aminoácido no natural particular, expresa polipéptido modificado con alta fidelidad y eficiencia. El polipéptido marcado con His que contiene un aminoácido no natural se produce por las células huésped de *E. coli* como cuerpos de inclusión o agregados. Los agregados se solubilizan y se purifican por afinidad bajo condiciones desnaturalizantes en HCl de guanidina 6 M. El replegamiento se realiza por diálisis a 4 ºC durante la noche en TRIS-HCl 50 mM, pH 8,0, $CuSO_4$ 40 µM y 2 % (peso/volumen) de Sarkosyl. A continuación, el material se dializa contra TRIS-HCl 20 mM, pH 8,0, NaCl 100 mM, $CaCl_2$ 2 mM, seguido de iminación de la marca de His. Véase Boissel y col., (1993) 268:15983-93. Los procedimientos de purificación de polipéptidos son muy conocidos en la técnica y se confirman por SDS-PAGE, análisis de transferencia Western o espectrometría de masas con trampa de iones por electropulverización-ionización y similares.

A modo de ejemplo, el polipéptido de hGH con tirosina sustituida por p-aminofenilalanina en la posición 35 (H6-hGH Y35pAF2) se generó en células de *E. coli* usando construcciones que codifican hGH, y un par de ARNt sintetasa ortogonal-ARNt ortogonal para el aminoácido no natural. La proteína expresada se purificó usando IMAC y cromatografía IEX. La proteína se dializó en fosfato de sodio 10 mM, 20 g/l de glicina, 5 g/l de manitol, pH 7,0, y a continuación se concentró a 350 uM. El pH de la muestra se ajustó a pH 4,0 usando 10 % de ácido acético.

La modificación postraduccional de H6-hGH Y35pAF2 se demostró por la pegilación de H6-hGH Y35pAF2 por

alquilación reductora de la p-aminofenilalanina con diversos PEG-aldehídos. Un procedimiento no limitante para tal pegilación se expone brevemente en la FIG. 39 y se describe en el presente documento del siguiente modo: se establecieron reacciones de PEGilación de 100 ul usando PEG de 20K, 30K y 40K-aldehído. El PEG-aldehído estuvo en tampón acetato 20 mM a pH 4,0. Cada reacción tuvo una relación molar 1:1 de PEG con respecto a proteína, y una relación molar 5:1 de $NaCNBH_3$ solubilizado en DMF:proteína. Las reacciones se incubaron tanto a 4 grados como a temperatura ambiente y se analizaron por SDS-PAGE después de 3, 4, 6 y 16 h. La electroforesis en gel de los péptidos resultantes se muestra en la FIG. 40

Los siguientes ejemplos describen procedimientos de modificación de polipéptidos por modificación postraduccional como se muestra en las FIG. 20-34, en las que X es p-aminofenilalanina.

Ejemplo 3a: Alquilaciones reductoras de urotensina-II (UT-1I-SH) reducida con propionaldehído y benzaldehído

Se añadió una mezcla de 3 equivalentes de $NaBCNH_3$ a pH 4 a una mezcla de propionaldehído 0,25 mM (o benzaldehído 0,25 mM) y urotensina reducida 0,25 mM y se dejó reaccionar durante 2 horas. El análisis de la mezcla de productos fue por HPLC con las siguientes condiciones: 0-50 % de B (A, 0,05 % de TFA en agua; B, 60 % de acetonitrilo y 0,05 % de TFA en agua), velocidad de flujo: 1,5 ml/min; columna: Zorbax Extend C18, 3,5 m, 4,6 x 50 mm.

Ejemplo 3b: Alquilaciones reductoras de urotensina-II (UT-II) con propionaldehído y benzaldehído

Se añadió una mezcla de 3 equivalentes de $NaBCNH_3$ a pH 4 a una mezcla de propionaldehído 0,25 mM (o benzaldehído 0,25 mM) y urotensina 0,25 mM y se dejó reaccionar durante 2 horas. El análisis de la mezcla de productos fue por HPLC con las siguientes condiciones: 5-60 % de B (A, 0,05 % de TFA en agua; B, 60 % de acetonitrilo y 0,05 % de TFA en agua), velocidad de flujo: 1,5 ml/min; columna: Zorbax Extend C18, 3,5 m, 4,6 x 50 mm.

Ejemplo 3c: Alquilaciones reductoras del péptido XT-8 con propionaldehído, benzaldehído, isobutaldehído y pivalaldehído

Se añadió una mezcla de 3 equivalentes de $NaBCNH_3$ a pH 4 a una mezcla de propionaldehído 0,25 mM (o benzaldehído 0,25 mM, o isobutaldehído 0,25 mM o pivalaldehído 0,25 mM) y XT-8 0,25 mM y se dejó reaccionar durante 2 horas. El análisis de la mezcla de productos fue por HPLC con las siguientes condiciones: 5-60 % de B (A, 0,05 % de TFA en agua; B, 60 % de acetonitrilo y 0,05 % de TFA en agua), velocidad de flujo: 1,5 ml/min; columna: Zorbax Extend C18, 3,5 m, 4,6 x 50 mm.

Ejemplo 3d: Alquilaciones reductoras del péptido SXT-9 con propionaldehído, benzaldehído, isobutaldehído y pivalaldehído

Se añadió una mezcla de 3 equivalentes de $NaBCNH_3$ a pH 4 a una mezcla de propionaldehído 0,25 mM (o benzaldehído 0,25 mM, o isobutaldehído 0,25 mM o pivalaldehído 0,25 mM) y SXT-9 0,25 mM y se dejó reaccionar durante 2 horas. El análisis de la mezcla de productos fue por HPLC con las siguientes condiciones: 5-60 % de B (A, 0,05 % de TFA en agua; B, 60 % de acetonitrilo y 0,05 % de TFA en agua), velocidad de flujo: 1,5 ml/min; columna: Zorbax Extend C18, 3,5 m, 4,6 x 50 mm.

Ejemplo 3e: Alquilaciones reductoras del péptido HXT-9 con propionaldehído, benzaldehído, isobutaldehído y pivalaldehído

Se añadió una mezcla de 3 equivalentes de $NaBCNH_3$ a pH 4 a una mezcla de propionaldehído 0,25 mM (o benzaldehído 0,25 mM, o isobutaldehído 0,25 mM o pivalaldehído 0,25 mM) y HXT-9 0,25 mM y se dejó reaccionar durante 2 horas. El análisis de la mezcla de productos fue por HPLC con las siguientes condiciones: 5-60 % de B (A, 0,05 % de TFA en agua; B, 60 % de acetonitrilo y 0,05 % de TFA en agua), velocidad de flujo: 1,5 ml/min; columna: Zorbax Extend C18, 3,5 m, 4,6 x 50 mm.

Ejemplo 3f: Alquilaciones reductoras del péptido WXT-9 con propionaldehído, benzaldehído e isobutaldehído

Se añadió una mezcla de 3 equivalentes de $NaBCNH_3$ a pH 4 a una mezcla de propionaldehído 0,25 mM (o benzaldehído 0,25 mM, o isobutaldehído 0,25 mM) y WXT-9 0,25 mM y se dejó reaccionar durante 2 horas. El análisis de la mezcla de productos fue por HPLC con las siguientes condiciones: 5-60 % de B (A, 0,05 % de TFA en agua; B, 60 % de acetonitrilo y 0,05 % de TFA en agua), velocidad de flujo: 1,5 ml/min; columna: Zorbax Extend C18, 3,5 m, 4,6 x 50 mm.

Ejemplo 3g: Alquilaciones reductoras del péptido NXT-9 con propionaldehído y benzaldehído

Se añadió una mezcla de 3 equivalentes de $NaBCNH_3$ a pH 4 a una mezcla de propionaldehído 0,25 mM (o benzaldehído 0,25 mM) y NXT-9 0,25 mM y se dejó reaccionar durante 2 horas. El análisis de la mezcla de

productos fue por HPLC con las siguientes condiciones: 5-60 % de B (A, 0,05 % de TFA en agua; B, 60 % de acetonitrilo y 0,05 % de TFA en agua), velocidad de flujo: 1,5 ml/min; columna: Zorbax Extend C18, 3,5 m, 4,6 x 50 mm.

Ejemplo 3h: Alquilaciones reductoras del péptido RXT-10 con propionaldehído o benzaldehído

Se añadió una mezcla de 3 equivalentes de $NaBCNH_3$ a pH 4 a una mezcla de propionaldehído 0,25 mM (o benzaldehído 0,25 mM) y RXT-10 0,25 mM y se dejó reaccionar durante 2 horas. El análisis de la mezcla de productos fue por HPLC con las siguientes condiciones: 5-60 % de B (A, 0,05 % de TFA en agua; B, 60 % de acetonitrilo y 0,05 % de TFA en agua), velocidad de flujo: 1,5 ml/min; columna: Zorbax Extend C18, 3,5 m, 4,6 x 50 mm.

Ejemplo 3i: Alquilaciones reductoras del péptido AXT-11 con propionaldehído o benzaldehído

Se añadió una mezcla de 3 equivalentes de $NaBCNH_3$ a pH 4 a una mezcla de propionaldehído 0,25 mM (o benzaldehído 0,25 mM) y AXT-11 0,25 mM y se dejó reaccionar durante 2 horas. El análisis de la mezcla de productos fue por HPLC con las siguientes condiciones: 5-60 % de B (A, 0,05 % de TFA en agua; B, 60 % de acetonitrilo y 0,05 % de TFA en agua), velocidad de flujo: 1,5 ml/min; columna: Zorbax Extend C18, 3,5 m, 4,6 x 50 mm.

Ejemplo 3j: Alquilaciones reductoras del péptido AXT-11 con 3-fenilpropanal, 2-fenilacetaldehído o cinamaldehído

Se añadió una mezcla de 3 equivalentes de $NaBCNH_3$ a pH 4 a una mezcla de 3-fenilpropanal 0,25 mM (o 2-fenilacetaldehído 0,25 mM o cinamaldehído 0,25 mM) y AXT-11 0,25 mM y se dejo reaccionar durante 5 horas. El análisis de la mezcla de productos fue por HPLC con las siguientes condiciones: 5-60 % de B (A, 0,05 % de TFA en agua; B, 60 % de acetonitrilo y 0,05 % de TFA en agua), velocidad de flujo: 1,5 ml/min; columna: Zorbax Extend C18, 3,5 m, 4,6 x 50 mm.

Ejemplo 3k: Alquilaciones reductoras del péptido AXT-11 con 1H-imidazol-5-carbaldehído, tiofeno-2-carbaldehído, picolinaldehído o quinolin-4-carbaldehído

Se añadió una mezcla de 3 equivalentes de $NaBCNH_3$ a pH 4 a una mezcla de 0,25 mM de 1H-imidazol-5-carbaldehído (o tiofeno-2-carbaldehído 0,25 mM, o picolinaldehído 0,25 mM, o quinolin-4-carbaldehído 0,25 mM) y AXT-11 0,25 mM y se dejó reaccionar durante 5 horas. El análisis de la mezcla de productos fue por HPLC con las siguientes condiciones: 5-60 % de B (A, 0,05 % de TFA en agua; B, 60 % de acetonitrilo y 0,05 % de TFA en agua), velocidad de flujo: 1,5 ml/min; columna: Zorbax Extend C18, 3,5 m, 4,6 x 50 mm.

Ejemplo 31: Alquilaciones reductoras del péptido AXT-11 con benzaldehído, 1-fenilbutano-1,3-diona, o una mezcla de benzaldehído y 1-fenilbutano-1,3-diona

Se añadió una mezcla de 3 equivalentes de $NaBCNH_3$ a pH 4 a una mezcla de 0,20 mM de benzaldehído (o 1-fenilbutano-1,3-diona 0,20 mM, o una mezcla de benzaldehído 0,20 mM y 1-fenilbutano-1,3-diona 0,20 mM) y AXT-11 0,20 mM y se dejó reaccionar durante 2 horas. El análisis de la mezcla de productos fue por HPLC con las siguientes condiciones: 5-60 % de B (A, 0,05 % de TFA en agua; B, 60 % de acetonitrilo y 0,05 % de TFA en agua), velocidad de flujo: 1,5 ml/min; columna: Zorbax Extend C18, 3,5 m, 4,6 x 50 mm.

Ejemplo 3m: Alquilaciones reductoras del péptido NXT-9 con benzaldehído, 1-fenilpropano-1,2-diona, o una mezcla de benzaldehído y 1-fenilpropano-1,2-diona

Se añadió una mezcla de 3 equivalentes de $NaBCNH_3$ a pH 4 a una mezcla de 0,20 mM de benzaldehído (o 1-fenilpropano-1,2-diona 0,20 mM, o una mezcla de benzaldehído 0,20 mM y 1-fenilpropano-1,2-diona 0,20 mM) y NXT-9 0,20 mM y se dejó reaccionar durante 2 horas. El análisis de la mezcla de productos fue por HPLC con las siguientes condiciones: 5-60 % de B (A, 0,05 % de TFA en agua; B, 60 % de acetonitrilo y 0,05 % de TFA en agua), velocidad de flujo: 1,5 ml/min; columna: Zorbax Extend C18, 3,5 m, 4,6 x 50 mm.

Ejemplo 3n: Alquilaciones reductoras del péptido MXT-9 con benzaldehído, acetofenona, una mezcla de benzaldehído y acetofenona, una mezcla de benzaldehído y propionaldehído, o una mezcla de benzaldehído y butan-2-ona

Se añadió una mezcla de 3 equivalentes de $NaBCNH_3$ a pH 4 a una mezcla de 0,20 mM de benzaldehído (o acetofenona 0,20 mM, o una mezcla de benzaldehído 0,20 mM y acetofenona 0,20 mM, o una mezcla de benzaldehído 0,20 mM y propionaldehído 0,20 mM, o una mezcla de benzaldehído 0,20 mM y butan-2-ona 0,20 mM) y MXT-9 0,20 mM y se dejó reaccionar durante 2 horas. El análisis de la mezcla de productos fue por HPLC con las siguientes condiciones: 5-60 % de B (A, 0,05 % de TFA en agua; B, 60 % de acetonitrilo y 0,05 % de TFA en agua), velocidad de flujo: 1,5 ml/min; columna: Zorbax Extend C18, 3,5 m, 4,6 x 50 mm.

Ejemplo 30: Reducción de péptido MXT-9-N_3 seguido de alquilaciones reductoras del péptido MXT-9-NH_2 con

propionaldehído o benzaldehído

Se redujo una mezcla de MXT-9-N_3 con TCEP dando el péptido MXT-9-NH_2. Se añadió una mezcla de 3 equivalentes de $NaBCNH_3$ a pH 4 a una mezcla de 0,20 mM de propionaldehído (o benzaldehído 0,20 mM) y MXT-9-NH_2 0,20 mM y se dejó reaccionar durante 1 hora. El análisis de la mezcla de productos fue por HPLC con las siguientes condiciones: 5-60 % de B (A, 0,05 % de TFA en agua; B, 60 % de acetonitrilo y 0,05 % de TFA en agua), velocidad de flujo: 1,5 ml/min; columna: Zorbax Extend C18, 3,5 m, 4,6 x 50 mm.

Los siguientes ejemplos describen procedimientos para medir y comparar la actividad *in vitro* e *in vivo* de un polipéptido de aminoácidos no naturales terapéuticamente activo modificado con la actividad *in vitro* e *in vivo* de un polipéptido de aminoácidos naturales terapéuticamente activo.

Ejemplo 4: Medición de la actividad y afinidad de polipéptidos de aminoácidos no naturales

Este ejemplo detalla la medición de la actividad y afinidad de polipéptidos de aminoácidos no naturales de polipéptidos de aminoácidos no naturales para su receptor, componente de unión o ligando.

La proteína para el receptor, componente de unión o ligando del polipéptido de aminoácidos no naturales se expresa y aísla según procedimientos conocidos para aquellos expertos habituales en la materia. El sistema Biacore™ se usa para analizar la unión del polipéptido de aminoácidos no naturales a su receptor. Similarmente, puede usarse un componente de unión o ligando en este ensayo.

Se inmovilizan aproximadamente 600-800 UR de receptor soluble sobre un chip CM5 Biacore™ usando un procedimiento de acoplamiento a amina convencional, como se recomienda por el fabricante. Se inyectan diversas concentraciones de polipéptido de aminoácidos no naturales no mutante o (modificado) en tampón HBS-EP (Biacore™, Pharmacia) sobre la superficie a una velocidad de flujo de 40 µl/min durante 4-5 minutos, y se monitoriza la disociación durante 15 minutos después de la inyección. La superficie se regenera por un pulso de 15 segundos de $MgCl_2$ 4,5 M. Solo se observa una pérdida de afinidad de unión mínima (1-5 %) después de al menos 100 ciclos de regeneración. Se usa una célula de referencia sin receptor inmovilizado para restar cualquier efecto volumétrico del tampón y unión no específica.

Los datos de unión cinética obtenidos de los experimentos de valoración de polipéptidos de aminoácidos no naturales (modificados) se procesan con el software BiaEvaluation 4.1 (BIACORE™). Las constantes de equilibrio de disociación (Kd) se calculan como relaciones de las constantes de velocidad individuales (k_{as}/k_{dis}).

Se establecen líneas celulares estables que expresan el receptor, componente de unión o ligando para el polipéptido de aminoácidos no naturales. Las células se electroporan con una construcción que contiene el ADNc del receptor, componente de unión o ligando. Se deja que las células transfectadas se recuperen durante 48 horas antes de la clonación. Los transfectantes que expresan receptor, componente de unión o ligando se identifican por tinción superficial con anticuerpo contra el receptor y se analizan sobre una matriz FACS (BD Biosciences, San Diego, CA). Se establecen clones de células establemente transfectadas tras rondas adicionales de subclonación repetida de transfectantes deseados. Tales células se usan en ensayos de unión a células.

Se incuban células ($3x10^6$) por duplicado en PBS/1 % de BSA (100 µl) en ausencia o presencia de diversas concentraciones (volumen: 10 µl) de polipéptido de aminoácidos naturales no marcado o un polipéptido de control negativo y en presencia de polipéptido de aminoácidos no naturales (modificado con)-^{125}I (aprox. 100.000 cpm o 1 ng) a 0 ºC durante 90 minutos (volumen total: 120 µl). A continuación, las células se resuspenden y se estratifican sobre 200 µl de SBF helado en un tubo de centrífuga de plástico de 350 µl y se centrifugan (1000 g; 1 minuto). El sedimento se recoge cortando el extremo del tubo y el sedimento y el sobrenadante se cuentan por separado en un contador gamma (Packard).

La unión específica (cpm) se determina como la unión total en ausencia de un competidor (media de duplicados) menos la unión no específica. La unión no específica se mide para cada uno de los tipos de células usadas. Los experimentos se realizan en días separados usando la misma preparación de polipéptido de aminoácidos no naturales (modificado con)-^{125}I y deben mostrar coherencia interna. El polipéptido de aminoácidos no naturales (modificado con)-^{125}I demuestra unión al receptor, proteína de unión o células productoras de ligando. La unión se inhibe en un modo dependiente de la dosis por polipéptido de aminoácidos naturales no marcado, pero no por un polipéptido de control negativo.

Ejemplo 5: Estudios *in vivo* de polipéptido de aminoácidos no naturales terapéuticamente activo modificado

Se administran polipéptido de aminoácidos no naturales terapéuticamente activo modificado, polipéptido de aminoácidos naturales terapéuticamente activo y disolución de tampón a ratones o ratas. Los resultados mostrarán actividad superior y semivida prolongada del polipéptido de aminoácidos no naturales terapéuticamente activo modificado en comparación con el polipéptido de aminoácidos naturales terapéuticamente activo.

Ejemplo 6: Medición de la semivida *in vivo* de polipéptido de aminoácidos no naturales terapéuticamente activo modificado conjugado y no conjugado y variantes del mismo

Toda la experimentación animal se realiza en una instalación acreditada por la AAALAC y bajo protocolos autorizados por el Comité Institucional de Cuidado y Uso Animal de la Universidad de St. Loues. Las ratas se alojan individualmente en jaulas en habitaciones con un ciclo de luz/oscuridad de 12 horas. Se proporciona acceso de los animales a pienso para roedores certificado 5001 de Purina y agua a voluntad.

Ejemplo 7: Estudios farmacocinéticos

La calidad de cada polipéptido de aminoácidos no naturales terapéuticamente activo modificado se evalúa por tres ensayos antes de entrar en experimentos en animales. La pureza del polipéptido de aminoácidos no naturales terapéuticamente activo modificado se examina migrando un gel NuPAGE Bis-Tris de 4-12 % de acrilamida con tampón de electroforesis MES-SDS bajo condiciones no reductoras (Invitrogen, Carlsbad, CA). Los geles se tiñen con azul de Coomassie. La banda del polipéptido de aminoácidos no naturales terapéuticamente activo modificado tiene más del 95 % pureza basándose en el barrido de densitometría. El nivel de endotoxina en cada polipéptido de aminoácidos no naturales terapéuticamente activo modificado se prueba por un ensayo de LAL cinético usando el kit KTA^2 de Charles River Laboratories (Wilmington, MA), y es inferior a 5 UE por dosis. La actividad biológica del polipéptido de aminoácidos no naturales terapéuticamente activo modificado se evalúa con los ensayos celulares que caracterizan la bioactividad del polipéptido.

Se comparan las propiedades farmacocinéticas de compuestos de polipéptido de aminoácidos no naturales terapéuticamente activo modificado entre sí y con polipéptido de aminoácidos naturales terapéuticamente activo en ratas Sprague-Dawley macho (261-425 g) obtenidas de Charles River Laboratories. Se colocan catéteres quirúrgicamente en la arteria carótida para la recogida de sangre. Tras la satisfactoria colocación del catéter, los animales se asignan a grupos de tratamiento (tres a seis por grupo) antes de la dosificación. Los animales se dosifican subcutáneamente con 1 mg/kg de compuesto en un volumen de dosis de 0,41-0,55 ml/kg. Se recogen muestras de sangre en diversos momentos de tiempo mediante el catéter permanente y en tubos de microcentrífuga recubiertos con EDTA. El plasma se recoge después de la centrifugación, y se guarda a -80 ºC hasta el análisis. Las concentraciones de compuesto se miden usando kits de ELISA de sándwich para anticuerpos de tanto BioSource International (Camarillo, CA) como de Diagnostic Systems Laboratories (Webster, TX). Las concentraciones se calculan usando patrones correspondientes al análogo que se dosifica. Se estiman parámetros farmacocinéticos usando el programa de modelado WinNonlin (Pharsight, versión 4.1). Se usa análisis no compartimental con integración trapezoidal lineal ascendente/logarítmica descendente, y los datos de concentración se ponderan uniformemente. A continuación, los datos se representan para obtener Cmáx: concentración máxima; $t_{1/2}$ terminal: semivida terminal; $ABC_{0\rightarrow inf}$: área bajo la curva de concentración-tiempo extrapolada al infinito; MRT: tiempo de residencia medio; Cl/f: eliminación de plasma total aparente; y Vz/f: volumen de distribución aparente durante la fase terminal.

Ejemplo 8: Estudios farmacodinámicos

Se obtienen ratas Sprague-Dawley macho de Charles River Laboratories. Se deja que los animales se aclimaten durante un periodo de tres semanas, tiempo durante el cual se monitorizan las características biológicas asociadas al polipéptido de aminoácidos naturales. Animales con un nivel aceptable de cambio en estas características biológicas se aleatorizan a grupos de tratamiento. Las ratas se administran tanto con una dosis en bolo como con dosis diaria subcutáneamente del polipéptido de aminoácidos no naturales modificado. Durante todo el estudio, las ratas se anestesian diariamente y secuencialmente, se sangran y se dosifican (cuando sea aplicable) y se miden las características biológicas de correlación. La sangre se recoge del seno orbital usando un tubo capilar heparinizado y se coloca en un tubo de microcentrífuga recubierto con EDTA. Se aísla plasma por centrifugación y se guarda a -80 ºC hasta el análisis. Se obtienen concentraciones en plasma tras una única dosis subcutánea en ratas.

Ejemplo 9: Ensayo clínico humano de la seguridad y/o eficacia de polipéptido de aminoácidos no naturales terapéuticamente activo modificado

Objetivo Comparar la seguridad y farmacocinética de un polipéptido de aminoácidos no naturales terapéuticamente activo modificado subcutáneamente administrado con la seguridad y farmacocinética de un polipéptido de aminoácidos naturales terapéuticamente activo.

Pacientes Se enrolan en el estudio dieciocho voluntarios sanos que oscilan entre 20-40 años de edad y que pesan entre 60-90 kg. Los sujetos tendrán valores de laboratorio no anormales clínicamente significativos para hematología o química del suero, y una prueba toxicológica de la orina, prueba del VIH y antígeno de superficie de la hepatitis B negativa. No deben tener ninguna evidencia de lo siguiente: hipertensión; una historia de cualquier enfermedad hematológica primaria; historia de enfermedad hepática, renal, cardiovascular, gastrointestinal, genitourinaria, metabólica, neurológica significativa; una historia de anemia o trastorno convulsivo; una sensibilidad conocida a productos bacterianos o derivados de mamífero, PEG o albúmina de suero humano; consumidor habitual y empedernido de bebidas que contienen cafeína; participación en cualquier otro ensayo clínico o que haya recibido

transfusión de sangre o donado en el plazo de 30 días desde la entrada en el estudio; haya tenido exposición a un polipéptido de aminoácidos naturales terapéuticamente activo en el plazo de tres meses desde la entrada en el estudio; haya tenido una enfermedad en el plazo de siete días desde la entrada en el estudio; y tenga anomalías significativas en la exploración física previa al estudio o las evaluaciones clínicas de laboratorio en el plazo de 14 días desde la entrada en el estudio. Todos los sujetos son evaluables para seguridad y todas las recogidas de sangre para análisis farmacocinético se recogen según lo programado. Todos los estudios se realizan con autorización del comité ético institucional y consentimiento del paciente.

Diseño del estudio Será un diseño cruzado de dos periodos, aleatorizado, de etiqueta abierta, unicéntrico de fase 1 en voluntarios masculinos sanos. Dieciocho sujetos se asignan aleatoriamente a uno de los grupos de la secuencia de tratamiento (nueve sujetos/grupo). Se administra un polipéptido de aminoácidos naturales terapéuticamente activo durante dos periodos de dosificación separados como una inyección s.c. en bolo en el muslo superior usando dosis equivalente del polipéptido de aminoácidos no naturales terapéuticamente activo modificado. Puede añadirse dosificación adicional, frecuencia de dosificación, u otro parámetro, según se desee, al estudio, incluyendo grupos adicionales de sujetos. Cada periodo de dosificación se separa por un periodo de lavado de 14 días. Los sujetos se confinan en el centro de estudio al menos 12 horas antes y 72 horas tras la dosificación durante cada uno de los dos periodos de dosificación, pero no entre los periodos de dosificación. Pueden añadirse grupos adicionales de sujetos si va a haber dosificación adicional, frecuencia, u otro parámetro, que también va a probarse para el polipéptido de aminoácidos no naturales terapéuticamente activo modificado.

Muestreo de sangre Se extraen muestras de sangre en serie por punción directa de la vena antes y después de la administración de polipéptido de aminoácidos no naturales terapéuticamente activo modificado o polipéptido de aminoácidos naturales terapéuticamente activo. Se obtienen muestras de sangre venosa (5 ml) para la determinación del polipéptido de aminoácidos no naturales terapéuticamente activo modificado en suero o concentraciones del polipéptido de aminoácidos naturales terapéuticamente activo aproximadamente 30, 20 y 10 minutos antes de la dosificación (3 muestras de referencia) y en aproximadamente los siguientes tiempos después de la dosificación: 30 minutos y 1, 2, 5, 8, 12, 15, 18, 24, 30, 36, 48, 60 y 72 horas. Cada muestra de suero se divide en dos alícuotas. Todas las muestras de suero se almacenan a -20 ºC. Las muestras de suero se transportan sobre nieve carbónica. Las pruebas de laboratorio clínico en ayunas (hematología, química del suero y análisis de orina) se realizan inmediatamente antes de la dosis inicial en el día 1, la mañana del día 4, inmediatamente antes de la dosificación en el día 16 y la mañana del día 19.

Procedimientos bioanalíticos Se usa un procedimiento con el kit ELISA (Diagnostic Systems Laboratory [DSL], Webster TX) para la determinación de concentraciones en suero.

Determinaciones de seguridad Se registran las constantes vitales inmediatamente antes de cada dosificación (días 1 y 16), y 6, 24, 48 y 72 horas después de cada dosificación. Las determinaciones de seguridad se basan en la incidencia y tipo de acontecimientos adversos y los cambios en las pruebas del laboratorio clínico desde el nivel inicial. Además, se evalúan cambios desde antes del estudio en mediciones de signos vitales, que incluyen tensión arterial y resultados de la exploración física.

Análisis de datos Los valores de concentración en suero después de la dosis se corrigen para las concentraciones del nivel inicial antes de la dosis restando de cada uno de los valores de después de la dosis la concentración del media nivel inicial determinada de promediar los niveles de las tres muestras recogidas 30, 20 y 10 minutos antes de la dosificación. Las concentraciones en suero antes de la dosis no se incluyen en el cálculo del valor medio si están por debajo del nivel de cuantificación del ensayo. Los parámetros farmacocinéticos se determinan a partir de los datos de concentración en suero corregidos para las concentraciones del nivel inicial. Los parámetros farmacocinéticos se calculan por procedimientos independientes del modelo en un sistema informático Digital Equipment Corporation VAX 8600 usando la última versión del software BIOAVL. Se determinan los siguientes parámetros farmacocinéticos: concentración pico en suero ($C_{máx}$); tiempo hasta la concentración pico en suero ($t_{máx}$); área bajo la curva de concentración-tiempo (ABC) desde tiempo cero hasta el último momento de muestreo de sangre (ABC_{0-72}) calculado usando la regla trapezoidal lineal; y semivida de eliminación terminal ($t_{1/2}$), calculada a partir de la constante de velocidad de eliminación. La constante de velocidad de eliminación se estima por regresión lineal de puntos de datos consecutivos en la región lineal terminal de la representación de concentración-tiempo lineal logarítmica. Se calculan la media, desviación estándar (DE) y coeficiente de variación (CV) de los parámetros farmacocinéticos para cada tratamiento. Se calcula la relación de las medias de parámetros (formulación preservada / formulación no preservada).

Resultados de seguridad La incidencia de acontecimientos adversos se distribuye igualmente a través de los grupos de tratamiento. No hay cambios clínicamente significativos desde el nivel inicial o pruebas de laboratorio clínico previas al estudio o tensiones arteriales, y no hay cambios notables de antes del estudio en los resultados de la exploración física y mediciones de constantes vitales. Los perfiles de seguridad para los dos grupos de tratamiento deben parecer similares.

Resultados farmacocinéticos Se comparan los perfiles de concentración-tiempo de polipéptido de aminoácidos no naturales terapéuticamente activo modificado o de polipéptido de aminoácidos naturales terapéuticamente activos

(sin corregir para los niveles de referencia) en suero medio en los 18 sujetos después de recibir una dosis única de polipéptido de aminoácidos no naturales terapéuticamente activo modificado o polipéptido de aminoácidos naturales terapéuticamente activo en cada momento de tiempo medido. Todos los sujetos deben tener concentraciones de referencia previas a la dosis dentro del intervalo fisiológico normal. Se determinan los parámetros farmacocinéticos a partir de los datos en suero corregidos para las concentraciones de referencia medias previas a la dosis y se determinan $C_{máx}$ y $t_{máx}$. El $t_{máx}$ medio para el polipéptido de aminoácidos naturales terapéuticamente activo es significativamente más corto que el $t_{máx}$ para el polipéptido de aminoácidos no naturales terapéuticamente activo modificado. Los valores de semivida terminal son significativamente más cortos para los polipéptido de aminoácidos naturales terapéuticamente activos en comparación con la semivida terminal para el polipéptido de aminoácidos no naturales terapéuticamente activo modificado.

Aunque el presente estudio se realiza en sujetos masculinos sanos, se anticiparían características de absorción y perfiles de seguridad similares en otras poblaciones de pacientes; tales como pacientes masculinos y femeninos con cáncer o insuficiencia renal crónica, pacientes pediátricos con insuficiencia renal, pacientes en programas de predepósito autólogo, o pacientes programados para cirugía electiva.

En conclusión, dosis únicas subcutáneamente administradas del polipéptido de aminoácidos no naturales terapéuticamente activo modificado serán seguras y bien toleradas por sujetos masculinos sanos. Basándose en una incidencia comparativa de acontecimientos adversos, valores de laboratorio clínico, constantes vitales y resultados de exploración física, los perfiles de seguridad del polipéptido de aminoácidos no naturales terapéuticamente activo modificado y el polipéptido de aminoácidos naturales terapéuticamente activo serán equivalentes. El polipéptido de aminoácidos no naturales terapéuticamente activo modificado proporciona potencialmente gran utilidad clínica a pacientes y personal sanitario.

LISTADO DE SECUENCIAS

ID SEC	Secuencia	Notas	Proteína, ácido nucleico, tRNA o RS
1	CCGGCGGTAGTTCAGCAGGGCAGAACGGCGGACTCTA AATCCGCATGGCGCTGGTTCAAATCCGGCCCGCGGAC CA	*M. jannaschii* $mtRNA^{Tyr}_{CUA}$	tRNA
2	CCCAGGGTAGCCAAGCTCGGCCAACGGCGACGGACTC TAAATCCGTTCTCGTAGGAGTTCGAGGGTTCGAATCCC TTCCC TGGGACCA	*HLAD03; un supresor ambar optimizado tRNA*	tRNA
3	GCGAGGGTAGCCAAGCTCGGCCAACGGCGACGGACTT CCTAATCCGTTCTCGTAGGAGTTCGAGGGTTCGAATCC CTCCCCTCGCACCA	*HL325A; un supresor AGGA de desplazamiento de marco optimizado tRNA*	tRNA
4	MDEFEMIKRNTSEIISEEELREVLKKDEKSAGIGFEPSGKIH LGHYLQIKKMIDLQNAGFDIIILLADLHAYLNQKGELDEIR KIGDYNKKVFEAMGLKAKYVYGSTFQLDKDYTLNVYRL ALKTTLKRARRSMELIARELLPKKVVCIHNPVLTGLD LGVDVAVGGMEQRKIHMLARELLPKKVVCIHNPVLTGLD GEGKMSSSKGNFIAVDDSPEEIRAKIKKAYCPAGVVEGNP IMEIAKYFLEYPLTIKRPEKFGGDLTVNSYEELESLFKNKE LHPMDLKNAVAEELIKILEPIRKRL	*p-Az-PheRS(6)*	RS
5	MDEFEMIKRNTSEIISEEELREVLKKDEKSAGIGFEPSGKIH LGHYLQIKKMIDLQNAGFDIIILLADLHAYLNQKGELDEIR KIGDYNKKVFEAMGLKAKYVYGSSFQLDKDYTLNVYRL ALKTTLKRARRSMELIARELLPKKVVCIHNPVLTGLD LGVDVAVGGMEQRKIHMLARELLPKKVVCIHNPVLTGLD GEGKMSSSKGNFIAVDDSPEEIRAKIKKAYCPAGVVEGNP IMEIAKYFLEYPLTIKRPEKFGGDLTVNSYEELESLFKNKE LHPMDLKNAVAEELIKILEPIRKRL	*Aminoacilo tRNA sintético para la incorporación de p-benzoilo-L-fenilalanina* *p-BpaRS(1)*	RS
6	MDEFEMIKRNTSEIISEEELREVLKKDEKAAIGFEPSGKIHL GHYLQIKKMIDL QNAGFDIIILLADLHAYLNQKGELDEIRKIGDYNKKVFEA MGLKAKYVYGSPFQLDKDYTLNVYRLALKTTLKRARRS MELIAREDENPKVAEVIYPIMQVNAIYLAVDVAVGGMEQ RKIHMLARELLPKKVVCIHNPVLTGLDGEGKMSSSKGNFI AVDDSPEEIRAKIKKAYCPAGVVEGNPIMEIAKYFLEYPLT IKRPEKFGGDLTVNSYEELESLFKNKELHPMDLKNAVAEE LIKILE PIRKR L	*Aminoacilo tRNA sintético para la incorporación de propargil-fenilalanina* *Propargyl-PheRS*	RS

(continuación)

SEC ID	Secuencia	Notas	Proteína, ácido nucleico, tRNA o RS
7	MDEFE MIKRN TSEII SEEEL REVLK KDEKS AAIGF EPSGK IHLGH YLQIK KMIDL QNAGF DIIIL LADLH AYLNQ KGELD EIRKI GDYNK KVFEA MGLKA KYVYG SPFQL DKDYT LNVYR LALKT TLKRA RRSME LIARE DENPK VAEVI YPIMQ VNIPY LPVD VAVGG MEQRK IHMLA RELLP KKVVC IHNPV LTGLD GEGKM SSSKG NFIAV DDSPE EIRAK IKKAY CPAGV VEGNP IMEIA KYFLE YPLTI KRPEK FGGDL TVNSY EELES LFKNK ELHPM DLKNA VAEEL IKILE PIRKR L	*Aminoacil tRNA sintético para la incorporación de propargil-fenilalanina* *Propargil-PheRS*	RS
8	MDEFE MIKRN TSEII SEEEL REVLK KDEKS AAIGF EPSGK IHLGH YLQIK KMIDL QNAGF DIIIL LADLH AYLNQ KGELD EIRKI GDYNK KVFEA MGLKA KYVYG SKFQL DKDYT LNVYR LALKT TLKRA RRSME LIARE DENPK VAEVI YPIMQ VNAIY LAVD VAVGG MEQRK IHMLA RELLP KKVVC IHNPV LTGLD GEGKM SSSKG NFIAV DDSPE EIRAK IKKAY CPAGV VEGNP IMEIA KYFLE YPLTI KRPEK FGGDL TVNSY EELES LFKNK ELHPM DLKNA VAEEL IKILE PIRKR L	*Aminoacil tRNA sintético para la incorporación de propargil-fenilalanina* *Propargil-PheRS*	RS
9	MDEFEMIKRNTSEIISEEELREVLKKDEKSATIGFEPSGKIH LGHYLQIKKMIDLQNAGFDIIILLADLHAYLNQKGELDEIR KIGDYNKKVFEAMGLKAKYVYGSNFQLDKDYTLNVYRL ALKTTLKRARRSMELIARERDENPKVAEVIYPIMQVNPLHY QGVDVAVGGMEQRKIHMLARELLPKKVVCIHNPVLTGL DGEGKMSSSKGNFIAVDDSPEEIRAKIKKAYCPAGVVEGN PIMEIAKYFLEYPLTIKRPEKFGGDLTVNSYEELESLFKNK ELHPMDLKNAVAEELIKILEPIRKRL	*Aminoacil tRNA sintético para la incorporación de p-acido-fenilalanina* *p-Az-PheRS(1)*	RS
10	MDEFEMIKRNTSEIISEEELREVLKKDEKSATIGFEPSGKIH LGHYLQIKKMIDLQNAGFDIIILLADLHAYLNQKGELDEIR KIGDYNKKVFEAMGLKAKYVYGSSFQLDKDYTLNVYRL ALKTTLKRARRSMELIARERDENPKVAEVIYPIMQVNPLHY QGVDVAVGGMEQRKIHMLARELLPKKVVCIHNPVLTGL DGEGKMSSSKGNFIAVDDSPEEIRAKIKKAYCPAGVVEGN PIMEIAKYFLEYPLTIKRPEKFGGDLTVNSYEELESLFKNK ELHPMDLKNAVAEELIKILEPIRKRL	*Aminoacil tRNA sintético para la incorporación de p-acido-fenilalanina* *p-Az-PheRS(3)*	RS

(continuación)

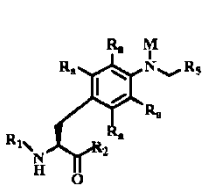
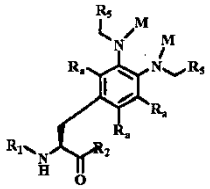
SEC ID	Secuencia	Notas	Protenia, acido nucleico, tRNA o RS
11	MDEFEMIKRNTSEIISEEELREVLKKDEKSALIGFEPSGKIH LGHYLQIKKMIDLQNAGFDIIILLADLHAYLNQKGELDEIR KIGDYNKKVFEAMGLKAKYVYGSTFQLDKDYTLNVYRL ALKTTLKRARRSMELIAREDENPKVAEVIYPIMQVNPVHY QGVDVAVGGMEQRKIHMLARELLPKKVVCIHNPVLTGL DGEGKMSSSKGNFIAVDDSPEEIRAKIKKAYCPAGVVEGN PIMEIAKYFLEYPLTIKRPEKFGGDLTVNSYEELESLFKNK ELHPMDLKNAVAEELIKILEPIRKRL	*Aminoacil tRNA sintético para la incorporación de p-acido-fenilalanina* *p-Az-PheRS(4)*	RS
12	MDEFEMIKRNTSEIISEEELREVLKKDEKSATIGFEPSGKIH LGHYLQIKKMIDLQNAGFDIIILLADLHAYLNQKGELDEIR KIGDYNKKVFEAMGLKAKYVYGSSFQLDKDYTLNVYRL ALKTTLKRARRSMELIAREDENPKVAEVIYPIMQVNPSHY QGVDVAVGGMEQRKIHMLARELLPKKVVCIHNPVLTGL DGEGKMSSSKGNFIAVDDSPEEIRAKIKKAYCPAGVVEGN PIMEIAKYFLEYPLTIKRPEKFGGDLTVNSYEELESLFKNK ELHPMDLKNAVAEELIKILEPIRKRL	*Aminoacil tRNA sintético para la incorporación de p-acido-fenilalanina* *p-Az-PheRS(2)*	RS
13	MDEFEMIKRNTSEIISEEELREVLKKDEKSALIGFEPSGKIH LGHYLQIKKMIDLQNAGFDIIILLADLHAYLNQKGELDEIR KIGDYNKKVFEAMGLKAKYVYGSEFQLDKDYTLNVYRL ALKTTLKRARRSMELIAREDENPKVAEVIYPIMQVNGCH YRGVDVAVGGMEQRKIHMLARELLPKKVVCIHNPVLTG LDGEGKMSSSKGNFIAVDDSPEEIRAKIKKAYCPAGVVEG NPIMEIAKYFLEYPLTIKRPEKFGGDLTVNSYEELESLFKN KELHPMDLKNAVAEELIKILEPIRKRL	*Aminoacil tRNA sintético para la incorporación de p-acido-fenilalanina (LW1)*	RS
14	MDEFEMIKRNTSEIISEEELREVLKKDEKSALIGFEPSGKIH LGHYLQIKKMIDLQNAGFDIIILLADLHAYLNQKGELDEIR KIGDYNKKVFEAMGLKAKYVYGSEFQLDKDYTLNVYRL ALKTTLKRARRSMELIAREDENPKVAEVIYPIMQVNGTHY RGVDVAVGGMEQRKIHMLARELLPKKVVCIHNPVLTGL DGEGKMSSSKGNFIAVDDSPEEIRAKIKKAYCPAGVVEGN PIMEIAKYFLEYPLTIKRPEKFGGDLTVNSYEELESLFKNK ELHPMDLKNAVAEELIKILEPIRKRL	*Aminoacil tRNA sintético para la incorporación de p-acido-fenilalanina (LW5)*	RS

(continuación)

SEC ID	Secuencia	Notas	Proteína, ácido nucleico, tRNA o RS
15	MDEFEMIKRNTSEIISEEELREVLKKDEKSAAIGFEPSGKIH LGHYLQIKKMIDLQNAGFDIIILLADLHAYLNQKGELDEIR KIGDYNKKVFEAMGLKAKYVYGSEFQLDKDYTLNVYRL ALKTTLKRARRSMELIAREDENPKVAEVIYPIMQVNGGH YLGVDVIVGGMEQRKIHMLARELLPKKVVCIHNPVLTGL DGEGKMSSSKGNFIAVDDSPEEIRAKIKKAYCPAGVVEGN PIMEIAKYFLEYPLTIKRPEKFGGDLTVNSYEELESLFKNK ELHPMDLKNAVAEELIKILEPIRKRL	*Aminoacil tRNA sintético para la incorporación de p-acido-fenilalanina (LW6)*	RS
16	MDEFEMIKRNTSEIISEEELREVLKKDEKSAAIGFEPSGKIH LGHYLQIKKMIDLQNAGFDIIILLADLHAYLNQKGELDEIR KIGDYNKKVFEAMGLKAKYVYGSRFQLDKDYTLNVYRL ALKTTLKRARRSMELIAREDENPKVAEVIYPIMQVNVIHY DGVDVAVGGMEQRKIHMLARELLPKKVVCIHNPVLTGL DGEGKMSSSKGNFIAVDDSPEEIRAKIKKAYCPAGVVEGN PIMEIAKYFLEYPLTIKRPEKFGGDLTVNSYEELESLFKNK ELHPMDLKNAVAEELIKILEPIRKRL	*Aminoacil tRNA* *sintético para la incorporación de p-acido-fenilalanina (AzPheRS-5)*	RS
17	MDEFEMIKRNTSEIISEEELREVLKKDEKSAGIGFEPSGKIH LGHYLQIKKMIDLQNAGFDIIILLADLHAYLNQKGELDEIR KIGDYNKKVFEAMGLKAKYVYGSTFQLDKDYTLNVYRL ALKTTLKRARRSMELIAREDENPKVAEVIYPIMQVNTYYY LGVDVAVGGMEQRKIHMLARELLPKKVVCIHNPVLTGLD GEGKMSSSKGNFIAVDDSPEEIRAKIKKAYCPAGVVEGNP IMEIAKYFLEYPLTIKRPEKFGGDLTVNSYEELESLFKNKE LHPMDLKNAVAEELIKILEPIRKRL	*Aminoacil tRNA sintético para la incorporación de p-acido-fenilalanina (AzPheRS-6)*	RS

REIVINDICACIONES

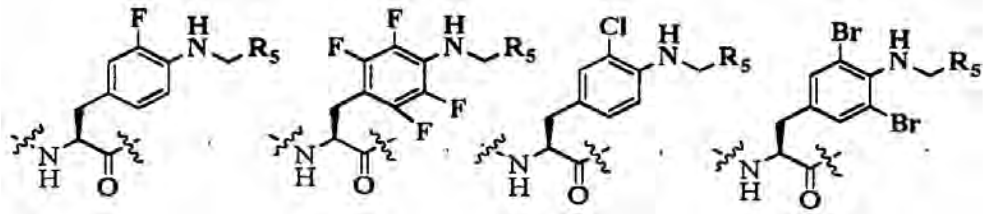
1. Un polipéptido o sal del mismo, que contiene al menos un compuesto seleccionado del grupo que consiste en:

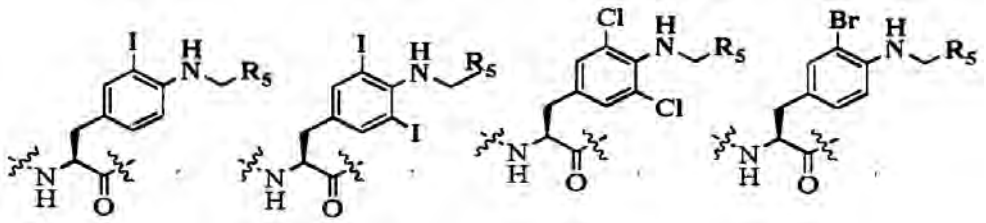

,

,

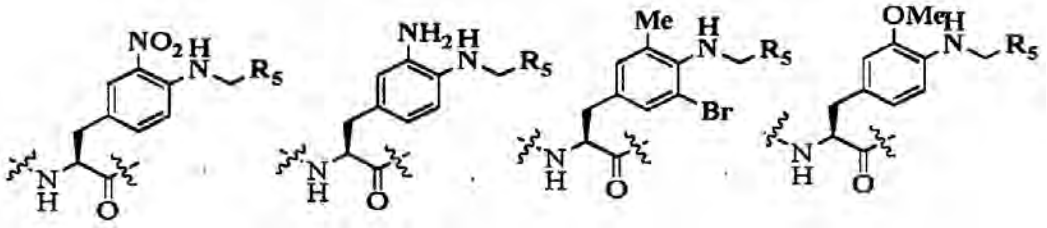
en el que cada R_a está seleccionado independientemente del grupo que consiste en H, halógeno, alquilo, $-NO_2$, -CN, alquilo sustituido, $-N(R')_2$, $-C(O)_kR'$, $-C(O)N(R')_2$, -OR' y $-S(O)_kR'$ en la que k es 1, 2 ó 3;

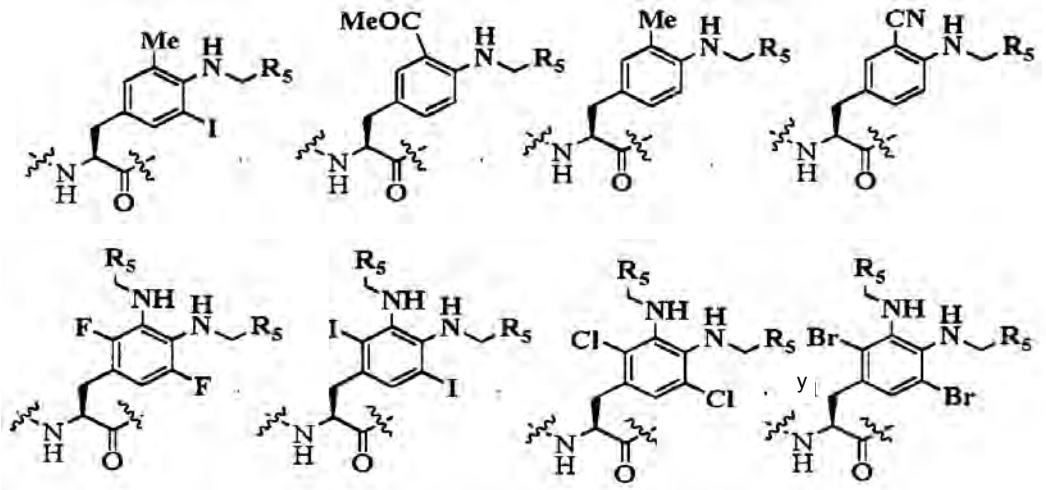
M es H o $-CH_2R_5$;
R_1 es H, un grupo protector de amino, resina, aminoácido, polipéptido o polinucleótido;
R_2 es OH, un grupo protector de éster, resina, aminoácido, polipéptido o polinucleótido;
R_5 es L-X, en la que X es un polímero que comprende alquilo, alquenilo, alquinilo, alcoxi, alquilalcoxi, poli(óxido de alquileno), arilo, heteroarilo, alcarilo o aralquilo; y L es opcional, y cuando está presente es un conector seleccionado del grupo que consiste en alquileno, alquenileno, -O-, -O-(alquilen)-, -(alquilen)-O-, -C(O)-, -C(O)-(alquilen)-, -(alquilen)-C(O)-, -C(O)N(R')-, -C(O)N(R')-(alquilen)-, -(alquilen)-C(O)N(R'), -OC(O)N(R')-, -OC(O)N(R')-(alquilen)-, -(alquilen)-OC(O)N(R')-, -N(R')C(O)-, -NR'C(O)-(alquilen)-, -(alquilen)-NR'C(O)-, -S-, -S-(alquilen)-, $-S(O)_k$- en la que k es 1, 2 ó 3, $-S(O)_k$(alquilen)-, -C(S)-, -C(S)-(alquilen)-, -CSN(R')-, -CSN(R')-(alquilen)-, -N(R')CO-(alquilen)-, -N(R')C(O)O-, -(alquilen)-ON=CR'-, -(alquilen)-C(O)NR'-(alquilen)-, -(alquilen)-$S(O)_k$-(alquilen o alquilen sustituido)-S-, -(alquilen)-S-S-, $-S(O)_kN(R')$-, -N(R')C(O)N(R')-, -N(R')C(S)N(R')-, -N(R')$S(O)_k$N(R')-, -N(R')-N=, -C(R')=N-, -C(R')=N-N(R')-, -C(R')=N-N=, $-C(R')_2$-N=N- y $-C(R')_2$-N(R')-N(R')-; y
cada R' es independientemente H, alquilo o alquilo sustituido;

en el que el compuesto está seleccionado del grupo:







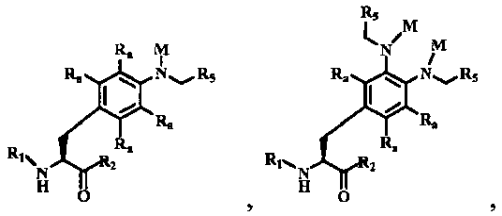


2. El compuesto de la reivindicación 1, en el que el polímero comprende poli(óxido de alquileno).

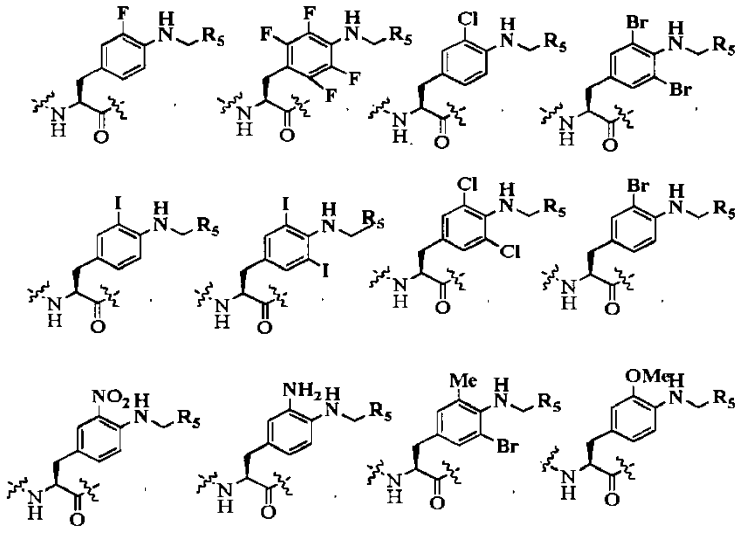
3. El compuesto de la reivindicación 1, en el que el polímero comprende -[(alquilen)-O-(hidrógeno o alquilo)]$_x$, en el que x es de 20-10.000.

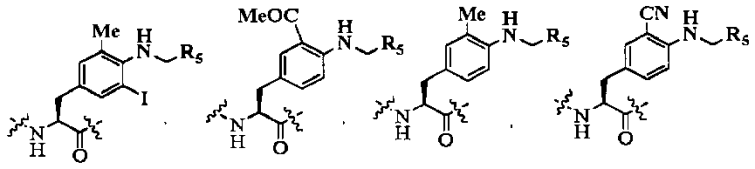
4. El compuesto de la reivindicación 1, en el que el polímero es m-PEG que tiene un peso molecular que oscila de aproximadamente 2 a aproximadamente 40 KDa.

5. Un procedimiento de preparación de un compuesto que contiene al menos un aminoácido no natural seleccionado del grupo que consiste en:



en el que el compuesto se forma por una alquilación reductora de un resto amino aromático sobre al menos un aminoácido no natural con al menos un reactante que comprende al menos un resto de aldehído, en el que R_1, R_2, R_5, R_a y M son como se han definido en la reivindicación 1;
en el que el al menos un aminoácido no natural está seleccionado de un grupo:



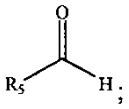


6. El procedimiento de la reivindicación 5, en el que el polímero comprende poli(óxido de alquileno) o un poli(óxido de alquileno) sustituido.

7. El procedimiento de la reivindicación 5, en el que el polímero comprende -[(alquilen o alquilen sustituido)-O-(hidrógeno, alquilo o alquilo sustituido)]$_x$, en el que x es de 20-10.000.

8. El procedimiento de la reivindicación 5, en el que el polímero es m-PEG que tiene un peso molecular que oscila de aproximadamente 2 a aproximadamente 40 KDa.

9. El procedimiento de las reivindicaciones 5, en el que el al menos un resto de aldehído tiene la estructura correspondiente a:



en la que:

R_5 es como se ha definido en la reivindicación 1.

10. El procedimiento de la reivindicación 5, en el que el aminoácido no natural se incorpora usando un sistema de traducción.

11. El procedimiento de la reivindicación 5, en el que el aminoácido no natural se incorpora en el polipéptido usando un sistema de traducción y un sistema de modificación postraduccional.

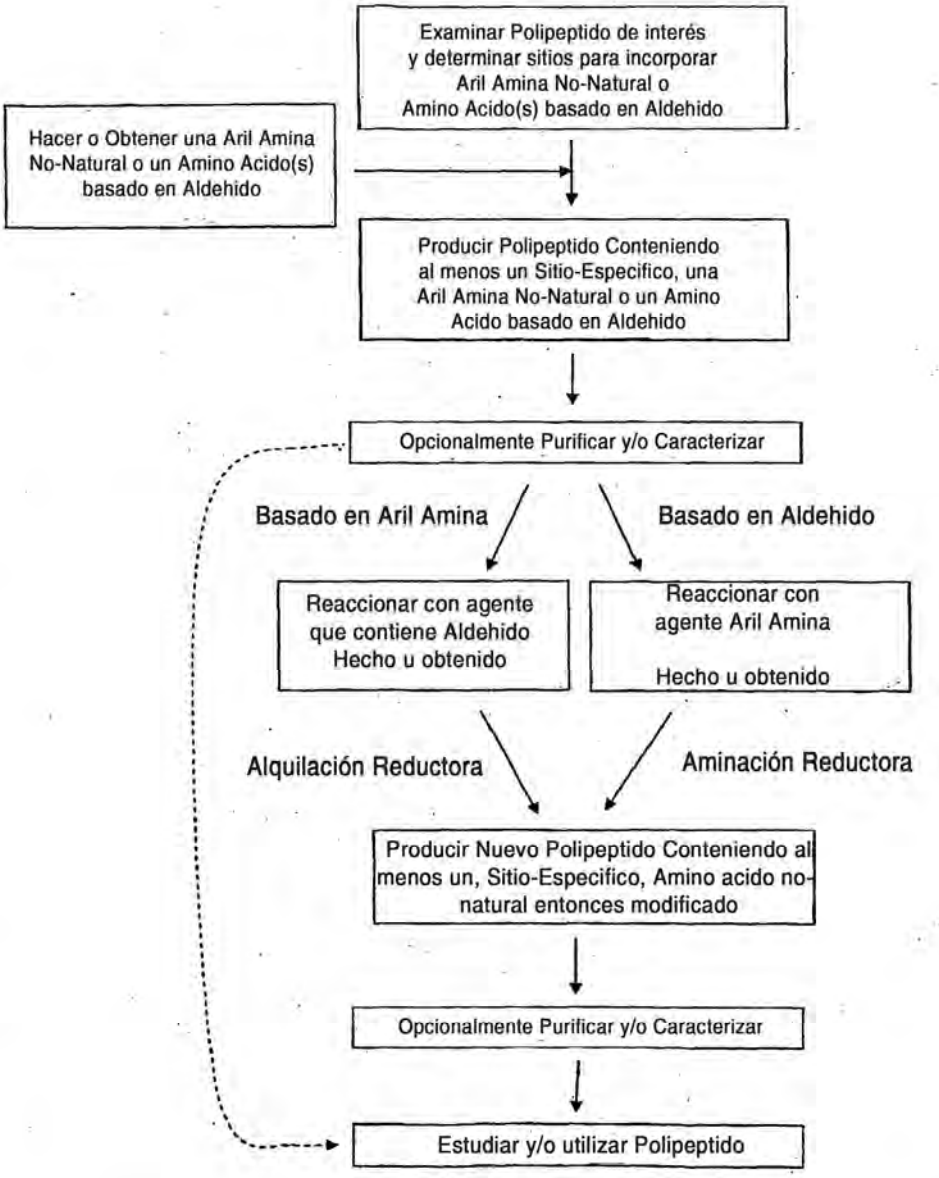

Examinar Polipeptido de interés y determinar sitios para incorporar Aril Amina No-Natural o Amino Acido(s) basado en Aldehido
Hacer o Obtener una Aril Amina No-Natural o un Amino Acido(s) basado en Aldehido
Producir Polipeptido Conteniendo al menos un Sitio-Especifico, una Aril Amina No-Natural o un Amino Acido basado en Aldehido
Opcionalmente Purificar y/o Caracterizar
Basado en Aril Amina
Basado en Aldehido
Reaccionar con agente que contiene Aldehido Hecho u obtenido
Reaccionar con agente Aril Amina Hecho u obtenido
Alquilación Reductora
Aminación Reductora
Producir Nuevo Polipeptido Conteniendo al menos un, Sitio-Especifico, Amino acido no-natural entonces modificado
Opcionalmente Purificar y/o Caracterizar
Estudiar y/o utilizar Polipeptido

Figura 1

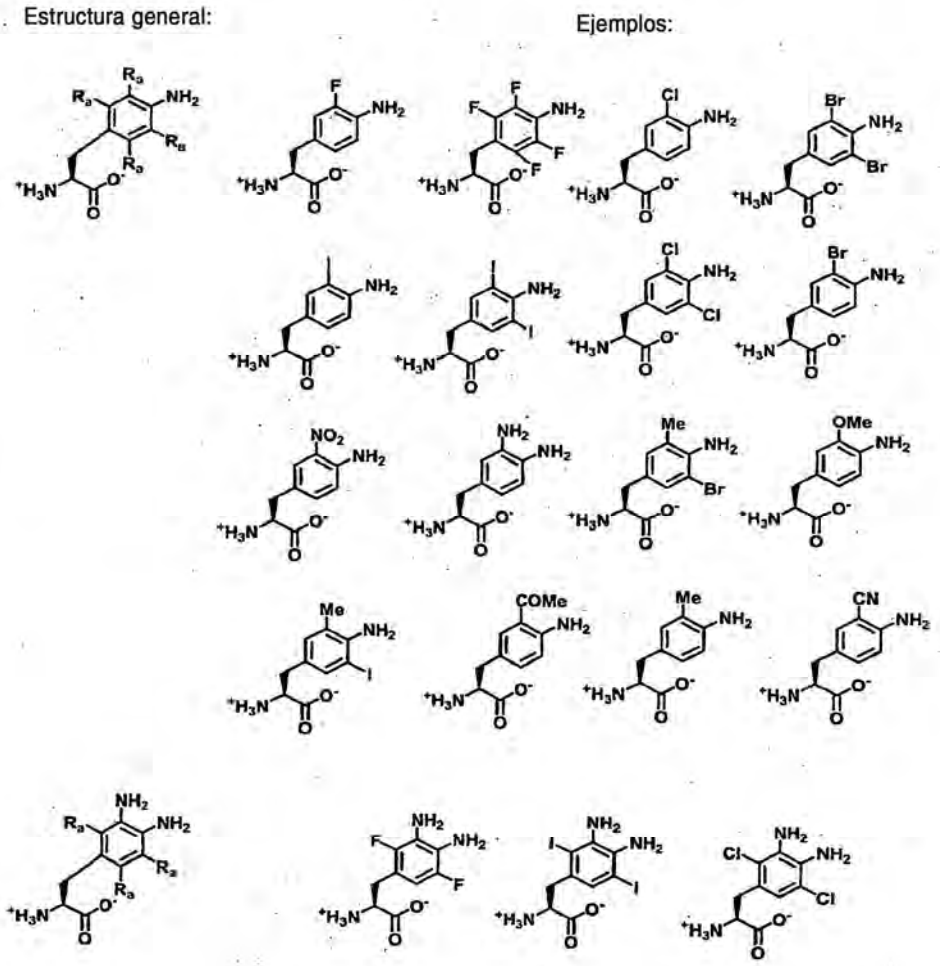


Figura 2

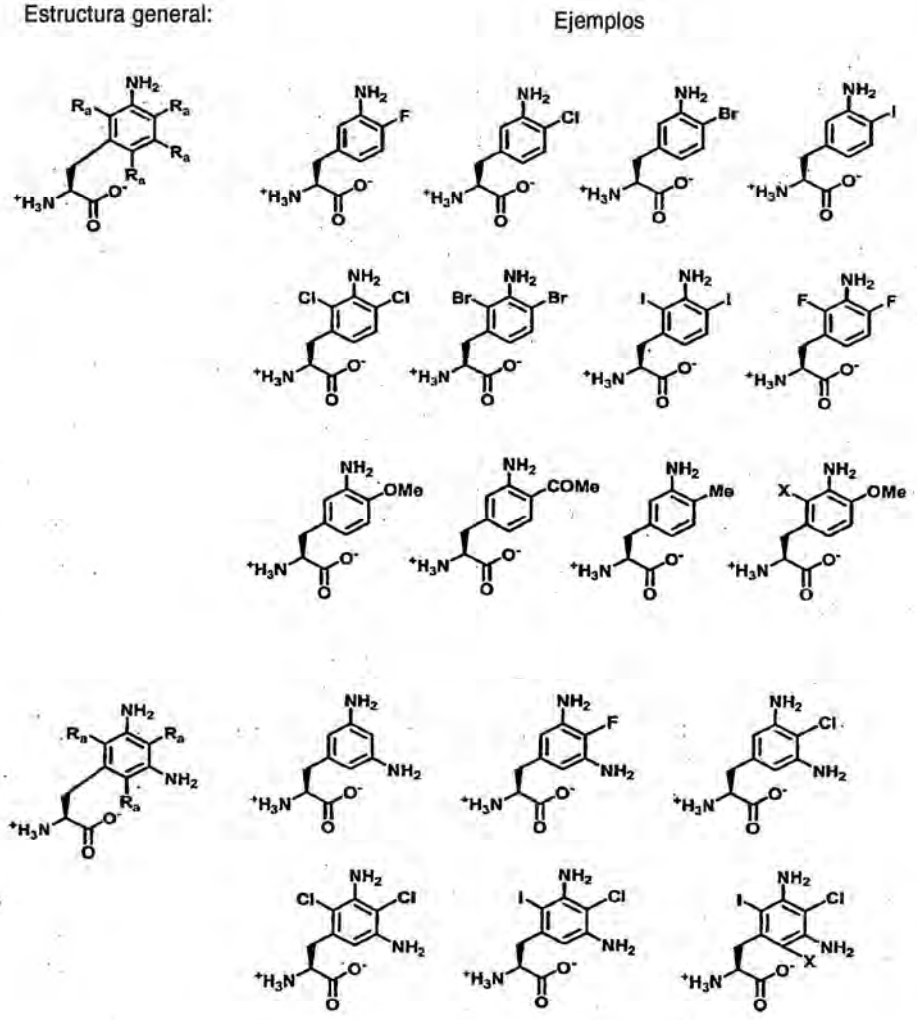


Figura 3

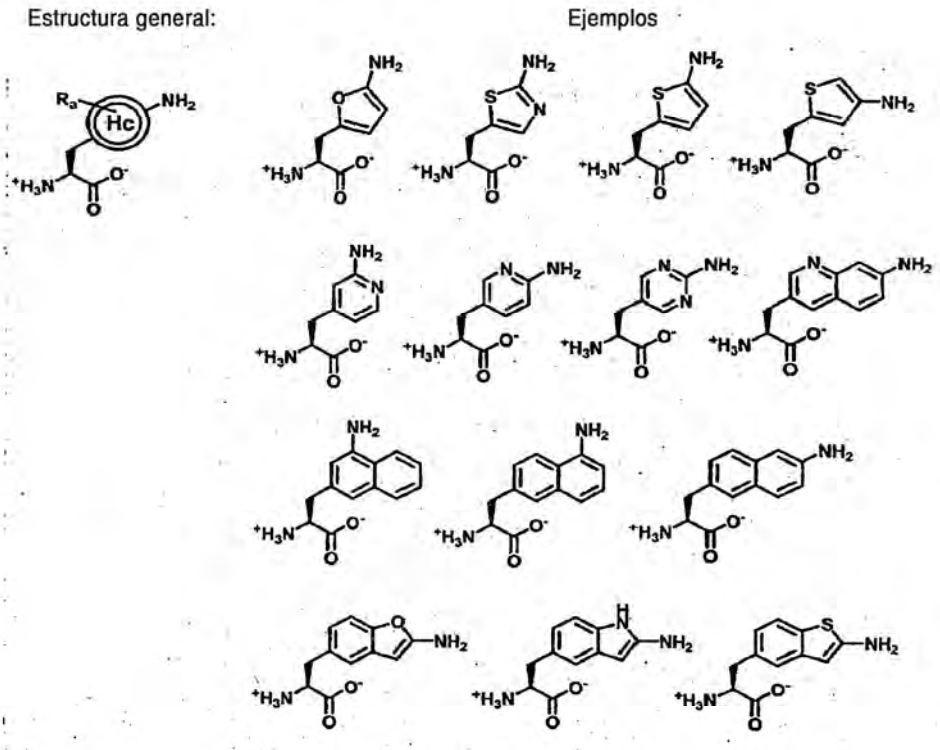


Figura 4

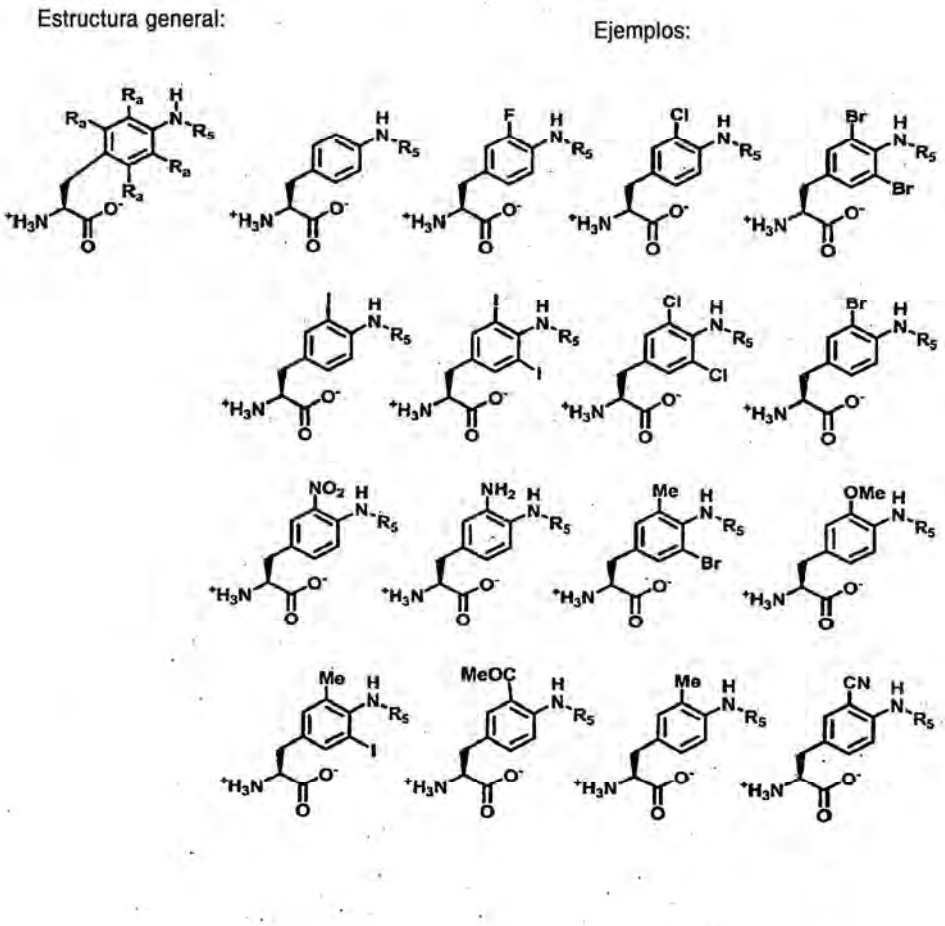


Figura 5

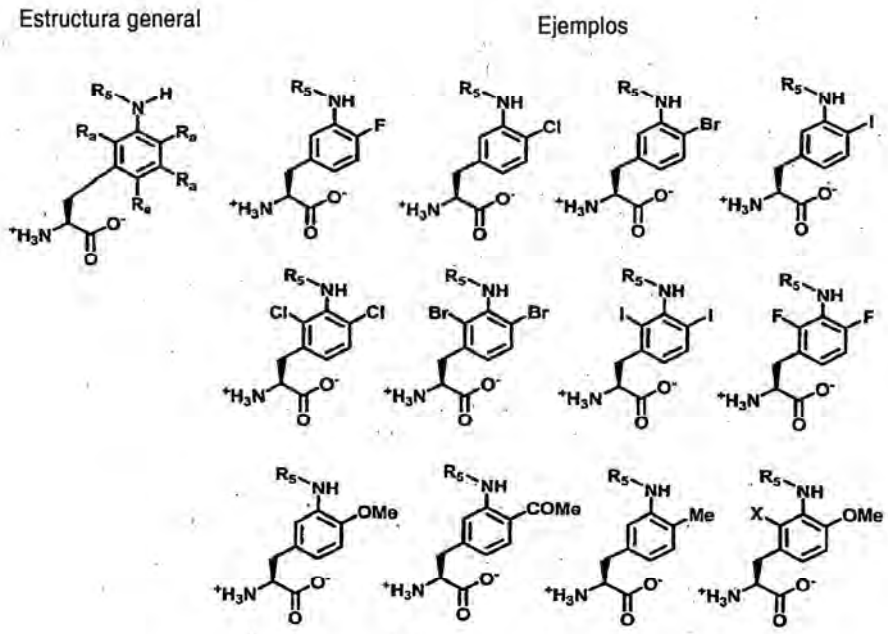


Figura 6

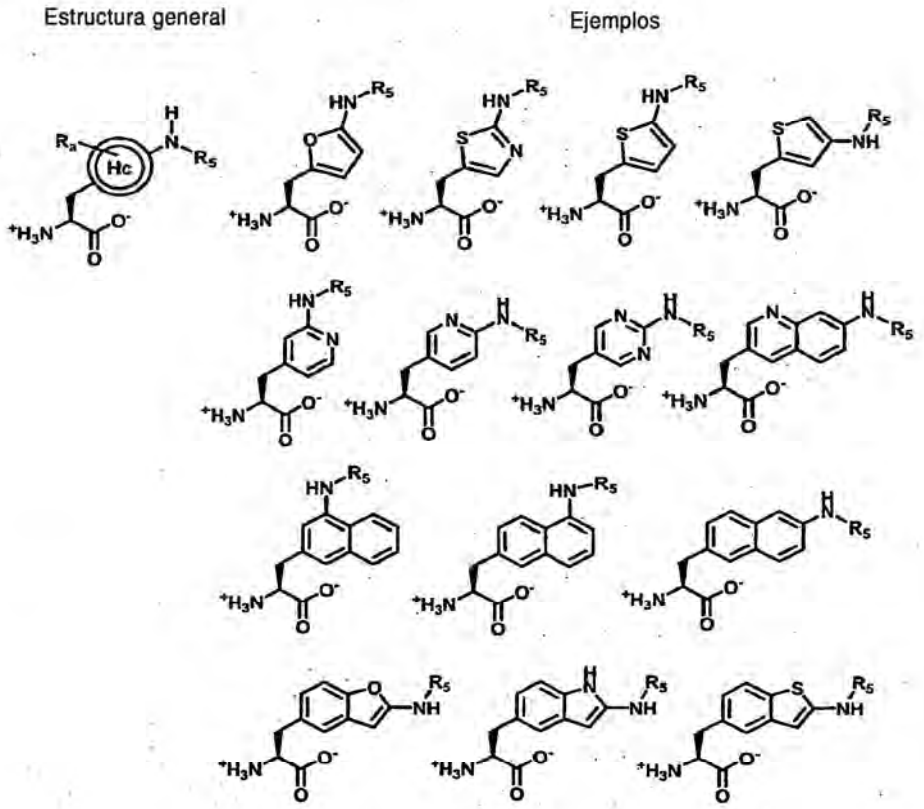


Figura 7

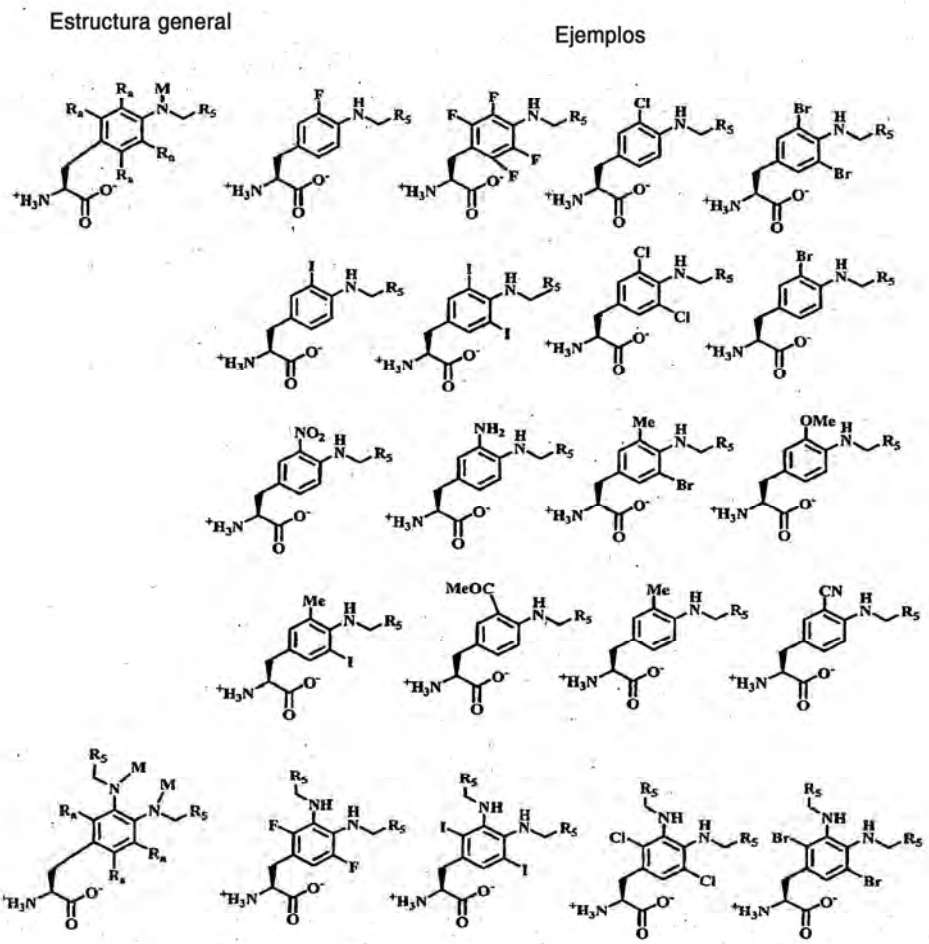


Figura 8

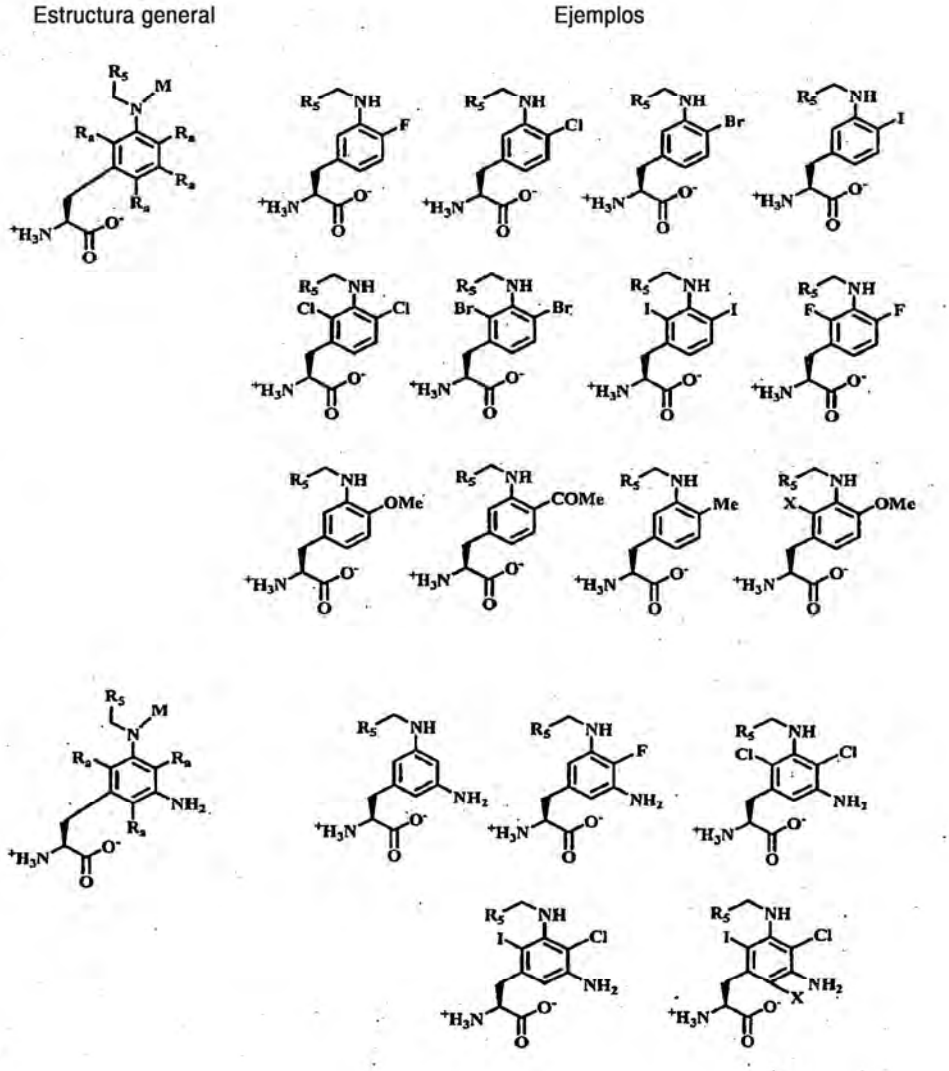


Figura 9

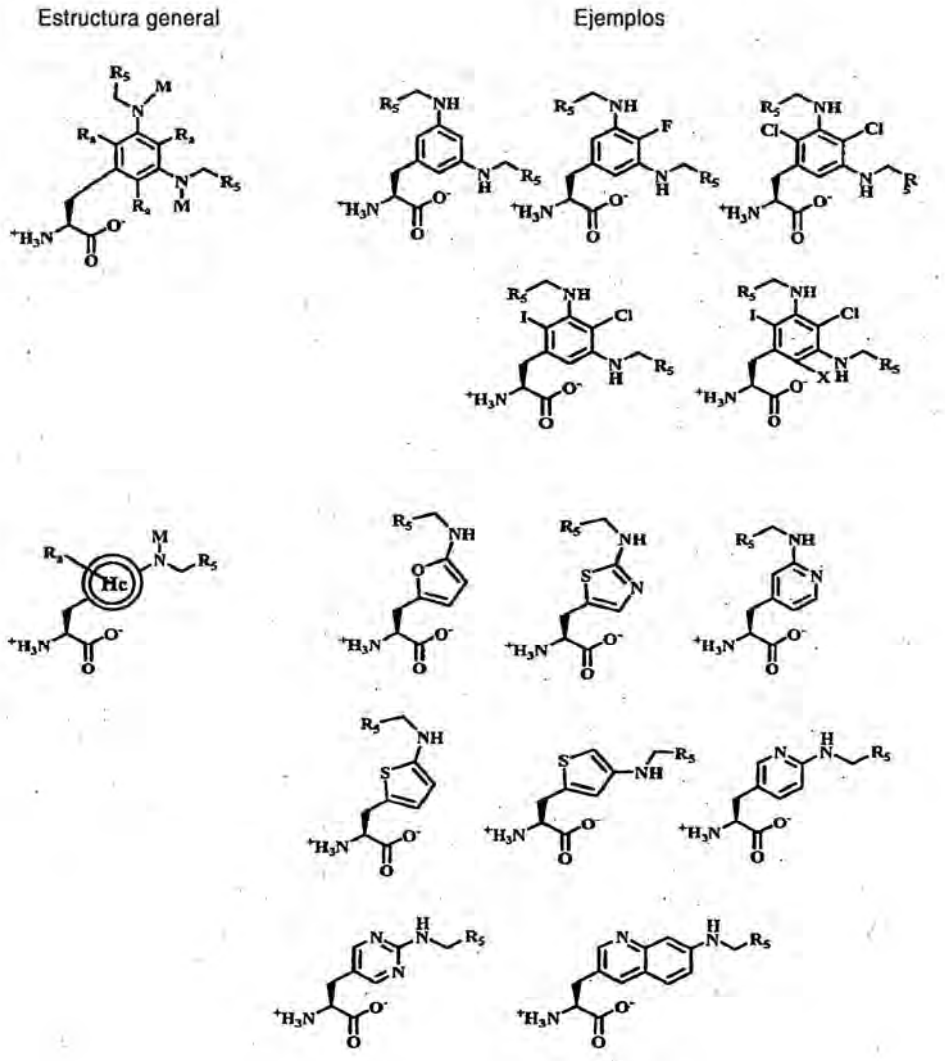
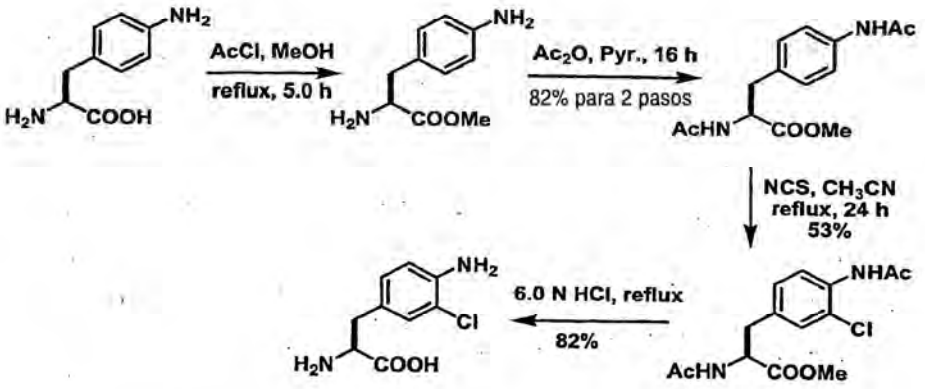


Figura 10

A. Sintesis de 3-cloro-4-amino-fenilalanina



B. Sintesis de 3-iodo-4-amino-fenilalanina

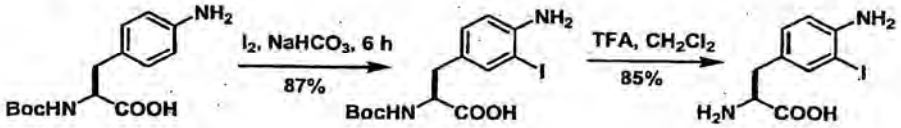
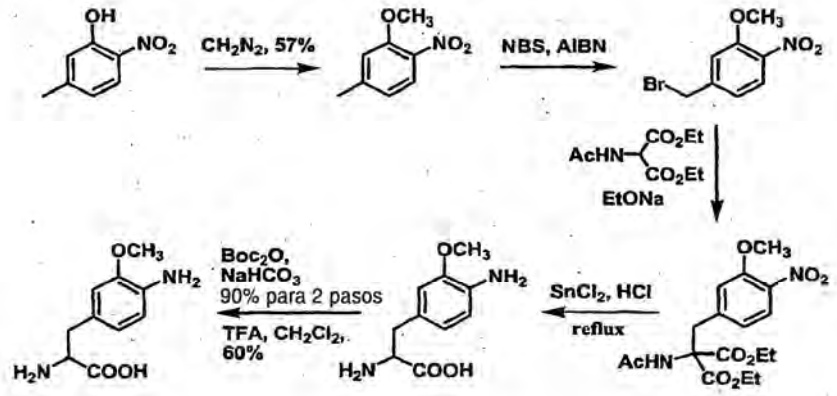


Figura 11

A. Sintesis de 3-metoxi-4-amino-fenilalanina



B. Sintesis de 3-fluoro-4-amino-fenilalanina

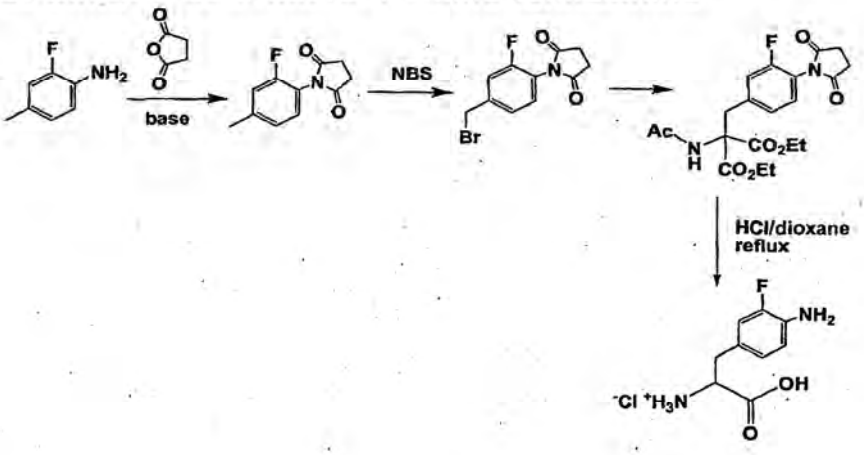
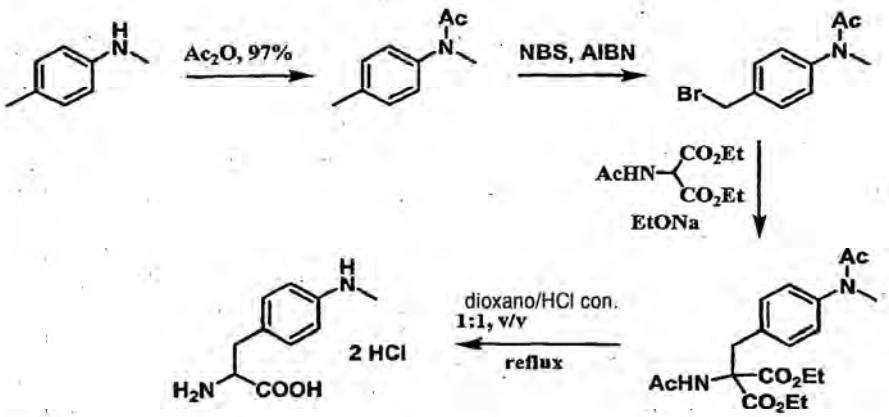


Figura 12

A. Sintesis de N-metil-p-amino-fenilalanina



B. Sintesis de N-etil-p-amino-fenilalanina

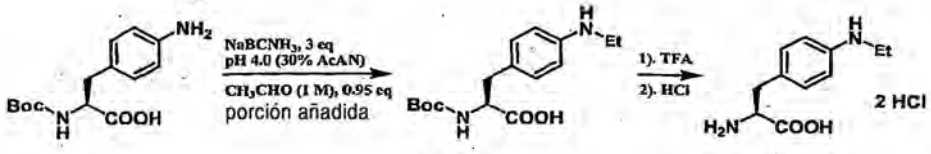


Figura 13

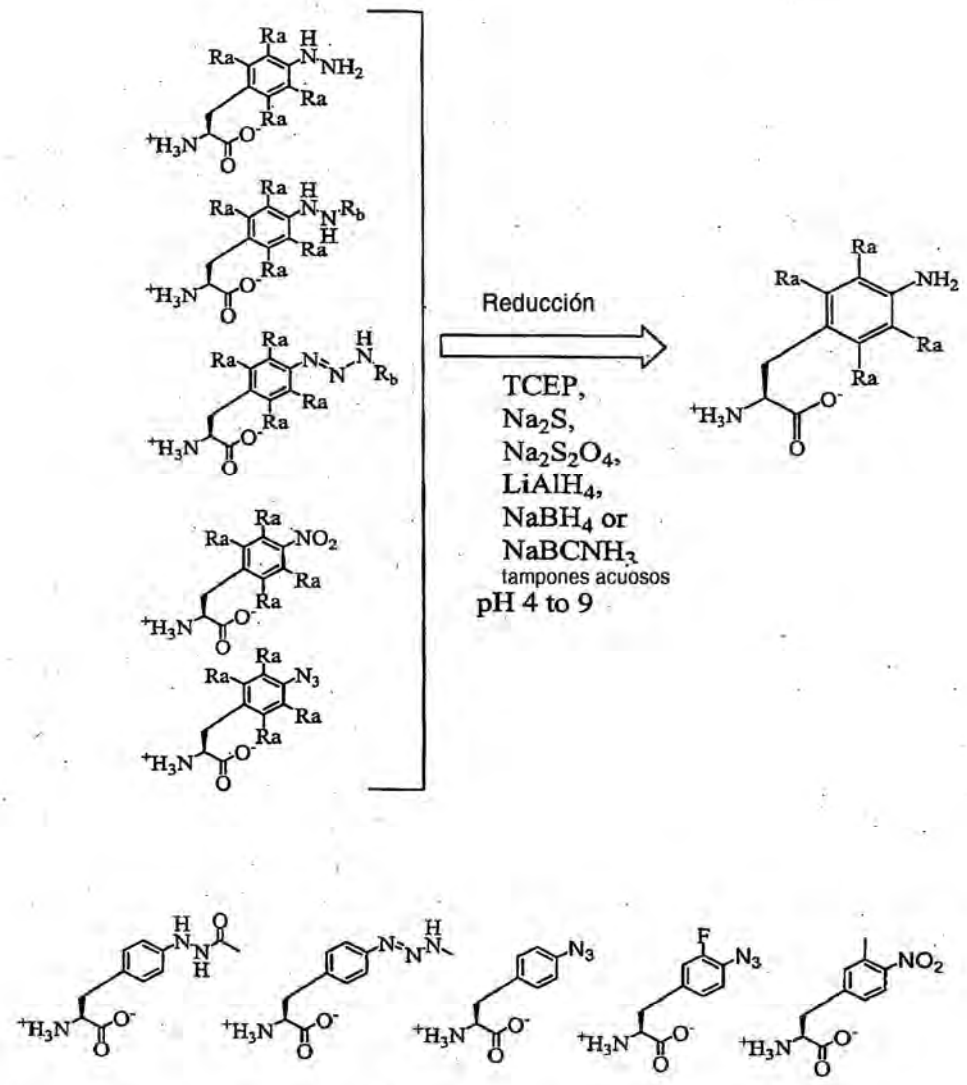

Reducción
TCEP,
Na_2S,
$Na_2S_2O_4$,
$LiAlH_4$,
$NaBH_4$ or
$NaBCNH_3$
tampones acuosos
pH 4 to 9

Figura 14

Formación de Polipeptidos con Amina Aromatica Conteniendo Amino Acidos no-naturales por reducción Post-translacional de los precursores enmascarados, seguidos por una reducccion de Alquilaciones con Reagentes que contienen Aldehido

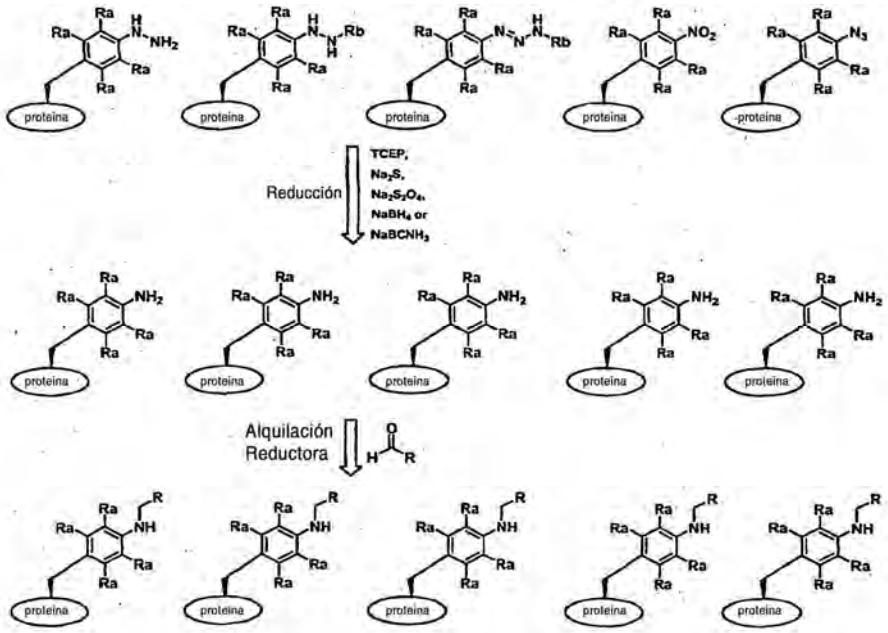


Figura 15

Alquilaciones Reductoras de Aminas Aromaticas Mono-alquiladas que contienen Amino Acido Peptidos no-naturales

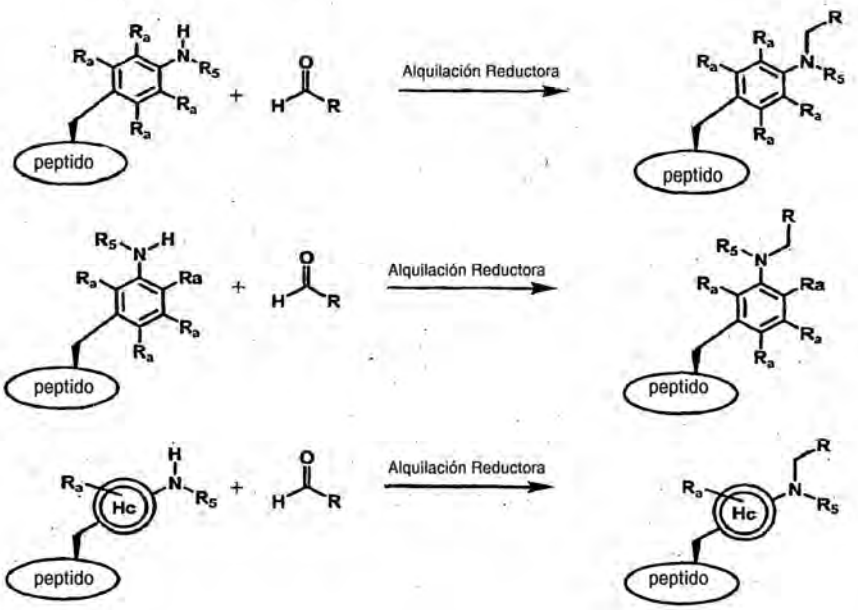


Figura 16

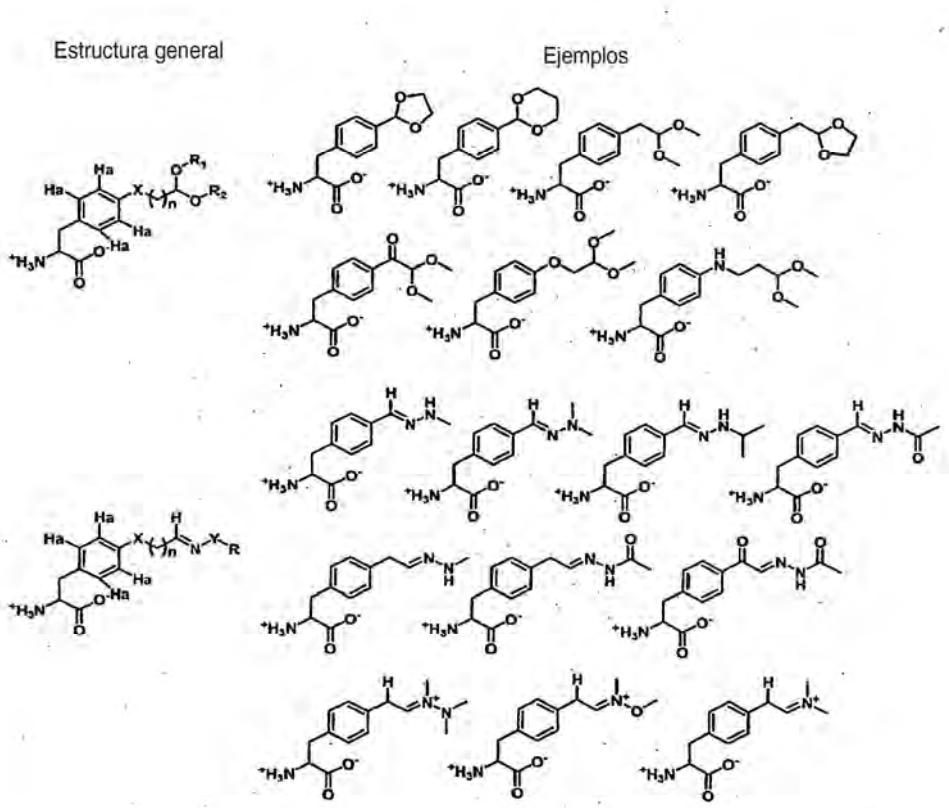


Figura 17

Formación de Polipeptidos con Aldehidos conteniendo Amino acidos no-naturales por desprotecciones post-translacionales de precursores de Aldehido enmascarados, seguidos por Aminaciones Reductoras con reagentes contenedores de aminas aromaticas

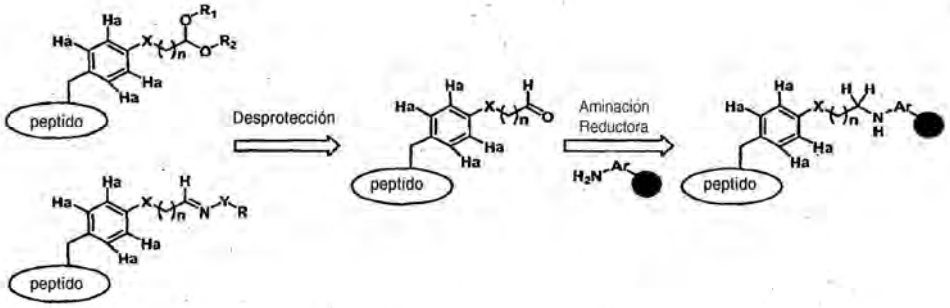


Figura 18

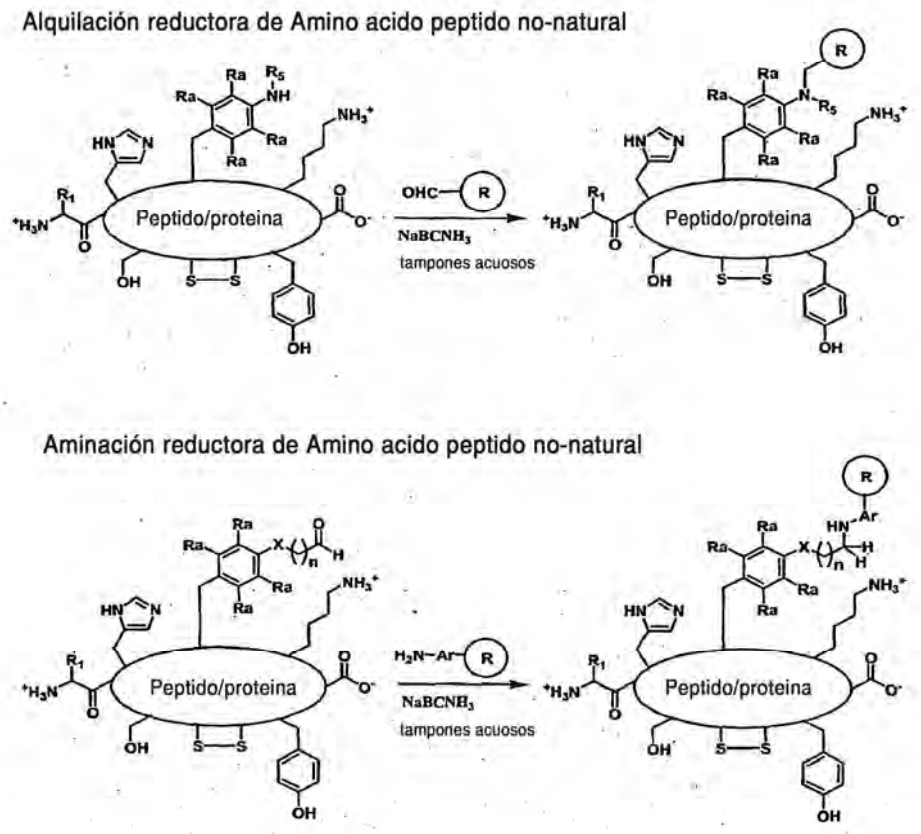
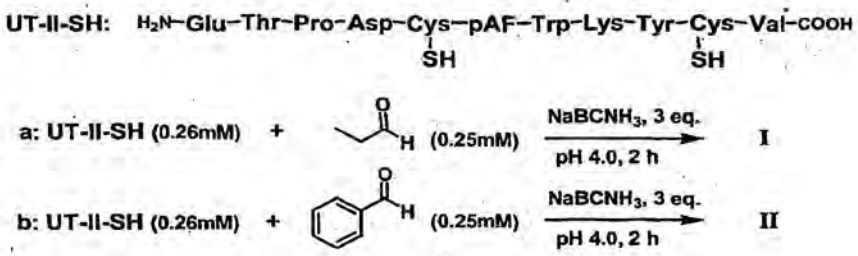

Alquilación reductora de Amino acido peptido no-natural
Peptido/proteina
OHC
R
NaBCNH3
tampones acuosos
Peptido/proteina
Aminación reductora de Amino acido peptido no-natural
Peptido/proteina
H2N-Ar
R
NaBCNH3
tampones acuosos
Peptido/proteina

Figura 19

Alquilaciobes reductoras de urotension-II (UT-II-SH) reducidas
con propionaldehido (a) y benzaldehido (b)



pAF= L-para-Amino-fenilalanina

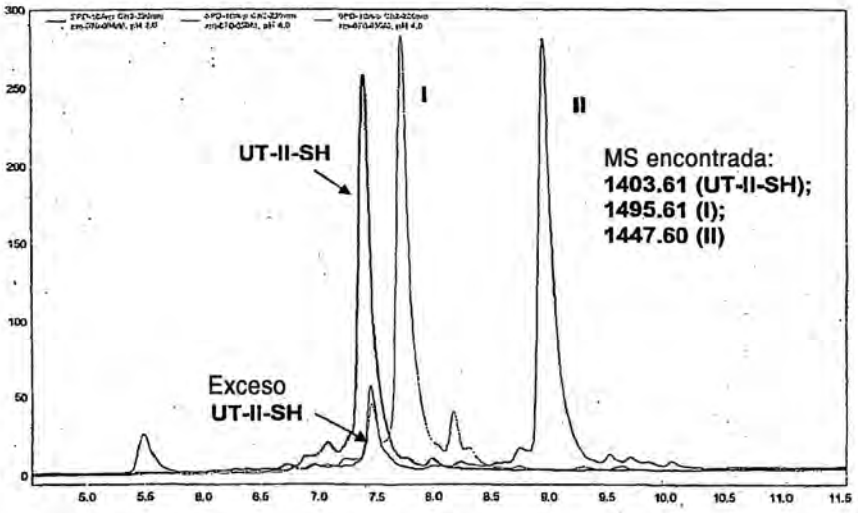
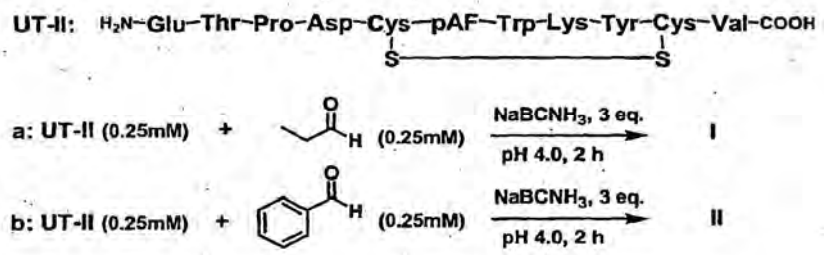


Figura 20

Alquilaciones Reductoras de urotension-II (UT-II) con propionaldehido (a) y benzonaldehido (b), respectivamente.



pAF = L-para-Amino-fenilalanina

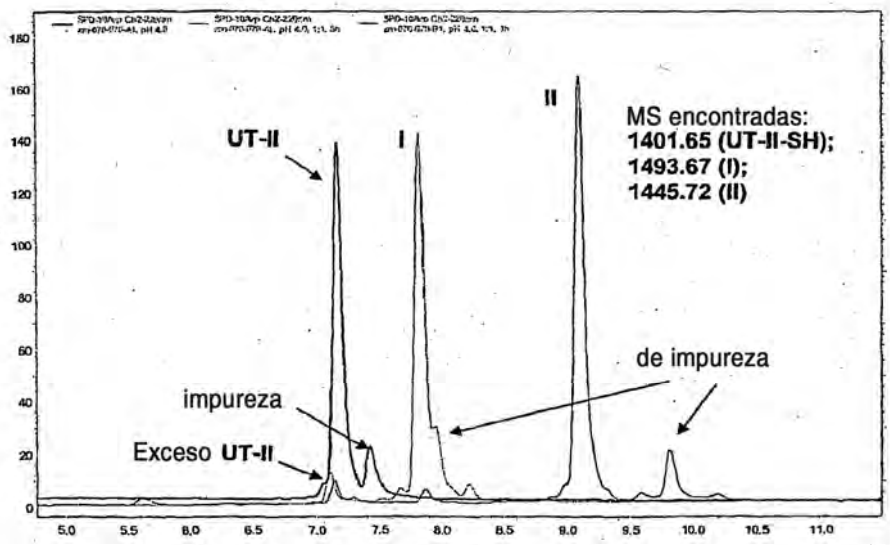
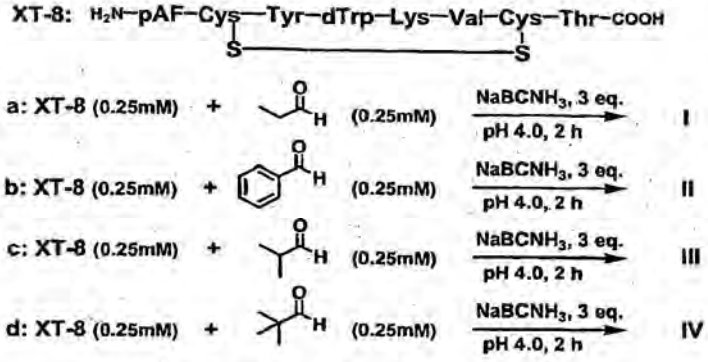


Figura 21

Alquilaciones reductoras de peptido XT-8 con propionaldehido (a) y benzaldehido (b), isobutaldehido (c) y pivalaldehido (d)



pAF= L-Amino-fenilalanina

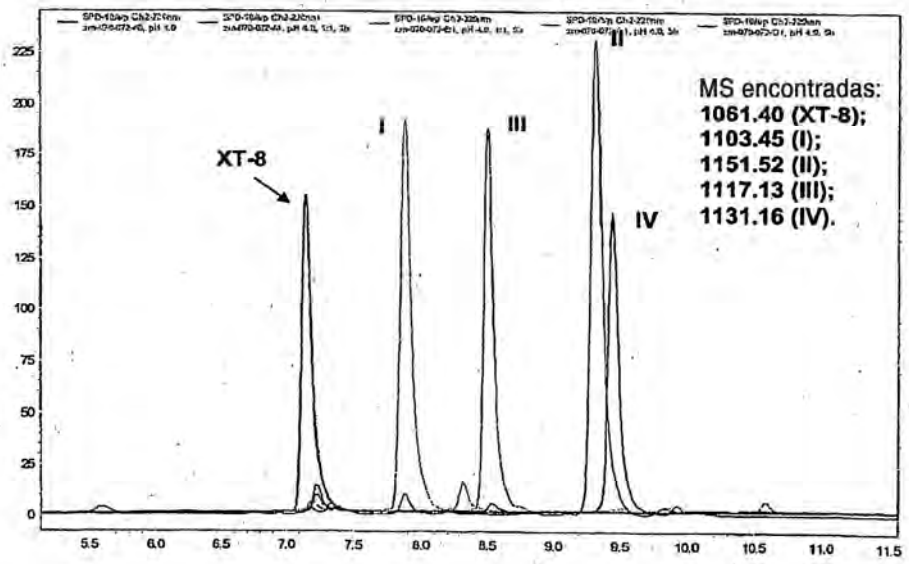
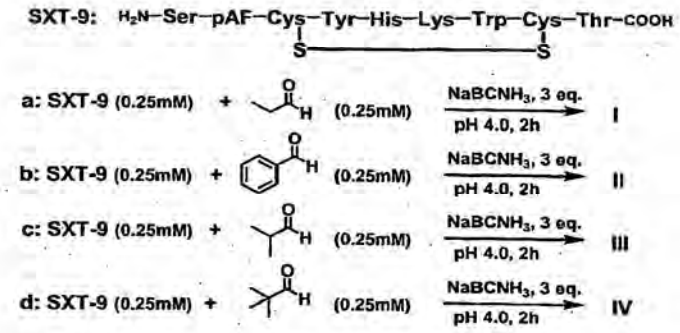


Figura 22

Alquilaciones reductoras de peptido SXT-9 con propionaldehido (a) y benzaldehido (b), isobutaldehido (c) y pivalaldehido (d)



pAF = L-para-Amino-fenilalanina

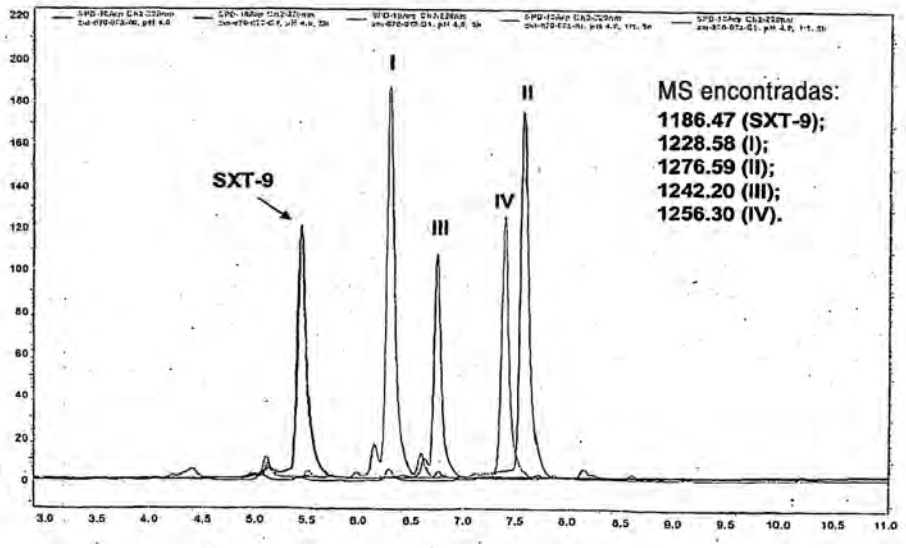
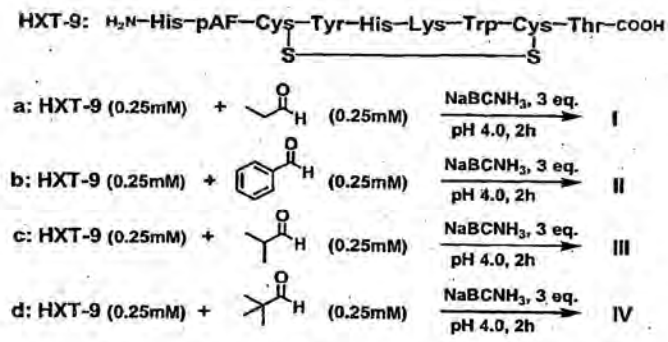


Figura 23

Alquilaciones reductoras de peptido HXT-9 con propionaldehido (a) y benzaldehido (b), isobutaldehido (c) y pivalaldehido (d)



pAF= L-para-Amino-fenilalanina

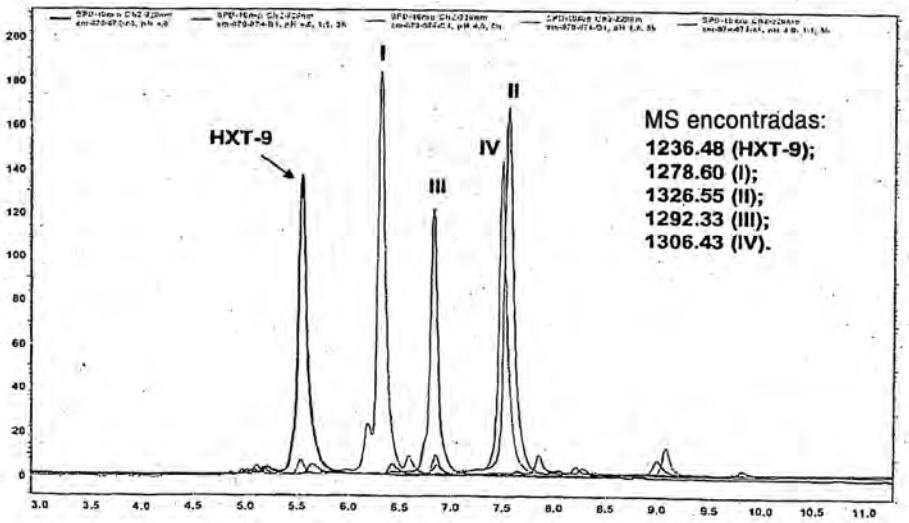
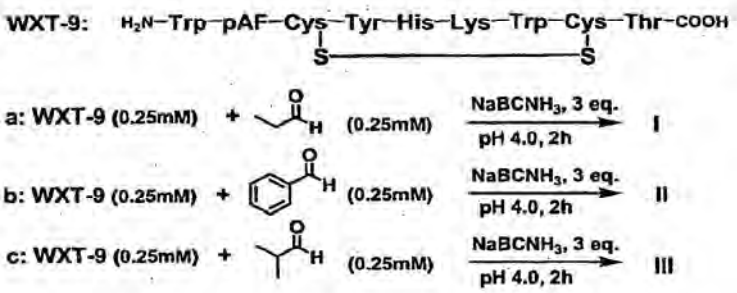


Figura 24

Alquilaciones reductoras de peptido WXT-9 con propionaldehido (a), benzaldehido (b), y isobutaldehido (c)



pAF= L-para-Amino-fenilalanina

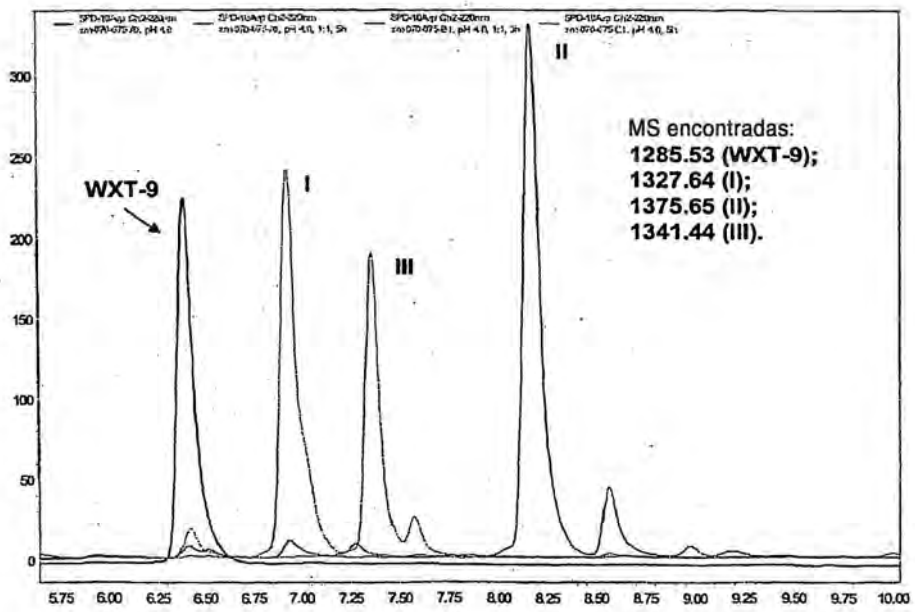
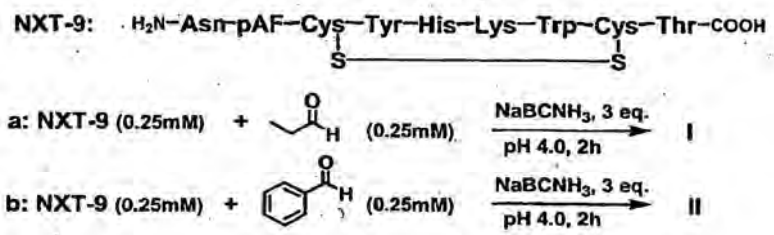


Figura 25

Alquilaciones reductoras de peptido NXT-9 con propionaldehido (a) y benzaldehido (b)



pAF = L-para-Amino-fenilalanina

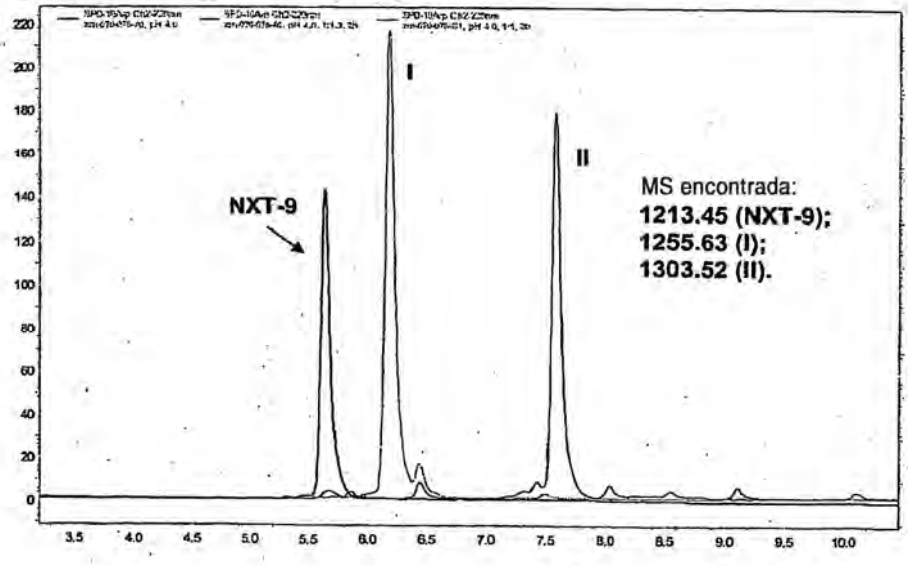
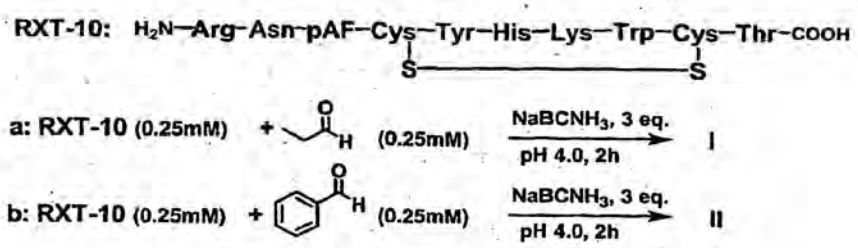


Figura 26

Alquilaciones reductoras de peptido RXT-10 con propionaldehido (a) y benzaldehido (b)



pAF = L-para-Amino-fenilalanina

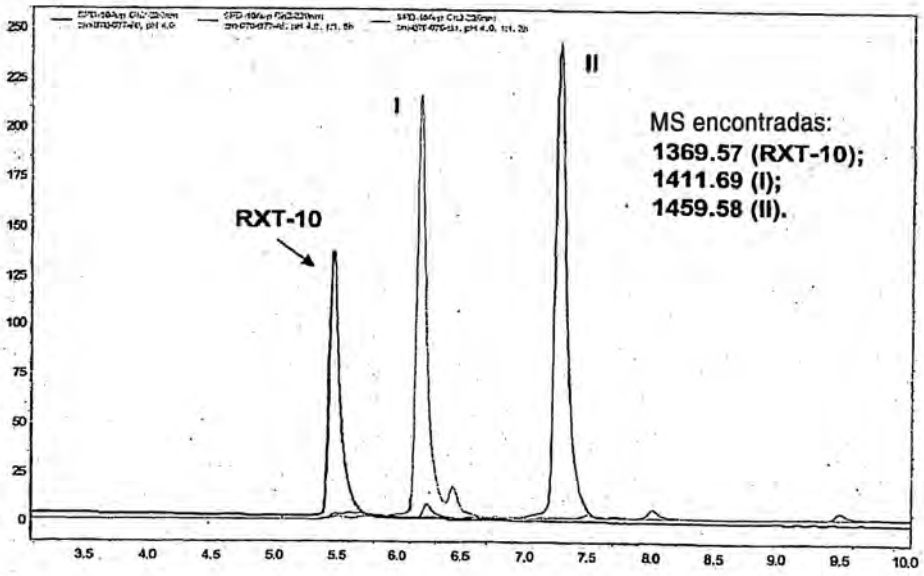
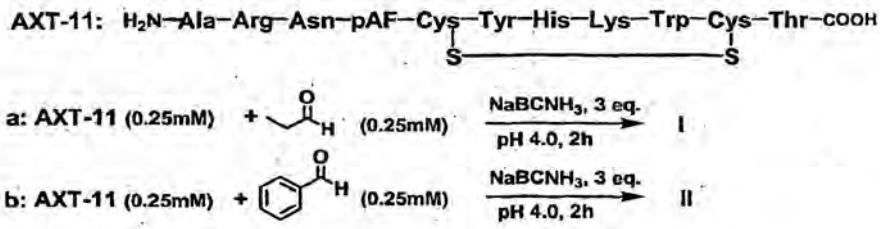


Figura 27

Alquilaciones reductoras de peptidos AXT-11 con propionaldehido (a) y benzaladehido (b)



pAF = L-para-Amino-fenilalanina

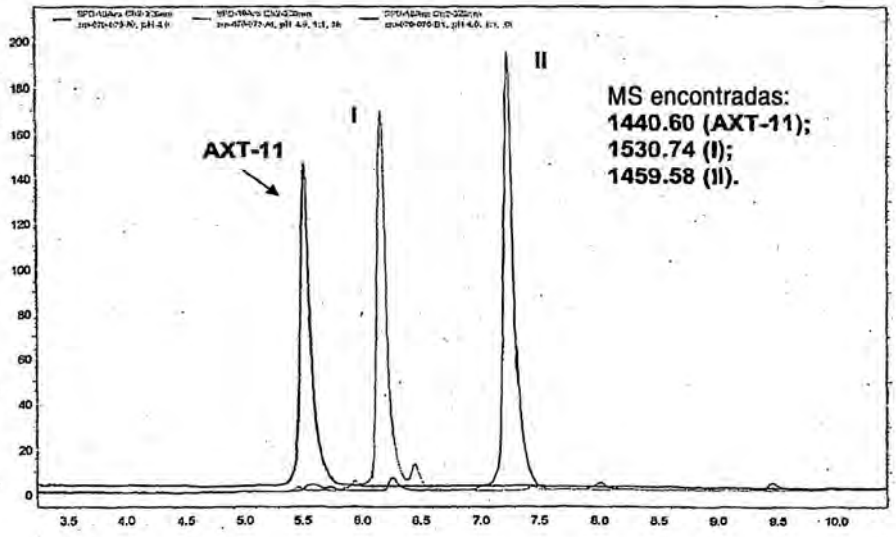
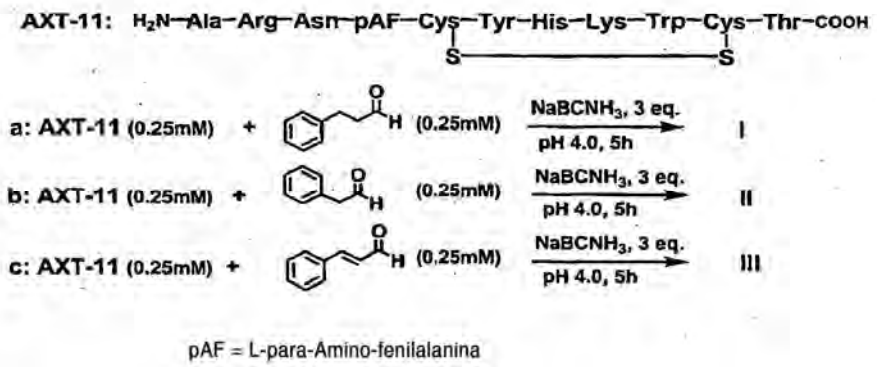


Figura 28

Alquilaciones reductoras de peptido AXT-11 con varios aldehidos (a-c)



pAF = L-para-Amino-fenilalanina

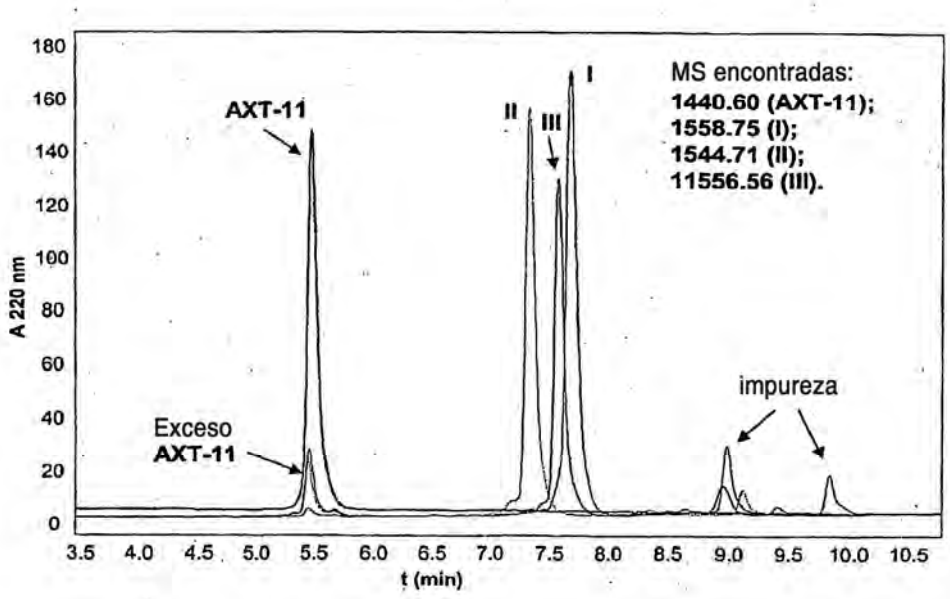
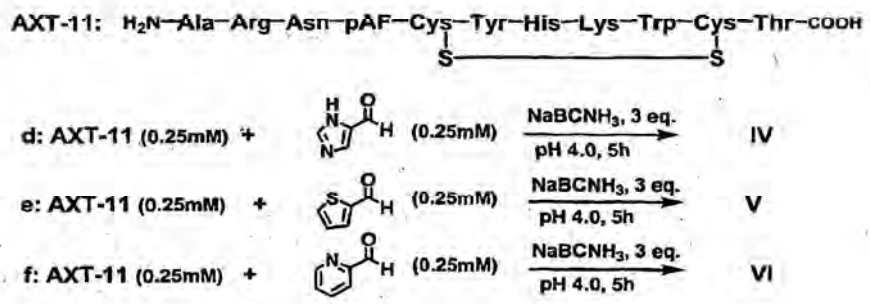


Figura 29

Alquilaciones reductoras de peptido AXT-11 con varios aldehidos heteroaromaticos (d-g)



pAF = L-para-Amino-fenilalanina

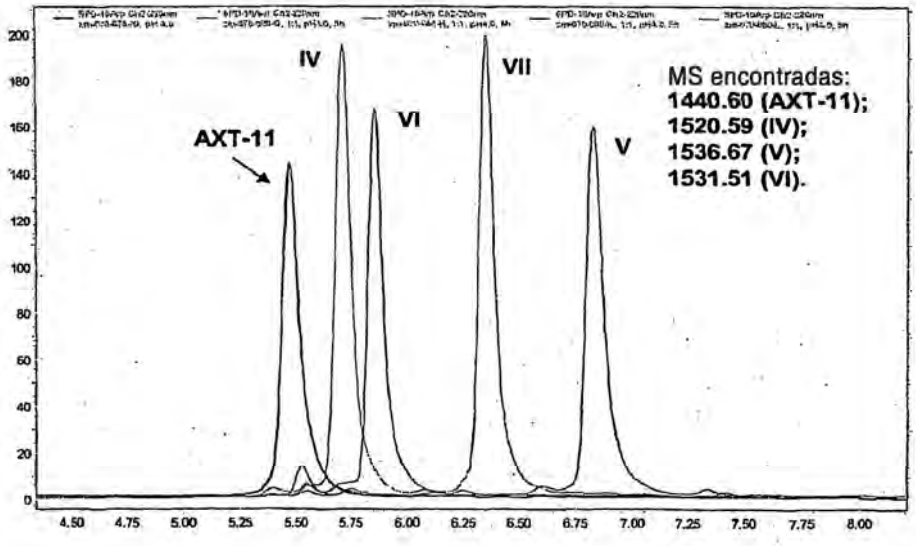
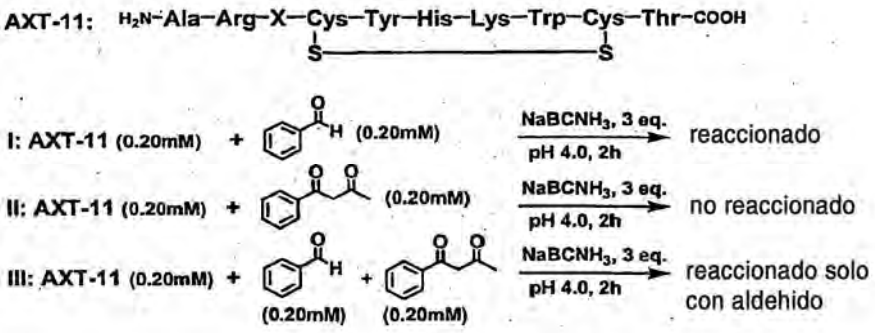


Figura 30

Reacciones competiticas de peptido AXT-11 con aldehido y 1,3-dicetona
(I-III)



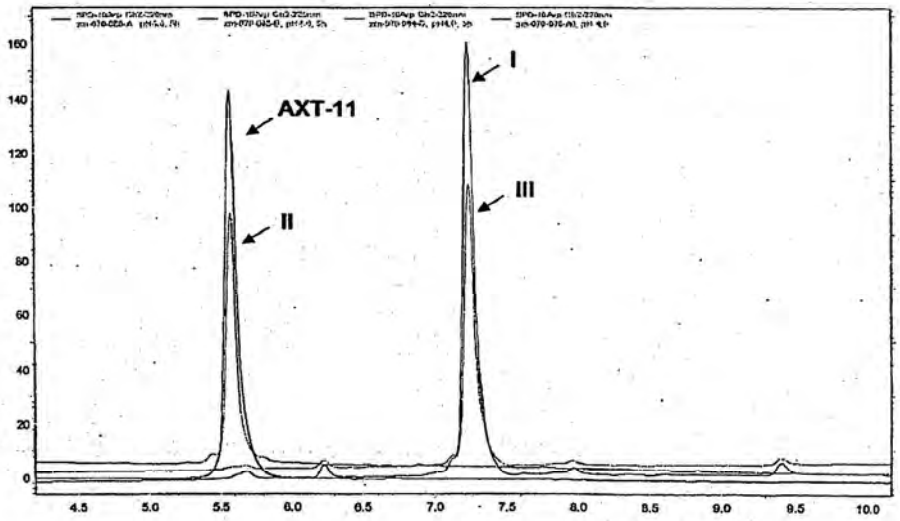
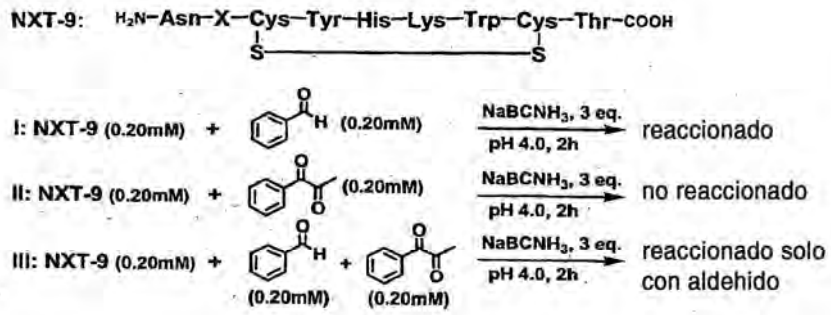


Figura 31

Reacciones competitivas con peptido NXT-9 con aldehido y 1,2-dicetona (I-III)



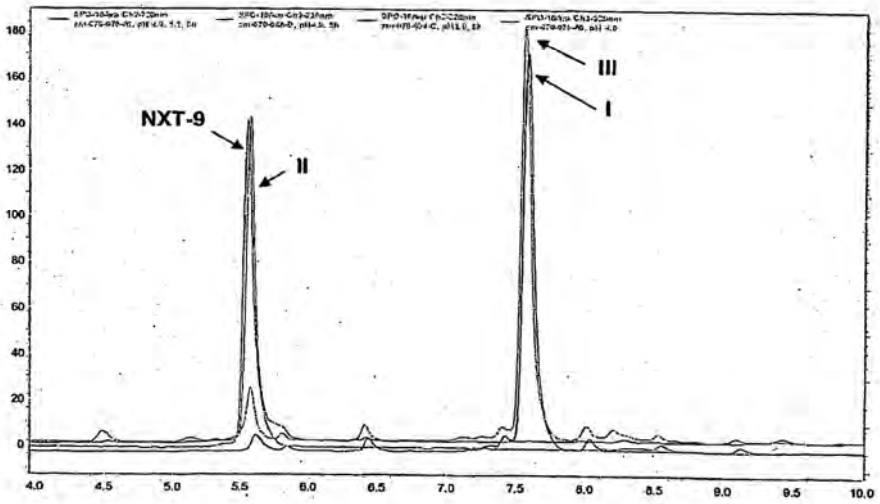
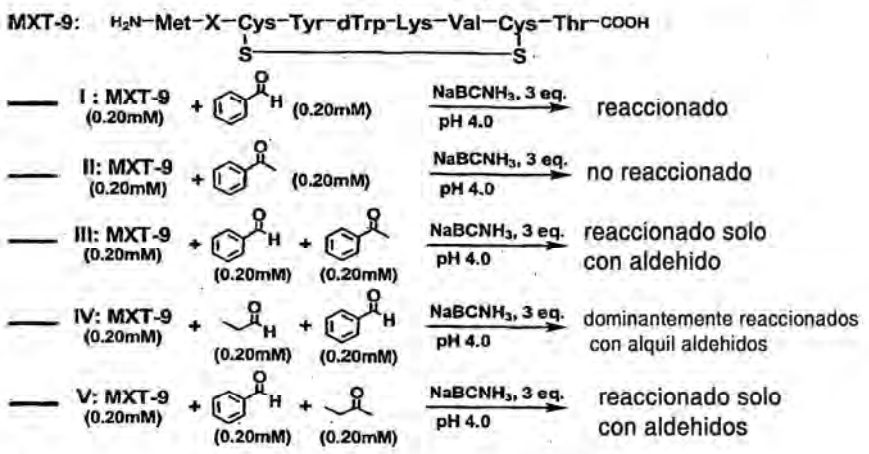


Figura 32

Reacciones competitivas de peptido MXT-9 con diferentes aldehidos y cetonas (I-V)



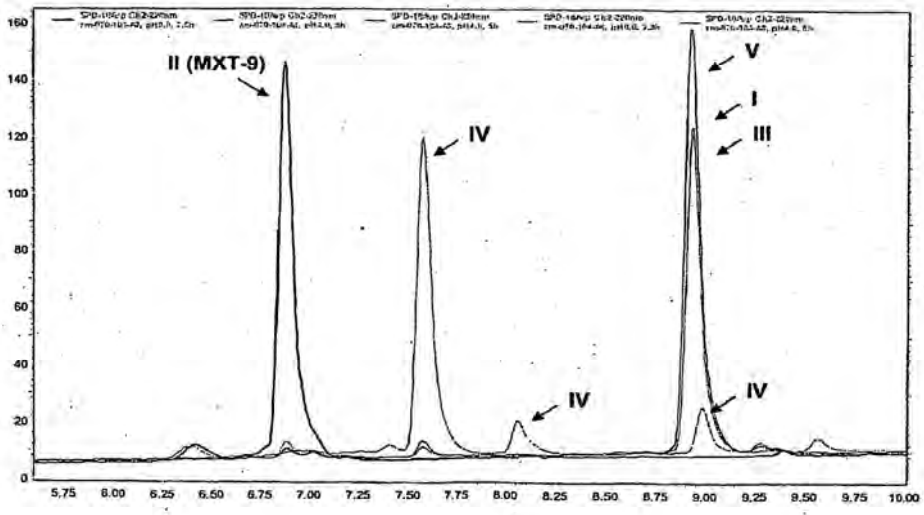
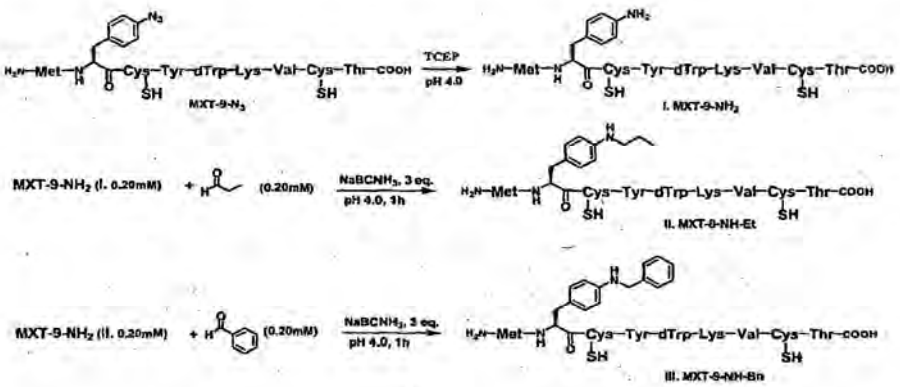


Figura 33

Reducción de peptido MXT-9-N3 a MXT-9NH2 (I), seguida por alquilaciones reductoras con propionaldehido (II) y benzaldehido (III)



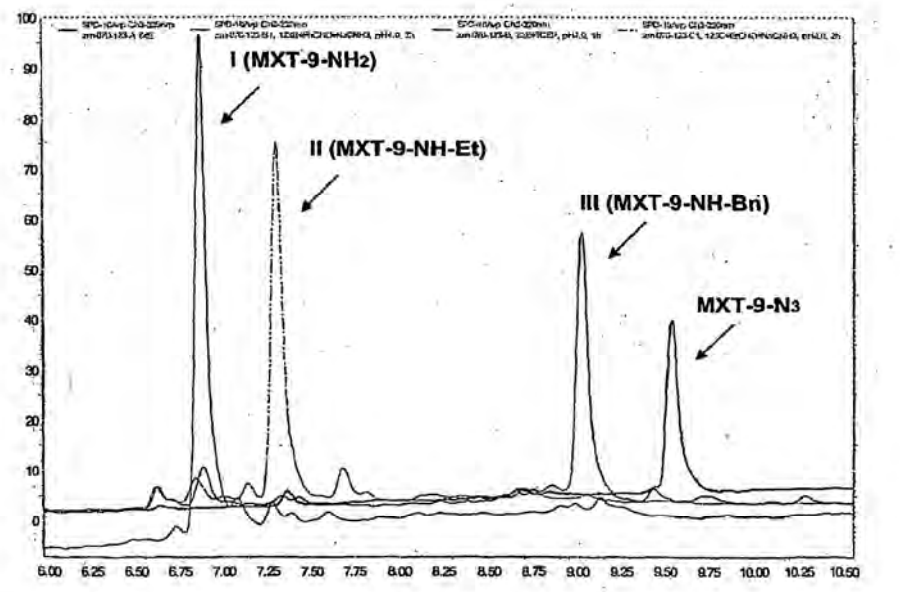


Figura 34

PEGilaciones de peptidos

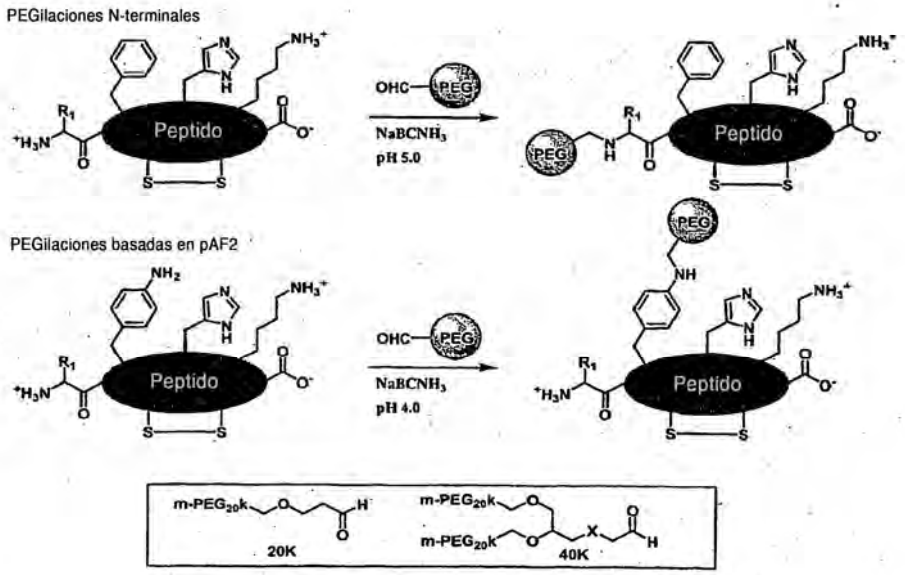


Figura 35

Aldehidos PEG usados para PEGilaciones

PEG aldehidos lineales:

PEG aldehidos ramificados derivados del glicerol:

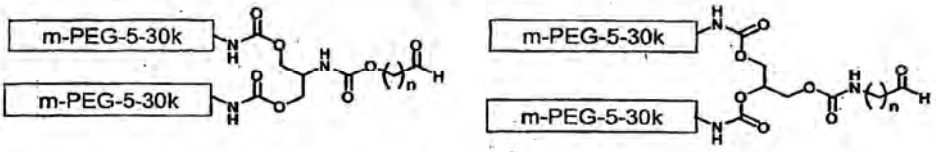


Figura 36

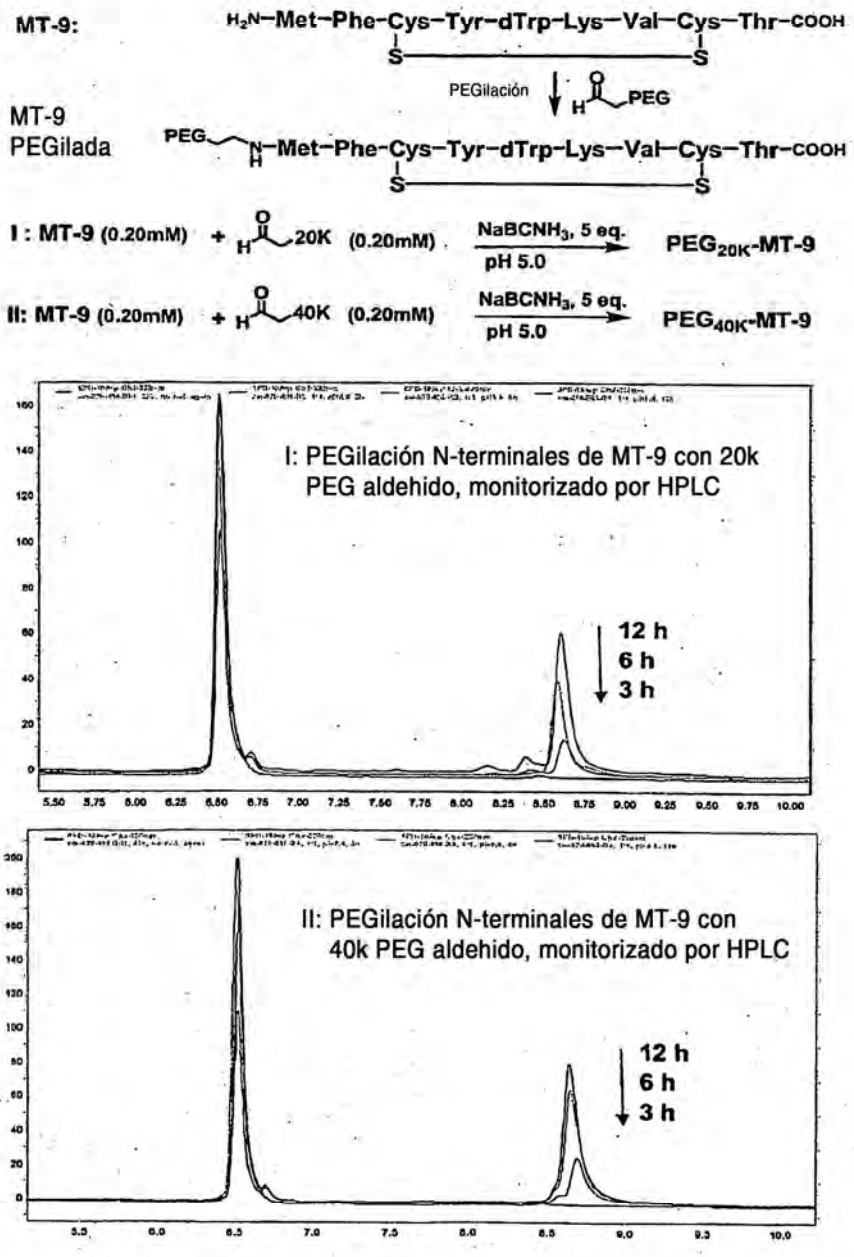
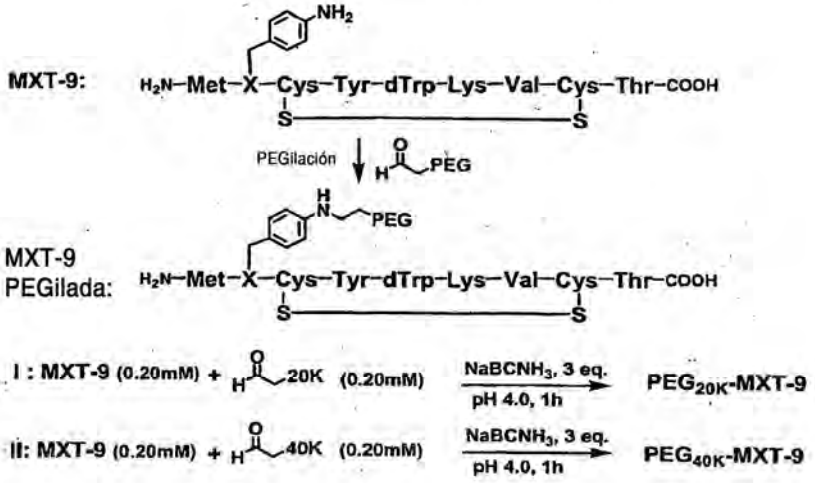

MT-9:
H2N–Met–Phe–Cys–Tyr–dTrp–Lys–Val–Cys–Thr–COOH
PEGilación
PEG
MT-9
PEGilada
PEG
N–Met–Phe–Cys–Tyr–dTrp–Lys–Val–Cys–Thr–COOH
I : MT-9 (0.20mM) + 20K (0.20mM)
NaBCNH3, 5 eq.
pH 5.0
PEG20K-MT-9
II: MT-9 (0.20mM) + 40K (0.20mM)
NaBCNH3, 5 eq.
pH 5.0
PEG40K-MT-9
I: PEGilación N-terminales de MT-9 con 20k PEG aldehido, monitorizado por HPLC
12 h
6 h
3 h
II: PEGilación N-terminales de MT-9 con 40k PEG aldehido, monitorizado por HPLC
12 h
6 h
3 h

Figura 37


MXT-9:
H2N–Met–X–Cys–Tyr–dTrp–Lys–Val–Cys–Thr–COOH
PEGilación
PEG
MXT-9
PEGilada:
H2N–Met–X–Cys–Tyr–dTrp–Lys–Val–Cys–Thr–COOH
I : MXT-9 (0.20mM) + 20K (0.20mM)
NaBCNH3, 3 eq.
pH 4.0, 1h
PEG20K-MXT-9
II: MXT-9 (0.20mM) + 40K (0.20mM)
NaBCNH3, 3 eq.
pH 4.0, 1h
PEG40K-MXT-9

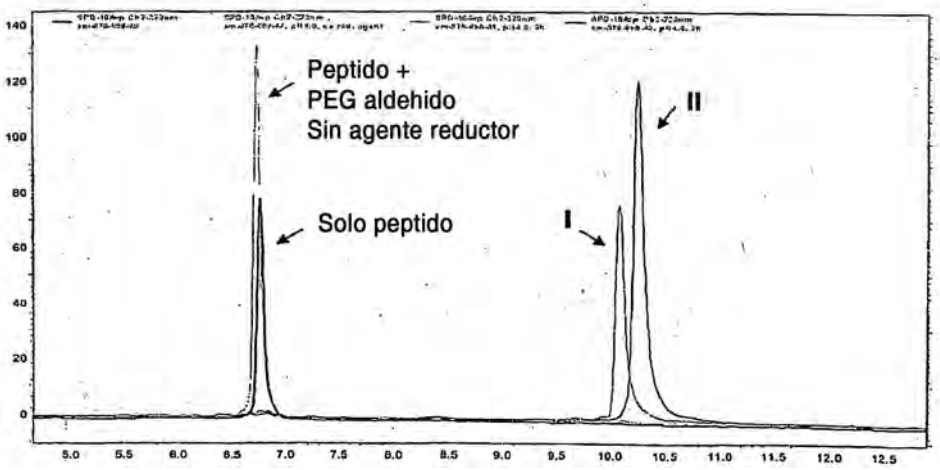

Peptido +
PEG aldehido
Sin agente reductor
Solo peptido
I
II
140
120
100
80
60
40
20
0
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0
9.5
10.0
10.5
11.0
11.5
12.0
12.5

Figura 38

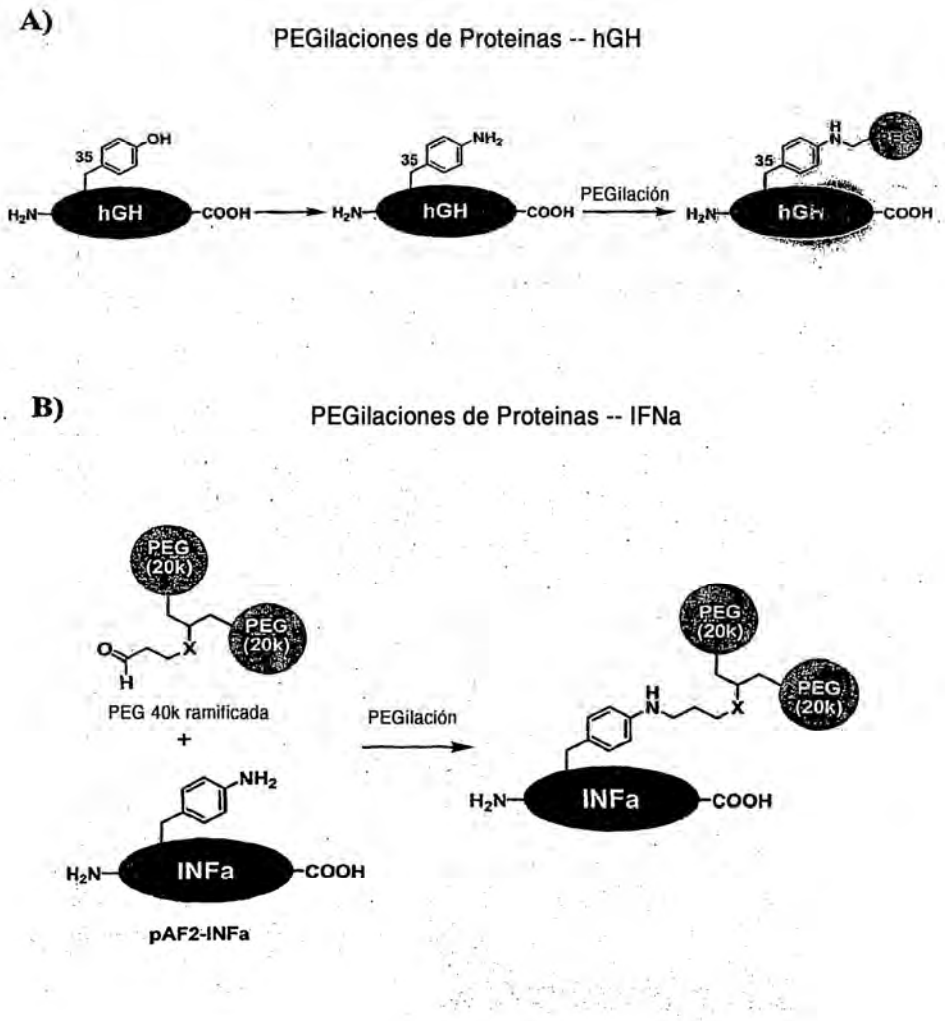

A)
PEGilaciones de Proteinas -- hGH
35
OH
hGH
COOH
PEGilación
B)
PEGilaciones de Proteinas -- IFNa
PEG (20k)
PEG 40k ramificada
+
INFa
pAF2-INFa
PEGilación

Figura 39

hGH Y35pAF2 PEGilación

Ratio PEG a proteina es 1:1
$NaCHBH_3$ es 5:1
pH = 4.0.
hGH esta a 350uM

1. H6-hGH Y35pAF2
2. H6-hGH Y35pAF2 + 20K PEG, 4ºC
3. H6-hGH Y35pAF2 + 30K PEG, 4ºC
4. H6-hGH Y35pAF2 + 40K, 4ºC
5. H6-hGH Y35pAF2 + 20K PEG, RT
6. H6-hGH Y35pAF2 + 30K PEG, RT
7. H6-hGH Y35pAF2 + 40K, RT
8. Met-hGH Y35pAF1
9. Met-hGH Y35pAF1 + 40K, 4ºC
10. Met-hGH Y35pAF1 + 40K, 4ºC
11. H6-hGH Y35pAF2
12. H6-hGH Y35pAF2 + 40K, 4ºC

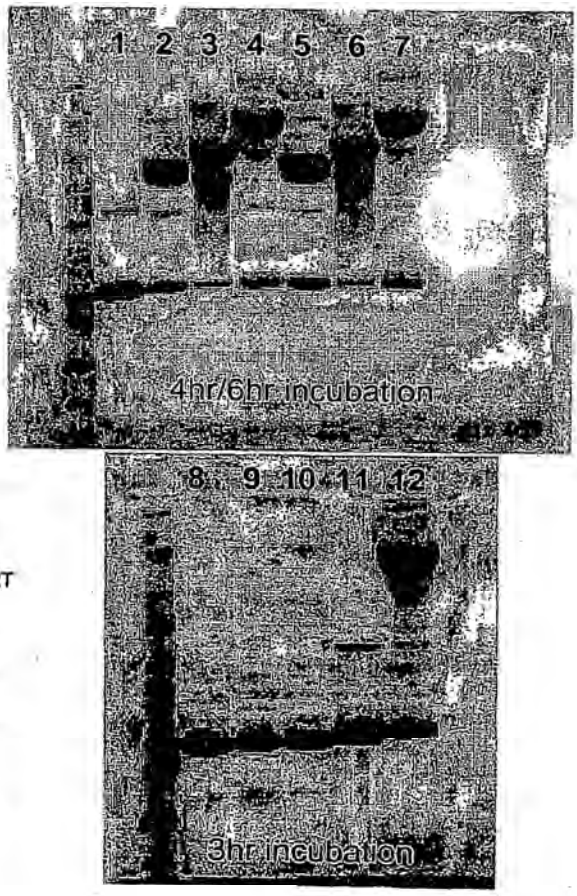


Figura 40